

ΔΙΑΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ (Δ.Π.Μ.Σ.) «ΔΙΚΑΙΟ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗ»

ΤΜΗΜΑΤΟΣ ΕΦΑΡΜΟΣΜΕΝΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΜΑΚΕΔΟΝΙΑΣ ΚΑΙ ΤΜΗΜΑΤΟΣ ΝΟΜΙΚΗΣ
ΔΗΜΟΚΡΙΤΕΙΟΥ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΘΡΑΚΗΣ



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**«ΠΡΟΓΝΩΣΤΙΚΗ ΑΝΑΛΥΣΗ ΜΑΖΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΥΓΕΙΑΣ
ΤΩΝ ΕΝΟΠΛΩΝ ΔΥΝΑΜΕΩΝ, ΣΤΗΝ ΕΠΟΧΗ ΤΟΥ
ΓΕΝΙΚΟΥ ΚΑΝΟΝΙΣΜΟΥ ΠΡΟΣΤΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ (ΓΚΠΔ)»**

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:

ΚΩΝΣΤΑΝΤΙΝΟΣ ΨΑΝΝΗΣ, kpsannis@uom.edu.gr

ΜΕΤΑΠΤΥΧΙΑΚΟΣ ΦΟΙΤΗΤΗΣ:

ΓΡΗΓΟΡΙΟΣ ΓΚΟΥΤΖΕΛΗΣ, mli19011@uom.edu.gr

ΦΕΒΡΟΥΑΡΙΟΣ 2021

ΠΕΡΙΕΧΟΜΕΝΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Σελίδα

<i>Ευχαριστίες</i>	4
<i>Βεβαίωση Εκπόνησης Διπλωματικής Εργασίας</i>	5
<i>1. Περίληψη</i>	6
<i>2. Abstract</i>	6
<i>3. Εισαγωγή</i>	7
<i>4. Βιβλιογραφική Επισκόπηση</i>	8
<i>5. Μεθοδολογία</i>	10
<i>6. Σύγκλιση Συναφών Τεχνολογιών - Δυνατοτήτων – Εργαλείων</i>	11
<i>6.1 Μαζικά Δεδομένα Υγείας (Health Big Data)</i>	12
<i>6.2 Ατομικός Ηλεκτρονικός Φάκελος Υγείας (ΑΗΦΥ)</i>	17
<i>6.3 Συστήματα Υποστήριξης Κλινικών Αποφάσεων (CDSS)</i>	19
<i>6.4 Τεχνητή Νοημοσύνη (Artificial Intelligence)</i>	20
<i>6.5 Μηχανική Μάθηση (Machine Learning)</i>	22
<i>6.5.1 Επιτηρούμενη Μηχανική Μάθηση (Supervised Machine Learning)</i>	27
<i>6.5.2 Μη Επιτηρούμενη Μηχανική Μάθηση (Unsupervised Machine Learning)</i>	29
<i>6.5.3 Ενισχυτική Μηχανική Μάθηση (Reinforcement Machine Learning)</i>	31
<i>6.5.4 Αυτοματοποιημένη Μηχανική Μάθηση (Automated Machine Learning)</i>	32
<i>6.6 Υπολογιστική Νέφους και Άκρου (Cloud and Edge Computing)</i>	33
<i>6.7 Διαδίκτυο των Πραγμάτων Υγείας (IoHT ή Internet of Healthcare Things)</i>	35
<i>6.8 Δίκτυο Κινητής Τηλεφωνίας 5G</i>	39
<i>6.9 Εφαρμογές Κινητής Υγείας (mHealth)</i>	40
<i>7. Ανάλυση Δεδομένων</i>	42
<i>7.1 Προγνωστική Ανάλυση (Predictive Analytics)</i>	45
<i>7.2 Προδιαγραφική Ανάλυση (Prescriptive Analytics)</i>	47
<i>7.3 Ιατρική Ακριβείας (Precision Medicine)</i>	48
<i>7.4 Μεταγραφική Ιατρική (Translational Medicine)</i>	51

<i>8. Ασφάλεια Δεδομένων Υγείας - Διατάξεις και Περιορισμοί του ΓΚΠΔ.....</i>	<i>52</i>
<i>8.1 Ασφάλεια Μαζικών Δεδομένων Υγείας στο Νέφος.....</i>	<i>54</i>
<i>8.2 Ασφάλεια Δεδομένων στο Διαδίκτυο των Πραγμάτων Υγείας.....</i>	<i>55</i>
<i>8.3 Ασφάλεια Μαζικών Δεδομένων Υγείας και Τεχνητή Νοημοσύνη.....</i>	<i>56</i>
<i>8.4 Εθνική Άμυνα – Εθνική Ασφάλεια</i>	<i>56</i>
<i>8.5 Δεδομένα Υγείας και Επιστημονική Έρευνα.....</i>	<i>59</i>
<i>8.6 Ανωνομοποίηση και Ψευδωνυμοποίηση Προσωπικών Δεδομένων.....</i>	<i>60</i>
<i>8.7 Συγκατάθεση και Εναντίωση των Υποκειμένων στην Προγνωστική Ανάλυση.....</i>	<i>64</i>
<i>8.8 Αξία Προγνωστικής Ανάλυσης Μαζικών Δεδομένων Υγείας για τα στελέχη των ΕΛ.....</i>	<i>66</i>
<i>9. Ανάπτυξη και Δοκιμή Μοντέλων Προγνωστικής Ανάλυσης.....</i>	<i>67</i>
<i>9.1 Συντελεστές Μοντελοποίησης.....</i>	<i>67</i>
<i>9.1.1 Δεδομένα Υγείας Δοκιμών.....</i>	<i>67</i>
<i>9.1.2 Πλατφόρμα Μηχανικής Μάθησης.....</i>	<i>68</i>
<i>9.2 Σενάρια Προγνωστικής Ανάλυσης – Αποτελέσματα.....</i>	<i>69</i>
<i>9.2.1 Πρόγνωση Κατάστασης Διαβητικός ή Όχι.....</i>	<i>69</i>
<i>9.2.2 Πρόγνωση Τιμής Γλυκοζυλιωμένης Αιμοσφαιρίνης.....</i>	<i>70</i>
<i>9.2.3 Πρόγνωση Κατάστασης Διαβήτη (4 Κατηγορίες).....</i>	<i>71</i>
<i>9.2.4 Πρόγνωση Κατάστασης Μεταβολικό Σύνδρομο ή Όχι.....</i>	<i>72</i>
<i>9.2.5 Πρόγνωση Κατάστασης Καρδιοπαθής ή Όχι.....</i>	<i>73</i>
<i>9.2.6 Πρόγνωση Κατάστασης Καρκίνος Πνεύμονα ή Όχι.....</i>	<i>74</i>
<i>10. Επίλογος</i>	<i>75</i>
<i>10.1 Συμπεράσματα.....</i>	<i>75</i>
<i>10.2 Προτάσεις για Μελλοντική Έρευνα.....</i>	<i>80</i>
<i>11. Πίνακας Συγκριτικής Ανάλυσης Βιβλιογραφικών Αναφορών.....</i>	<i>I</i>
<i>12. Βιβλιογραφία.....</i>	<i>i</i>
<i>12.1 Άρθρα – Συγγράμματα.....</i>	<i>i</i>
<i>12.2 Διαδίκτυο.....</i>	<i>iv</i>

Θερμές ευχαριστίες οφείλω στον Καθηγητή μου κ. Ψάννη Κωνσταντίνο, για την πολύτιμη καθοδήγησή του, κατά την εκπόνηση αυτής της εργασίας.

Επίσης, ευχαριστώ τη σύζυγό μου Δήμητρα για την υπομονή και στήριξή της, καθώς και τις κόρες μου Ζωή και Ελευθερία, για τη βοήθειά τους.

Η εργασία είναι αφιερωμένη στη μνήμη του πατρός μου Κωνσταντίνου.

Βεβαίωση Εκπόνησης Διπλωματικής Εργασίας

Δηλώνω υπεύθυνα ότι η συγκεκριμένη Διπλωματική Εργασία, για τη λήψη του Μεταπτυχιακού Τίτλου σπουδών, του ΔΠΜΣ «Δίκαιο και Πληροφορική», του Τμήματος Εφαρμοσμένης Πληροφορικής του Πανεπιστημίου Μακεδονίας και του Τμήματος Νομικής του Δημοκρίτειου Πανεπιστημίου Θράκης, έχει συγγραφεί από εμένα προσωπικά και δεν έχει υποβληθεί ούτε έχει εγκριθεί στο πλαίσιο κάποιου άλλου Μεταπτυχιακού ή Προπτυχιακού τίτλου σπουδών, στην Ελλάδα ή στο εξωτερικό.

Η εργασία αυτή έχοντας εκπονηθεί από εμένα, αντιπροσωπεύει αποκλειστικά τις προσωπικές μου απόψεις επί του θέματος. Κατά τη συγγραφή της ακολούθησα την πρόπυσα ακαδημαϊκή δεοντολογία και απέφυγα οποιαδήποτε ενέργεια, που συνιστά παράπτωμα λογοκλοπής.

Οι βιβλιογραφικές πηγές στις οποίες ανέτρεξα για την εκπόνηση της συγκεκριμένης διπλωματικής εργασίας, αναφέρονται στο σύνολό τους, με πλήρεις αναφορές στους συγγραφείς τους, συμπεριλαμβανομένων και των πηγών που αντλήθηκαν από το Διαδίκτυο.

Τα σύνολα δεδομένων υγείας που χρησιμοποιήθηκαν για τις δοκιμές Προγνωστικής Ανάλυσης, με το εργαλείο Μηχανικής Μάθησης Knowi, αναζητήθηκαν και αντλήθηκαν από τον ιστότοπο data.world, αποτελούν δεδομένα δοκιμών, πλήρως ανωνυμοποιημένα και σε καμία περίπτωση δεν μπορούν να οδηγήσουν, άμεσα ή έμμεσα, σε ταυτοποίηση φυσικών προσώπων.

Ο Δηλών

Γρηγόριος Κ. Γκουτζέλης
(Α.Μ.: MLI19011)

«Το διοικείν εστί προβλέπειν», Αλκιβιάδης (450-404 π.Χ.)

1. Περίληψη

Τα Μαζικά Δεδομένα Υγείας των στελεχών των Ελληνικών Ενόπλων Δυνάμεων, προερχόμενα από το ιατρικό ιστορικό των υποχρεωτικών τους εξετάσεων, συνδυαζόμενα με περιβαλλοντικά καθώς και με δεδομένα γονιδιώματος, μπορούν μέσα από ειδική προγνωστική ανάλυση, να οδηγήσουν σε ασφαλείς προγνώσεις ασθενειών, με μεθόδους Μηχανικής Μάθησης. Αυτή η δυνατότητα είναι πολύ σημαντική, αφού η καλή κατάσταση της υγείας τους, αποτελεί αποφασιστικής σημασίας παράγοντα μαχητικής ισχύος. Νέες τάσεις στην επιστήμη των δεδομένων υγείας, συνδυαζόμενες με νέες σύγχρονες ιατρικές δυνατότητες, μπορούν να αποτελέσουν ανεκτίμητα εργαλεία επιλογής της ορθής ιατρικής θεραπείας, εξατομικευμένης για κάθε περίπτωση ασθενή. Τέτοιες τάσεις και δυνατότητες είναι η Προγνωστική και η Προδιαγραφική Ανάλυση, η Ιατρική Ακριβείας και η Μεταγραφική Ιατρική. Η γνώση της κατάστασης της υγείας του στρατιωτικού προσωπικού, μπορεί να αποτελέσει ιδιαίτερα πολύτιμη πληροφορία για τον εκάστοτε επιχειρησιακό διοικητή, για την ανάληψη, συνέχιση ή μεταβολή των πολεμικών επιχειρήσεων κι αποστολών που του ανατίθενται. Τα δεδομένα υγείας πρέπει να τυγχάνουν κατάλληλου χειρισμού και προστασίας, σύμφωνα με όσα προβλέπει ο Γενικός Κανονισμός Προστασίας Δεδομένων (ΓΚΠΔ). Τεχνικές όπως η κρυπτογράφηση, η ανωνυμοποίηση και η ψευδωνυμοποίηση, προσφέρουν μεγαλύτερη προστασία των προσωπικών δεδομένων υγείας. Η παρούσα έρευνα λαμβάνει υπόψη τις νέες αυτές τάσεις, πάντοτε μέσα στα σαφή όρια που θέτει ο ΓΚΠΔ, ενώ παράλληλα μελετά τη χρήση και σύγκλιση σύγχρονων τεχνολογιών, εργαλείων και δυνατοτήτων, οι οποίες ενισχύουν τις τάσεις αυτές. Εκτιμάται ότι μπορεί να συμβάλει στη διαμόρφωση προτάσεων και ιδεών προς την κατεύθυνση του εκσυγχρονισμού του στρατιωτικού συστήματος υγείας. Η σημαντικότερη συνεισφορά της εστιάζεται στην έγκαιρη διάγνωση και πρόληψη χρόνιων σοβαρών ασθενειών, μέσω της Ανάλυσης Δεδομένων, επ' ωφελεία τόσο των στελεχών, όσο και της Εθνικής Άμυνας της χώρας μας.

Λέξεις κλειδιά: *Προγνωστική Ανάλυση, Προδιαγραφική Ανάλυση, Μαζικά Δεδομένα Υγείας, Προσωπικά Δεδομένα Υγείας, Ιατρική Ακριβείας, Μεταγραφική Ιατρική, Ανάλυση Δεδομένων.*

2. Abstract

The Big Health Data of the Greek Armed Forces executives, derived from their medical history of their mandatory exams, combined with environmental as well as genome data, can lead, through special predictive analysis, to safe disease prediction, using Machine Learning methods. This capability is very significant, since their good health is a crucial factor of combat power. New trends in medical data science combined with new modern medical possibilities can be invaluable tools for choosing the right medical treatment, personalized for each specific patient case. Such trends are the Predictive and the Prescriptive Analytics, the Precision and Translational Medicine. The situational awareness of the personnel health status can be particularly valuable information for the respective operational commander, in terms of undertaking, sustainment or alteration of the military operations and the undertaken missions. Health data must be properly handled and protected in accordance with the General Data Protection Regulation (GDPR). Techniques such as cryptography, anonymization and pseudonymization offer greater protection of the personal health data. The present research takes into account these new trends, always within the clear limits set by the GDPR, while in parallel studies the use and convergence of modern technologies, tools and capabilities, which are being utilized to facilitate these trends. It is estimated that it may contribute to the formation of recommendations and ideas towards the modernization of the military health system. It's most important contribution is focused on the early diagnosis and prevention of chronic serious diseases, through data analytics, for the benefit of both the executives and the National Defense of our country.

Keywords: *Predictive Analysis, Prescriptive Analytics, Big Health Data, Personal Health Data, Precision Medicine, Transcription Medicine, Data Analytics.*

3. Εισαγωγή

Τα Μαζικά Δεδομένα (Big Data) σχετίζονται με τους διαθέσιμους πόρους σε δεδομένο χρονικό σημείο και χαρακτηρίζονται γενικά από όγκο, ταχύτητα και ποικιλομορφία. Πηγές Μαζικών Δεδομένων Υγείας (ΜΔΥ) μπορούν να αποτελέσουν, μεταξύ άλλων, τα δεδομένα των ιατρικών φακέλων των ασθενών και οι υπολογιστικές αναλύσεις των δεδομένων αυτών. Σύνθετοι αλγόριθμοι, που υποστηρίζονται από υψηλής απόδοσης υπολογιστικά συστήματα, επιτρέπουν την μετατροπή αυτών των μεγάλων συνόλων δεδομένων σε γνώση. Για τα στελέχη των Ενόπλων Δυνάμεων (ΕΔ) της χώρας μας, τα ΜΔΥ τους αρχίζουν να συλλέγονται από τη στιγμή της κατάταξής τους, αφού η συνεχής παρακολούθηση της υγείας τους και η διατήρησή της σε άριστο επίπεδο, αποτελεί κύριο μέλημα της στρατιωτικής ηγεσίας. Συνδυαζόμενα με περιβαλλοντικά καθώς και με δεδομένα γονιδιώματος, μπορούν να οδηγήσουν σε ασφαλείς προγνώσεις ασθενειών.

Πρόγνωση και πρόληψη είναι έννοιες αλληλένδετες, ειδικά όταν πρόκειται για δραστηριότητες που σχετίζονται με τα δεδομένα υγείας. Στον ευαίσθητο αυτό χώρο, η πρόγνωση μέσω της ανάλυσης ΜΔΥ, μπορεί να προλάβει ή και να αποτρέψει, την εκδήλωση σοβαρών χρόνιων ασθενειών. Μια τέτοια δυνατότητα μπορεί να αξιοποιηθεί περαιτέρω, όταν υποκείμενα των εν λόγω δεδομένων είναι στελέχη των ΕΔ, αφού η καλή κατάσταση της υγείας τους είναι προαπαιτούμενο της πολεμικής τους ετοιμότητας, αποτελεί δε αποφασιστικής σημασίας παράγοντα μαχητικής ισχύος. Η γνώση της κατάστασης της υγείας του προσωπικού, μπορεί να αποτελέσει ιδιαίτερα πολύτιμη πληροφορία για τον εκάστοτε επιχειρησιακό διοικητή, για την ανάληψη, συνέχιση ή μεταβολή των πολεμικών επιχειρήσεων ή των ανατιθέμενων αποστολών.

Τα δεδομένα υγείας, προσωπικά και ταυτόχρονα ευαίσθητα, αφενός μπορούν και πρέπει να χρησιμοποιούνται προς όφελος των υποκειμένων, αφετέρου πρέπει να τυγχάνουν κατάλληλου χειρισμού και προστασίας, σύμφωνα με όσα προβλέπει ο Γενικός Κανονισμός Προστασίας Δεδομένων (ΓΚΠΔ) και ο αντίστοιχος Ν 4624/2019, που ψηφίστηκε με σκοπό τη λήψη μέτρων εφαρμογής του, για την προστασία των φυσικών προσώπων, έναντι της επεξεργασίας των δεδομένων προσωπικού χαρακτήρα και την ελεύθερη κυκλοφορία τους. Μολονότι υπάρχουν εξαιρέσεις στην προστασία των δεδομένων, όταν ιδίως προτάσσεται η εθνική ασφάλεια και το δημόσιο συμφέρον ή εφόσον είναι απαραίτητη η επεξεργασία, για σκοπούς προληπτικής ή επαγγελματικής ιατρικής, εκτίμησης της ικανότητας προς εργασία κ.λπ., τεχνικές όπως αυτές της ανωνυμοποίησης και της ψευδωνυμοποίησης των δεδομένων, προσφέρουν μεγαλύτερη προστασία στα υποκείμενα.

Για την πραγματοποίηση της παρούσας έρευνας χρησιμοποιήθηκε κυρίως η τεχνική αναζήτησης βιβλιογραφίας με τη μέθοδο «pearl-growing» (ανάπτυξης μαργαριταριού), που χρησιμοποιείται συχνά για αναζήτηση και χρήση βιβλιογραφικών αναφορών, σχετικών με το προς μελέτη θέμα. Στόχος της είναι να οδηγήσει στην εξαγωγή χρήσιμων συμπερασμάτων, ως προς τις δυνατότητες της Προγνωστικής Ανάλυσης των ΜΔΥ, για την πρόγνωση και την πρόληψη σοβαρών ασθενειών των στελεχών των ΕΔ. Η ανάλυση αυτή λαμβάνει αφενός υπόψη τα όρια που θέτει ο ΓΚΠΔ και αφετέρου, ενδυναμώνεται από τη χρήση και σύγκλιση σύγχρονων τεχνολογιών, όπως της Τεχνητής Νοημοσύνης, της Μηχανικής Μάθησης, της Υπολογιστικής Νέφους και του Διαδικτύου των Πραγμάτων Υγείας (IoHT), με την υποστήριξη του δικτύου κινητής τηλεφωνίας 5ης γενιάς (5G). Εξάλλου, η Προδιαγραφική Ανάλυση, ως συμπλήρωμα της Προγνωστικής, συμβάλει στην επίλυση προβλημάτων βελτιστοποίησης στο χώρο της υγείας, με τη χρήση δυνατοτήτων παρεχομένων από τις επιστήμες της στατιστικής και των μαθηματικών.

Οι ΕΔ της χώρας μας βρίσκονται σε μια διαρκή προσπάθεια αναβάθμισης των προσφερόμενων υγειονομικών υπηρεσιών τους και συχνά έρχονται αντιμέτωποι με νέες προκλήσεις, όπως αυτή της έγκαιρης πρόγνωσης και πρόληψης χρόνιων ασθενειών. Ως εκ τούτου, η εν λόγω έρευνα εκτιμάται ότι θα συμβάλει στη διαμόρφωση προτάσεων και ιδεών προς την κατεύθυνση του εκσυγχρονισμού του στρατιωτικού συστήματος υγείας, γενικά, καθώς και ειδικότερα, προς την κατεύθυνση της πρόληψης χρόνιων σοβαρών ασθενειών, επ' ωφελεία τόσο των στελεχών, όσο και της Εθνικής μας Άμυνας.

4. Βιβλιογραφική Επισκόπηση

Το γενικότερο πρόβλημα, που η παρούσα εργασία καλείται να λύσει είναι η Προγνωστική Ανάλυση των Μαζικών Δεδομένων Υγείας των στελεχών των Ενόπλων Δυνάμεων (ΕΔ) και μάλιστα υπό το περιοριστικό νομοθετικό πλαίσιο που θέτει ο Γενικός Κανονισμός Προστασίας των Δεδομένων (ΓΚΠΔ). Έτσι, γίνεται σαφές ότι στην εργασία κυριαρχούν δύο θεμελιώδεις έννοιες: η έννοια της Προγνωστικής Ανάλυσης και η έννοια των Μαζικών Δεδομένων Υγείας.

Η Προγνωστική Ανάλυση (Predictive Analytics) είναι μια κατηγορία Ανάλυσης Δεδομένων (Data Analytics), που στοχεύει στη διενέργεια προβλέψεων για μελλοντικά αποτελέσματα, βάσει ιστορικών δεδομένων και τεχνικών ανάλυσης, όπως η στατιστική μοντελοποίηση (Statistical Modelling) και η Μηχανική Μάθηση (Machine Learning). Η επιστήμη της Προγνωστικής Ανάλυσης μπορεί να παράγει μελλοντικές προβλέψεις με σημαντικό βαθμό ακρίβειας. Με τη βοήθεια εξελιγμένων εργαλείων και μοντέλων πρόβλεψης, κάθε οργανισμός μπορεί πλέον να χρησιμοποιεί παλαιότερα και πιο πρόσφατα δεδομένα, για αξιόπιστες προβλέψεις τάσεων και συμπεριφορών στο μέλλον (CIO, 2019).

Μαζικά Δεδομένα είναι εκείνα τα δεδομένα «των οποίων η κλίμακα, η ποικιλομορφία και η πολυπλοκότητα απαιτούν νέα αρχιτεκτονική, τεχνικές, αλγόριθμους και αναλυτικά στοιχεία, για τη διαχείριση και εξαγωγή αξίας και κρυμμένης γνώσης, μέσα από αυτά» (Harper, 2014).

Η Προγνωστική Ανάλυση σε Μαζικά Δεδομένα, ασφαλώς δεν είναι μια καινούργια ιδέα. Χρησιμοποιείται εδώ και αρκετές δεκαετίες, καθώς με τη βοήθεια μεθόδων Τεχνητής Νοημοσύνης και Μηχανικής Μάθησης μπορούν να πραγματοποιηθούν ασφαλείς προγνώσεις για καταστάσεις ή αριθμητικές τιμές, των οποίων η έγκαιρη και αξιόπιστη πρόγνωση ενδεχομένως να είναι πολύτιμη.

Στην ιδέα αυτή ακριβώς της έγκαιρης, αξιόπιστης και πολύτιμης πρόγνωσης βασίστηκε και η παρούσα έρευνα, εστιάζοντας στα δεδομένα υγείας, προσωπικά και συνάμα ευαίσθητα, του προσωπικού των ΕΔ, αφού εντόπισε αντίστοιχο ερευνητικό κενό κατά την αναζήτηση της βιβλιογραφίας, γεγονός που καθιστά το προς διερεύνηση ζήτημα, σημαντικό και σχετικά πρωτότυπο. Εκτιμήθηκε συναφώς ότι η έγκαιρη γνώση, μέσω της πρόγνωσης, της κατάστασης της υγείας, είναι μάλλον προς όφελος, πρωτίστως του υποκειμένου των δεδομένων υγείας και δευτερευόντως των ΕΔ, στις οποίες υπηρετεί. Κι αυτό γιατί η καλή κατάσταση της υγείας των υποκειμένων (στελεχών των ΕΔ) είναι προαπαιτούμενο της πολεμικής τους ετοιμότητας και αποτελεί αποφασιστικής σημασίας παράγοντα μαχητικής ισχύος του στρατεύματος. Η γνώση της κατάστασης της υγείας του προσωπικού, μπορεί να αποτελέσει ιδιαίτερα πολύτιμη πληροφορία για τον εκάστοτε επιχειρησιακό διοικητή, για την ανάληψη, συνέχιση ή μεταβολή των πολεμικών επιχειρήσεων ή των ανατιθέμενων, σε αυτόν, αποστολών.

Επιπλέον, η ηγεσία των ΕΔ εκτιμάται ότι θα επιθυμούσε να έχει στη διάθεσή της αυτή την πολύτιμη γνώση, όχι μόνο για να μπορεί να αξιολογήσει ανά πάσα χρονική στιγμή το αξιόμαχο του προσωπικού της, αλλά και διότι υγιή στελέχη σημαίνει θετικό αντίκτυπο στην εθνική μας οικονομία, αφού από έρευνες έχει προκύψει ότι η καλή υγεία του πληθυσμού, έχει θετική και στατιστικά σημαντική επίδραση, στην οικονομική ανάπτυξη μιας χώρας (Bloom et al., 2004). Αν λοιπόν ο υπεύθυνος επεξεργασίας των Μαζικών Δεδομένων Υγείας, εν προκειμένω το Υπουργείο Εθνικής Άμυνας με τις Διευθύνσεις Υγειονομικού των Κλάδων, γνωρίζει στοιχεία της πρόγνωσης, που δείχνουν πού κυμαίνεται η κατάσταση της υγείας του προσωπικού του, ενδεχομένως θα επιδιώξει να το συμβουλευτεί εγκαίρως, να τροποποιήσει τρόπους εκπαίδευσης, διατροφής, άσκησης, φαρμακευτικής αγωγής κ.λπ. των υποκειμένων, ώστε να τα βοηθήσει να διατηρήσουν ή να ανακτήσουν την καλή τους υγεία.

Η εργασία αναπτύχθηκε περιγράφοντας αρχικά όλες τις συγκλίνουσες τεχνολογίες (όπως η Τεχνητή Νοημοσύνη, η Μηχανική Μάθηση, η Υπολογιστική Νέφους και Άκρου, το Διαδίκτυο των Πραγμάτων Υγείας, το Δίκτυο Κινητής Τηλεφωνίας 5G), οι οποίες συνδράμουν στην υλοποίηση της Προγνωστικής Ανάλυσης, στη σύγχρονη εποχή.

Στη συνέχεια αναπτύχθηκε η ενότητα της Ανάλυσης Δεδομένων, καταδεικνύοντας τόσο τη σημασία της όσο και τις σύγχρονες επιστημονικές τάσεις στην Ιατρική και στην Ανάλυση, με υποενότητες σχετικές με την Προγνωστική Ανάλυση, την Προδιαγραφική Ανάλυση, την Ιατρική Ακριβείας και την Μεταγραφική Ιατρική.

Επειδή τα εν λόγω δεδομένα ανήκουν στις επονομαζόμενες «ειδικές κατηγορίες» κατά τον ΓΚΠΔ, αναλύθηκε στη συνέχεια το νομοθετικό περιβάλλον, στο οποίο η προτεινόμενη ιδέα, ως αντικείμενο έρευνας, καλείται να εφαρμοστεί. Έτσι εξετάστηκαν τα δικαιώματα των υποκειμένων και οι εκ του νόμου δυνατότητες του υπευθύνου επεξεργασίας, υπό το πρίσμα των συνταγματικών ελευθεριών, της διαφύλαξης των εθνικών συμφερόντων και της εθνικής ασφάλειας.

Η προτεινόμενη ιδέα, αποδείχθηκε στη συνέχεια, με πρακτικό και εφαρμόσιμο τρόπο, ότι μπορεί όντως και υπό προϋποθέσεις να υλοποιηθεί. Με τη βοήθεια ανωνυμοποιημένων δεδομένων υγείας, που αναζητήθηκαν και αντλήθηκαν από ανοικτές πηγές δεδομένων στο Διαδίκτυο (data.world) και εισήχθησαν σε εργαλείο Προγνωστικής Ανάλυσης (Knowi.com), επιχειρήθηκαν δοκιμές και προγνώσεις. Εξετάστηκαν έξι σενάρια Προγνωστικής Ανάλυσης (κατάστασης διαβητικός ή όχι, τιμής γλυκοζυλιωμένης αιμοσφαιρίνης, κατάστασης τεσσάρων κατηγοριών διαβήτη, κατάστασης πάσχων από μεταβολικό σύνδρομο ή όχι, κατάστασης καρδιοπαθής ή όχι, κατάστασης πάσχων από καρκίνο του πνεύμονα ή όχι) και πραγματοποιήθηκαν προγνώσεις, οι οποίες σημείωσαν πολύ ικανοποιητικά ποσοστά ακρίβειας. Έτσι καταδείχθηκε ότι αυτό που η εργασία προτείνει και ερευνά, μπορεί πραγματικά να επιτευχθεί, υπό την προϋπόθεση ότι υπάρχουν επαρκή και κατάλληλα σύνολα δεδομένων υγείας, που εισάγονται στα κατάλληλα εργαλεία Μηχανικής Μάθησης και υφίστανται ανάλυση με μεθόδους όπως η Ταξινόμηση και η Παλινδρόμηση.

Ως προς τη βιβλιογραφία, το κυρίως μέρος της εργασίας βασίστηκε σε 30 σύγχρονες βιβλιογραφικές πηγές, οι οποίες αντλήθηκαν από τις Διαδικτυακές πύλες «ScienceDirect», «ResearchGate» και «Frontiers in Medicine» και απετέλεσαν τόσο σημεία αναφοράς όσο και έμπνευσης, για τη σύνταξη μεγάλων τμημάτων της εργασίας. Δεν είναι αυτές οι μόνες πηγές που χρησιμοποιήθηκαν, είναι όμως αυτές που διαμόρφωσαν τον πυρήνα, για την ανάπτυξη των επιμέρους εννοιών, με σκοπό την ολοκληρωμένη και ολιστική κατανόηση των προς διερεύνηση ζητημάτων.

Τα κύρια άρθρα που προσδιορίστηκαν, με επίκεντρο την Ανάλυση των Μαζικών Δεδομένων Υγείας (Big Health Data Analytics), αξιολογήθηκαν αρχικά μεμονωμένα, ενώ καταγράφηκε μια συνοπτική περιγραφή του καθενός, με παράθεση των ερευνητικών στόχων, των κυριότερων ευρημάτων τους, των προβλημάτων που μελετήθηκαν και των αντίστοιχων λύσεών τους. Για αυτή τη συνοπτική περιγραφή χρησιμοποιήθηκαν οι περιλήψεις (abstracts) και τα συμπεράσματα κάθε άρθρου, ενώ η μελέτη των επιμέρους προβλημάτων και λύσεων, προέκυψε από έρευνα του κυρίως μέρους τους. Στη συνέχεια τα άρθρα κατηγοριοποιήθηκαν σε πέντε βασικές ομάδες, όπου το γενικότερο θέμα-τίτλος της κάθε ομάδας, εξετάστηκε από διάφορες οπτικές γωνίες και διενεργήθηκε συγκριτική, μεταξύ τους, ανάλυση. Οι βασικές ομάδες των άρθρων είναι: 1) Analytics, 2) Big Data, 3) Precision Medicine, 4) IoT και 5) Machine Learning.

Στο τέλος της εργασίας, και κατόπιν υπόδειξης και σταδιακής έγκρισης του επιβλέποντος καθηγητή, σχεδιάστηκε και τοποθετήθηκε ένας συγκεντρωτικός και συγκριτικός πίνακας κατανομής της βιβλιογραφίας κατά κατηγορία, όπου αναλύονται έξι άρθρα για καθεμιά, από τις παραπάνω αναφερόμενες, κατηγορίες. Η ανάλυση των 30 άρθρων περιλαμβάνει για καθένα από αυτά, τον ερευνητικό του σκοπό, την χρησιμοποιούμενη μεθοδολογία, τα κυριότερα ευρήματα που προέκυψαν από την έρευνα, τα προβλήματα που μελετήθηκαν, τις λύσεις που προτείνονται, καθώς και τα ανοικτά θέματα για μελλοντική διερεύνηση. Ο πίνακας αυτός εκτιμάται ότι θα αποτελέσει ένα χρήσιμο εργαλείο για την ερευνητική κοινότητα, ή για τον αναγνώστη αυτής της επισκόπησης.

5. Μεθοδολογία

Η μέθοδος που χρησιμοποιήθηκε στην παρούσα εργασία είναι η ερμηνευτική έρευνα, βασισμένη σε δεδομένα βιβλιογραφικής ανασκόπησης, τα οποία προέκυψαν και διασταυρώθηκαν από πρόσφατα έγκυρα δημοσιευμένα άρθρα, μέσω αναζήτησης με κατάλληλες λέξεις κλειδιά, από τις Διαδικτυακές πύλες «ScienceDirect», «ResearchGate» και «Frontiers in Medicine».

Για την ανεύρεση των κύριων βιβλιογραφικών αναφορών επιλέχθηκε η μέθοδος «pearl-growing» (ανάπτυξης μαργαριταριού), η οποία χρησιμοποιείται συχνά στην αναζήτηση και χρήση άρθρων και εργασιών, σχετικών με το εκάστοτε προς μελέτη θέμα. Έτσι αναζητήθηκαν αναφορές, τόσο προς τα πίσω (backwards), δηλαδή από παραπομπές που βρέθηκαν σε ένα συγκεκριμένο άρθρο, όσο και προς τα εμπρός (forwards), δηλαδή βρίσκοντας ποιοι άλλοι ερευνητές παραπέμπουν στο προς μελέτη άρθρο. Επίσης νέοι όροι που εντοπίστηκαν σε κάποια αντιπροσωπευτικά άρθρα και σχετίζονταν με το κεντρικό θέμα της εργασίας, αναζητήθηκαν στις προαναφερόμενες πύλες στο Διαδίκτυο, για ανεύρεση επιπλέον περισσότερο εξειδικευμένης βιβλιογραφίας.

Συνεπώς, με βάση την ιδέα ότι ένα μαργαριτάρι σχηματίζεται σταδιακά, από πολλαπλά στρώματα κρυστάλλινης ουσίας, έτσι και η συγκεκριμένη βιβλιογραφική ανασκόπηση αναπτύχθηκε σε στάδια, με επιστημονικά δεδομένα, από μια παραπομπή σε άλλη ή με αναζήτηση νέων σχετικών όρων. Κατά τα στάδια αυτά, ιδιαίτερη έμφαση δόθηκε στο να συμπεριληφθούν όλες εκείνες οι συναφείς και συγκλίνουσες τεχνολογικές παράμετροι, οι οποίες περιγράφουν σφαιρικά το κεντρικό θέμα της εργασίας, υπό το περιοριστικό πλαίσιο, της σχετικής με την προστασία των προσωπικών δεδομένων, νομοθεσία (ΓΚΠΔ).

Τα κύρια άρθρα που προσδιορίστηκαν, με επίκεντρο την Ανάλυση των Μαζικών Δεδομένων Υγείας (Big Health Data Analytics), αξιολογήθηκαν αρχικά μεμονωμένα, ενώ καταγράφηκε μια συνοπτική περιγραφή του καθενός, με παράθεση των ερευνητικών στόχων, των κυριότερων ευρημάτων τους, των προβλημάτων που μελετήθηκαν και των αντίστοιχων λύσεών τους. Για αυτή τη συνοπτική περιγραφή χρησιμοποιήθηκαν οι περιλήψεις (abstracts) και τα συμπεράσματα κάθε άρθρου, ενώ η μελέτη των επιμέρους προβλημάτων και λύσεων, προέκυψε από έρευνα του κυρίως μέρους τους. Στη συνέχεια τα άρθρα κατηγοριοποιήθηκαν σε πέντε βασικές ομάδες, όπου το γενικότερο θέμα-τίτλος της κάθε ομάδας, εξετάστηκε από διάφορες οπτικές γωνίες και διενεργήθηκε συγκριτική, μεταξύ τους, ανάλυση. Οι βασικές ομάδες των άρθρων είναι: 1) Analytics, 2) Big Data, 3) Precision Medicine, 4) IoT και 5) Machine Learning.

Η περιγραφική μελέτη του θέματος έγινε με σκοπό την ολοκληρωμένη και ολιστική κατανόηση των προς διερεύνηση ζητημάτων, για επέκταση της υπάρχουσας γνώσης με τρόπο διερευνητικό και επαγωγικό, ώστε να καταστεί δυνατή, η περαιτέρω χρήση και εφαρμογή των εξαχθέντων αποτελεσμάτων στο μέλλον. Έτσι στο τέλος της εργασίας παρατίθεται ο πίνακας συγκριτικής ανάλυσης των σημαντικότερων και αντιπροσωπευτικότερων άρθρων, στα οποία αυτή βασίστηκε. Η ανάλυση αυτή περιλαμβάνει, για κάθε κύριο άρθρο, τον ερευνητικό του σκοπό, την χρησιμοποιούμενη μεθοδολογία, τα κυριότερα ευρήματα που προέκυψαν από την έρευνα, τα προβλήματα που μελετήθηκαν, τις λύσεις που προτείνονται, καθώς και τα ανοικτά θέματα για μελλοντική διερεύνηση.

Μετά την περιγραφική μελέτη και προς τεκμηρίωση της προτεινόμενης τεχνολογίας, επιχειρήθηκε η δημιουργία μοντέλων Μηχανικής Μάθησης και Ανάλυσης Μαζικών Δεδομένων Υγείας, με τη βοήθεια του εργαλείου Knowi (knowi.com). Τα δεδομένα αντλήθηκαν από τον ιστότοπο data.world και επιχειρήθηκαν επιτυχείς προγνωστικές αναλύσεις με μοντέλα Ταξινόμησης, για προβλέψεις διακριτών καταστάσεων (π.χ. ασθενής έχει ή δεν έχει διαβήτη), καθώς και με μοντέλα Παλινδρόμησης, για προβλέψεις αριθμητικών τιμών (π.χ. της τιμής της γλυκοζυλιωμένης αιμοσφαιρίνης ασθενούς), μέσα από άλλα σχετιζόμενα χαρακτηριστικά (features). Στα μοντέλα Ταξινόμησης χρησιμοποιήθηκε η μέθοδος των Δέντρων Αποφάσεων (Decision Trees), ενώ στα μοντέλα Παλινδρόμησης, χρησιμοποιήθηκε η μέθοδος των Ελαχίστων Τετραγώνων (Ordinary Least Squares - OLS).

6. Σύγκλιση Συναφών Τεχνολογιών - Δυνατοτήτων - Εργαλείων

Η λεγόμενη τέταρτη βιομηχανική επανάσταση (Industry 4.0 ή I4.0) βασίζεται, από τεχνικής άποψης, κυρίως στην έννοια των Cyber-Physical Systems (CPS), μέσω σύγκλισης των τομέων πληροφορικής, επικοινωνιών και ελέγχου και υποστηρίζεται, σε μεγάλο βαθμό, από τρεις ομάδες ή πυλώνες τεχνολογιών: 1) το Διαδίκτυο των Πραγμάτων (Internet of Things), που χαρακτηρίζεται από τη διεισδυτική παρουσία μιας ποικιλίας συνεργαζόμενων αντικειμένων, όπως κινητά τηλέφωνα, αισθητήρες και ενεργοποιητές, 2) την Υπολογιστική Νέφους και Ομίχλης (Cloud and Fog Computing), που παρέχουν σχεδόν απεριόριστους υπολογιστικούς πόρους, αποθηκευτικούς χώρους και πόρους επικοινωνίας, ως βοηθητικά προγράμματα, δηλαδή κατ' απαίτηση και πληρωμή ανά χρήση και 3) την Ανάλυση των Μαζικών Δεδομένων (Big Data Analytics), για εξαγωγή αξίας, μέσα από μεγάλο όγκου δεδομένα (Aceto et al., 2020).

Ο τομέας της υγειονομικής περίθαλψης δεν θα μπορούσε να μείνει ανεπηρέαστος από τις φρενήρεις τεχνολογικές εξελίξεις. Οι από κοινού κυβερνητικές και επιστημονικές προσπάθειες, σε παγκόσμιο επίπεδο, συνετέλεσαν ώστε και η ηλεκτρονική υγεία (e-Health) να οδηγηθεί, με τη σειρά της, σε μια αντίστοιχη επανάσταση, την λεγόμενη «Υγειονομική Περίθαλψη 4.0» (Healthcare 4.0).

Ως εκ τούτου, η «Υγειονομική Περίθαλψη 4.0» χαρακτηρίζεται κι αυτή από την υιοθέτηση τριών κύριων και συγκλινουσών τεχνολογιών, δηλαδή από το Διαδίκτυο των Πραγμάτων Υγείας (IoHT), τα Μαζικά Δεδομένα Υγείας (ΜΔΥ) και την Υπολογιστική Νέφους Υγείας (H-Cloud Computing). Ως συνέπεια δε της υιοθέτησης των ΜΔΥ και της εκμετάλλευσης της αξίας τους, μέσα από την «προγνωστική» (predictive) ανάλυσή τους, απαιτείται η σύγκλιση επιπρόσθετων σύγχρονων τεχνολογιών αιχμής, η εκμετάλλευση δυνατοτήτων, καθώς και η ανάπτυξη και χρήση σύγχρονων κατάλληλων εργαλείων, όπως θα παρουσιαστούν στις υποενότητες που ακολουθούν.

Παράλληλα, η πρώτη γενιά Πληροφορικής και Ανάλυσης Δεδομένων Υγείας (Healthcare Informatics and Analytics ή HCI&A 1.0), αξιοποίησε παραδοσιακές τεχνολογίες σχεσιακών βάσεων δεδομένων και συστημάτων RDBMS, για την υποστήριξη της συλλογής, του μετασχηματισμού, της εξαγωγής, της φόρτωσης και της ανάλυσης των δεδομένων, σε διάφορες βάσεις δεδομένων υγειονομικής περίθαλψης.

Με τη δημοτικότητα του Web 2.0, εμφανίστηκε η δεύτερη γενιά πληροφορικής και ανάλυσης δεδομένων (HCI&A 2.0), για την υποστήριξη της «συνεργατικής υγειονομικής περίθαλψης» (collaborative healthcare), συνδέοντας πλήθη χρηστών, με κοινά ενδιαφέροντα και ανάγκες υγειονομικής περίθαλψης, μέσω δημόσιων ή εταιρικών κοινωνικών δικτύων. Αυτά τα δεδομένα συλλέγονται με προηγμένα εργαλεία και τεχνικές ανάλυσης μεγάλου όγκου, μη δομημένων δεδομένων, μετατρέποντάς τα σε χρήσιμες πληροφορίες για την υγειονομική περίθαλψη, όπως τα συναισθήματα των ασθενών.

Η τρίτη γενιά πληροφορικής και ανάλυσης δεδομένων (HCI&A 3.0) αντιμετωπίζει τη μεγάλη πρόκληση στο χώρο των δεδομένων υγείας, υποσχόμενη την υιοθέτηση μεγάλης κλίμακας εργαλείων διαχείρισης κι ανάλυσης δεδομένων υγειονομικής περίθαλψης (Pramanik et al., 2020).

Η σύγκλιση των υπό εξέταση τεχνολογιών, μέσα από τις πολύ υψηλές ταχύτητες, τις χαμηλές καθυστερήσεις και την ευελιξία του δικτύου 5G, μπορούν να διευκολύνουν και να εκτοξεύσουν την Προγνωστική Ανάλυση των ΜΔΥ, με σκοπό την μελλοντική πρόγνωση και πρόληψη ασθενειών του πληθυσμού γενικά και των στελεχών των ΕΔ, ειδικότερα. Απαραίτητες προϋποθέσεις αποτελούν η ασφάλεια, η ιδιωτικότητα των δεδομένων καθώς και η διατήρησή τους σε ψηφιακή μορφή, όπως στους ατομικούς ηλεκτρονικούς φακέλους υγείας.

Η Μηχανική Μάθηση με τη χρησιμοποίηση εργαλείων βαθιάς μάθησης (deep learning), λόγω της ικανότητάς της να αφομοιώνει μεγάλα σύνολα δεδομένων και να εξάγει συμπεράσματα για τις σύνθετες σχέσεις μεταξύ μεταβλητών, μπορεί να χρησιμοποιηθεί στον χώρο των ΜΔΥ και να μας δώσει ασφαλείς και πολύτιμες προγνώσεις, όπως κατέδειξε και η έρευνα των Ngiam and Khor (2019).

Στο περιβάλλον του Νέφους Υγείας (H-Cloud), τα ΜΔΥ μπορούν να αποθηκεύονται, αφού προηγουμένως συλλεχθούν μέσω συσκευών ΙοΗΤ. Σε αυτό συντελεί και η δυνατότητα σύνδεσης των υποκειμένων, μέσω κατάλληλων συσκευών και εφαρμογών στο Διαδίκτυο, σχεδόν «εν παντί τόπω και χρόνω» (anywhere-and-anytime connectivity), με εκμετάλλευση και των προσφερόμενων δυνατοτήτων του δικτύου 5G. Επειδή όμως η διασφάλιση των δεδομένων στο Νέφος αποτελεί μεγάλη πρόκληση ασφαλείας, απαιτείται η χρησιμοποίηση ασφαλών αλγορίθμων αποθήκευσης δεδομένων ή προηγμένων αλγορίθμων κρυπτογράφησης τους, που υπόσχονται λύσεις στο θέμα της ασφάλειας των ιατρικών δεδομένων, όπως κατέδειξε και η έρευνα των Kumar et al., (2018).

Η απομακρυσμένη επιτήρηση ασθενών, μέσα από τοπικές συσκευές ΙοΗΤ, που μπορούν να εντοπίζουν επικίνδυνες καταστάσεις, να εκτελούν επεξεργασία και ανάλυση δεδομένων από αισθητήρα τεχνολογίας Edge, καταδείχθηκε από την έρευνα των Mrozek et al., (2020), χρησιμοποιώντας επεκτάσιμη αρχιτεκτονική κι επιτυγχάνοντας μείωση της ποσότητας δεδομένων και της κίνησης του δικτύου.

Εξάλλου, τεχνικές όπως η απόκρυψη δεδομένων, η παρακολούθηση δραστηριότητας, η ομοιομορφική κρυπτογράφηση, ο ενσωματωμένος έλεγχος ταυτότητας και ο λεπτομερής έλεγχος πρόσβασης, εγγυώνται ασφάλεια των δεδομένων υγείας, όπως φάνηκε και στην έρευνα των Kaur et al., (2018) και πρέπει να λαμβάνονται υπόψη, κατά την ανάπτυξη συστημάτων Μηχανικής Μάθησης σε ΜΔΥ.

Εν τω μεταξύ, αυτός ο κατακλυσμός δεδομένων υγείας δημιούργησε, όπως ήταν αναμενόμενο, σοβαρές ανησυχίες σχετικά με την ανάκτηση, την αποθήκευση, την ανάλυση και την ασφάλεια των πληροφοριών των επαγγελματιών υγείας (Pramanik et al., 2020).

6.1 Μαζικά Δεδομένα Υγείας (Health Big Data)

Μαζικά Δεδομένα (Big Data) είναι εκείνα τα δεδομένα «των οποίων η κλίμακα, η ποικιλομορφία και η πολυπλοκότητα απαιτούν νέα αρχιτεκτονική, τεχνικές, αλγόριθμους και αναλυτικά στοιχεία, για τη διαχείριση και εξαγωγή αξίας και κρυμμένης γνώσης, μέσα από αυτά» (Harper, 2014).

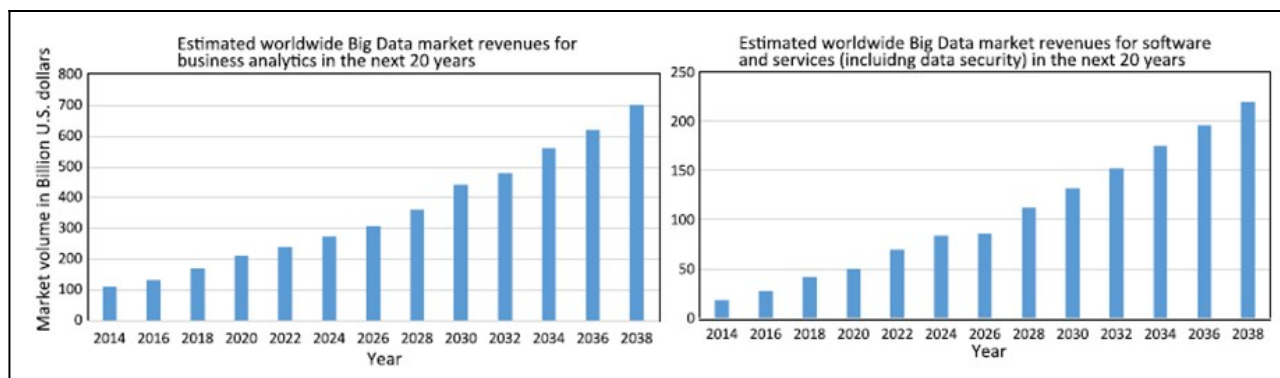
Ο όρος «Big Data» χρησιμοποιήθηκε αρχικά από τους ερευνητές Cox και Ellsworth της NASA, όταν το 1997 προσπαθούσαν να περιγράψουν τη δυσκολία προβολής συνόλων δεδομένων, πολύ μεγάλων για να αποθηκευτούν στην κύρια μνήμη του υπολογιστή. Η έννοια των Big Data είναι σχετική με τους διαθέσιμους πόρους σε δεδομένο χρονικό σημείο, κι ενώ γενικότερα χαρακτηρίζονται από τα τρία V (Volume, Velocity, Variety), δηλαδή τον μεγάλο τους όγκο, την ταχύτητα δημιουργίας και ανάλυσής τους και την ποικιλομορφία τους, σύμφωνα με τον Anil Jain (2016), αντιπρόεδρο της IBM Watson Health, στα Μαζικά Δεδομένα Υγείας (ΜΔΥ) θα πρέπει να συμπεριλάβουμε ακόμη δύο V.

Το τέταρτο V είναι αυτό της μεταβλητότητας (Variability), καθώς ο τρόπος παροχής της θεραπείας, αλλά ακόμη περισσότερο, ο τρόπος συλλογής των δεδομένων υγείας, διαφέρει σε τόπο και χρόνο. Το πέμπτο V, ανήκει στην αξία (Value), εφόσον αν πρόκειται να επενδύσει κανείς στην υποδομή που απαιτείται, για τη συλλογή και την ερμηνεία δεδομένων ευρείας κλίμακας, είναι σημαντικό να διασφαλίσει ότι οι πληροφορίες που δημιουργούνται βασίζονται σε ακριβή δεδομένα και οδηγούν σε μετρήσιμες βελτιώσεις.

Τα δεδομένα υγείας ανήκουν στα δεδομένα προσωπικού χαρακτήρα και μάλιστα συμπεριλαμβάνονται στις ειδικές κατηγορίες αυτών, δηλαδή στα λεγόμενα ευαίσθητα δεδομένα, τα οποία χαιρούν αυξημένης προστασίας σε σχέση με τα απλά δεδομένα προσωπικού χαρακτήρα. Ο όγκος, η πολυπλοκότητά τους καθώς και η συνεπακόλουθη γνώση από την επεξεργασία τους, αυξάνεται συνεχώς.

Ο συνδυασμός τέτοιων δεδομένων προερχομένων από πηγές πρωτοβάθμιας ή δευτεροβάθμιας φροντίδας υγείας ή ακόμη κι από διοικητικές υγειονομικές πηγές, μπορεί να αποκαλύψει νέες προσεγγίσεις στις μεθόδους θεραπείας του ασθενούς και καταγραφής των συμπεριφορών του, λαμβάνοντας υπόψη τις πολύπλευρες πτυχές της κλινικής και χρόνιας περίθαλψης.

Στατιστικές εκτιμήσεις δείχνουν, όπως στο Σχ. 1, τη δυναμική των Μαζικών Δεδομένων. Σύμφωνα με αυτές, τα έσοδα της παγκόσμιας αγοράς Μαζικών Δεδομένων για επιχειρηματικές αναλύσεις θα φτάσουν τα 700 δισεκατομμύρια δολάρια σε βάθος εικοσαετίας, ενώ τα αντίστοιχα έσοδα για λογισμικό – υπηρεσίες, συμπεριλαμβανομένης της ασφάλειας δεδομένων, θα ξεπεράσουν τα 200 δισεκατομμύρια δολάρια.



Σχ. 1 Στατιστικές εκτιμήσεις εσόδων της παγκόσμιας αγοράς Μαζικών Δεδομένων για επιχειρηματικές αναλύσεις (αριστερά) καθώς και για λογισμικό – υπηρεσίες, συμπεριλαμβανομένης της ασφάλειας δεδομένων, στα επόμενα 20 χρόνια (δεξιά). (Πηγή: Tao et al., 2019).

Πηγές Μαζικών Δεδομένων Υγείας ή Health Big Data μπορούν να αποτελέσουν τα δεδομένα των ιατρικών φακέλων των ασθενών και οι υπολογιστικές αναλύσεις των δεδομένων αυτών, ενώ πολύ σημαντική πηγή τέτοιων δεδομένων είναι τα δεδομένα PROM (Patient Related Outcome Measurements) και τα δεδομένα PREM (Patient Related Experience Measurements), τα οποία συλλέγονται από τους ασθενείς μέσω μετρήσεων από εφαρμογές κινητών τηλεφώνων ή υπολογιστών, στο πλαίσιο παρακολούθησής τους από τους φορείς υγείας ή με πρωτοβουλία των ίδιων των ασθενών (Willems et al., 2019).

Σύμφωνα με πρόσφατη έρευνα (Alam και Rahmani, 2020), υπάρχει πληθώρα μεταδεδομένων (metadata) στον χώρο της υγείας, τα οποία μπορούν να σχετίζονται με τους ασθενείς, με τους παρόχους υγείας, με τις φαρμακευτικές αγωγές, με τους πληρωτές των παρεχόμενων υπηρεσιών υγείας, με τις υπηρεσίες διακυβέρνησης για το ρυθμιστικό πλαίσιο λειτουργίας, με τους παρόχους υπηρεσιών δεδομένων υγείας, με την ιατρική έρευνα και με τους κατασκευαστές ιατρικών συσκευών.

Τεράστιος εξάλλου είναι ο όγκος μια άλλης κατηγορίας πηγών ΜΔΥ, που δεν σχετίζονται με τα δεδομένα αυτά καθαυτά των ασθενών, αλλά με την διεθνή επιστημονική βιβλιογραφία, καθώς υπολογίζεται ότι ετησίως δημοσιεύονται πάνω από ένα εκατομμύριο άρθρα, σχετικά με την υγεία. Πρακτικά είναι αδύνατο να αναγνωστεί έστω κι ένα μικρό κλάσμα όλων αυτών των δημοσιεύσεων, μη υπολογιζόμενων και άλλων σχετικών κειμένων ή αναρτήσεων στο Διαδίκτυο (Willems et al., 2019).

Τα ΜΔΥ βασίζονται σε τέσσερις πυλώνες (Archena et al., 2015): 1) Στην περίθαλψη με επίκεντρο τον ασθενή, όπου χρησιμοποιούνται διαγνώσεις βάσει κλινικών δεδομένων, σε πρώιμο στάδιο, ελαχιστοποιώντας τη δοσολογία των φαρμάκων, για την αποφυγή παρενεργειών στους ασθενείς. Αυτό βοηθά στη μείωση των ποσοστών επανεισδοχής στα νοσοκομεία, καθώς και στη μείωση του κόστους για τους ασθενείς. 2) Στην προγνωστική ανάλυση ασθενειών, όπου πραγματοποιείται πρόβλεψη των ιογενών διαταραχών σε πρώιμο στάδιο, πριν από την εξάπλωσή τους, με βάση ανάλυσή τους σε πραγματικό χρόνο. Αυτό μπορεί να επιτευχθεί αναλύοντας τα κοινωνικά αρχεία των ασθενών, που πάσχουν από μια ασθένεια, σε μια συγκεκριμένη τοποθεσία και θα βοηθήσει περαιτέρω τους επαγγελματίες υγείας να συμβουλευτούν τους πιθανούς ασθενείς, λαμβάνοντας τα απαραίτητα προληπτικά μέτρα. 3) Στην παρακολούθηση ασθενών σε πραγματικό χρόνο, όπου ελέγχεται κατά πόσον τα νοσοκομεία λειτουργούν σύμφωνα με τους κανόνες που έχουν ρυθμιστεί από το εθνικά ιατρικά συμβούλια. Αυτός ο περιοδικός έλεγχος μπορεί να βοηθήσει τις εθνικές κυβερνήσεις να λάβουν τα απαραίτητα μέτρα, για τον αποκλεισμό ακατάλληλων νοσοκομείων. 4) Στη βελτίωση των μεθόδων θεραπείας, όπου πραγματοποιείται προσαρμοσμένη παρακολούθηση της θεραπείας των ασθενών και μετά από την ανάλυση των φαρμάκων τους, οι δοσολογίες μπορούν να αλλάζουν, για ταχεία ανακούφισή τους.

Η ανάλυση των δεδομένων των ασθενών που υποφέρουν από παρόμοια συμπτώματα βοηθά τους γιατρούς να παρέχουν αποτελεσματικότερα φάρμακα σε νέους ασθενείς.

Σύμφωνα με πρόσφατη βιβλιογραφική ανασκόπηση (Rajabion et al., 2019), για τα ΜΔΥ χρησιμοποιείται ευρύτατα πλέον η Υπολογιστική Νέφος (Cloud Computing). Με την τεχνολογία αυτή διευκολύνεται ο χειρισμός και η επεξεργασία των δεδομένων υγείας, ενώ υπάρχουν τέσσερις κύριες κατηγορίες, όπου τα ΜΔΥ βρίσκουν εφαρμογή. Οι κατηγορίες αυτές είναι: 1) η εξ αποστάσεως θεραπεία, 2) η λήψη-διαμοιρασμός-αποθήκευση πληροφοριών του ασθενούς, 3) η προστασία του περιβάλλοντος και 4) η παρακολούθηση της φροντίδας.

Στην εξ αποστάσεως θεραπεία, η τεχνολογία παρέχει τη δυνατότητα στους ασθενείς να λαμβάνουν υπηρεσίες υγείας, χωρίς να πρέπει να μετακινηθούν προς τα νοσηλευτικά ιδρύματα, όταν για παράδειγμα κατοικούν σε απομακρυσμένες περιοχές ή δεν μπορούν να ταξιδέψουν για να επισκεφτούν κάποιον γιατρό. Με προηγμένες εφαρμογές και μέσω της υπολογιστικής νέφος, διακινούνται και αποθηκεύονται ογκώδη δεδομένα απεικονιστικών εξετάσεων, γίνονται τηλεδιασκέψεις μεταξύ γιατρών και ασθενών για παροχή οδηγιών και συμβουλών, καθώς και συνεχής εκπαίδευση σχετικά με τη θεραπεία τους.

Έχει διαπιστωθεί ότι τα παραδοσιακά συστήματα Τεχνολογιών Πληροφορικής και Επικοινωνιών (ΤΠΕ), καταστρέφουν το περιβάλλον λόγω των ζημιόγων εκπομπών άνθρακα. Ωστόσο, με τη χρήση μεγάλων κέντρων δεδομένων (data centers), που εκμεταλλεύονται τις τεχνολογικές δυνατότητες της Υπολογιστικής Νέφος για τη διακίνηση και αποθήκευση των ΜΔΥ, εξοικονομείται ενέργεια και προστατεύεται το περιβάλλον, ενώ ταυτόχρονα μειώνονται δραστικά, τα κόστη λειτουργίας των νοσηλευτικών φορέων.

Στην εξ αποστάσεως παρακολούθηση των ασθενών, τα ΜΔΥ χρειάζονται εκτός από τις υποδομές της Υπολογιστικής Νέφος, τη χρήση συσκευών του Διαδικτύου των Πραγμάτων (IoT), όπου κατάλληλοι αισθητήρες που φέρονται στο χέρι ή σε άλλη θέση του σώματος του ασθενούς, δίνουν πληροφορίες σε διασυνδεδεμένες με το Νέφος πλατφόρμες παρακολούθησης, παρέχοντας συνεχή ενημέρωση για την κατάσταση του ασθενούς, σε θεράποντες ιατρούς και νοσηλευτές.

Η ιατρική επιστήμη προσδοκά από τα ΜΔΥ ουσιαστικά δύο κύρια αποτελέσματα: την ικανότητα δημιουργίας προγνωστικών μοντέλων και τη δυνατότητα εφαρμογής τους, ώστε να ληφθούν ακριβή προφίλ κινδύνου ασθενών ή/και να ακολουθηθούν σωστές θεραπευτικές οδοί (Carobianco, 2020). Εξάλλου, σύμφωνα με τους Praba και Srividhya (2016), υπάρχουν δύο μεγάλες κατηγορίες τύπων ΜΔΥ, τα Big Data των Κλινικών Λειτουργιών και τα Big Data Έρευνας κι Ανάπτυξης. Τα Big Data των Κλινικών Λειτουργιών αναφέρονται σε όλες τις δραστηριότητες που σχετίζονται με τους ασθενείς και περιλαμβάνουν δεδομένα αναγνώρισης ή διάγνωσης ασθενειών, παρακολούθησης της εξέλιξής τους, παρατηρήσεις, θεραπείες και αναφορές, σχετικά με τις εφαρμοζόμενες θεραπευτικές αγωγές.

Τα Big Data Έρευνας κι Ανάπτυξης διαδραματίζουν σπουδαίο ρόλο στη βιομηχανία των δεδομένων υγείας. Τα δεδομένα αυτά προϋποθέτουν χρήση αλγορίθμων και εργαλείων για την παροχή ποιοτικής θεραπείας και συνεχή βελτίωση της σχεδίασης των κλινικών δοκιμών. Επίσης περιλαμβάνουν δεδομένα προσαρμοστικών κλινικών δοκιμών για την αξιολόγηση κάποιας θεραπείας ή κάποιας ιατρικής συσκευής, μέσω της παρατήρησης των αποτελεσμάτων των ασθενών, καθώς και αλλαγών των παραμέτρων του πρωτοκόλλου δοκιμών, σε σύγκριση με παλαιότερες παρατηρήσεις. Στην κατηγορία αυτή εξάλλου, ανήκουν και τα δεδομένα Προγνωστικής Μοντελοποίησης, με τα οποία παρασκευάζονται αποτελεσματικότερα φάρμακα και ιατρικές συσκευές.

Εκτός από την Έρευνα κι Ανάπτυξη, τα Big Data μπορούν να συνεισφέρουν στη γονιδιωματική ανάλυση, στην ανάλυση και σύγκριση ποικιλόμορφων δομημένων και αδόμητων δεδομένων, στην ανάλυση του προφίλ των ασθενών, στην εξ αποστάσεως παρακολούθηση ασθενών και στην επιτάχυνση της καινοτομίας. Σωστές αποφάσεις και καινοτόμες ιδέες προάγουν την ανάπτυξη, την πρόοδο και την ασφάλεια στο χώρο της υγείας (Praba και Srividhya, 2016). Παράδειγμα δομημένων δεδομένων αποτελούν οι κωδικοί ICD (International Classification of Diseases), ενώ παράδειγμα αδόμητων δεδομένων αποτελούν οι καταγραφές συμπτωμάτων του ασθενούς (Cirillo και Valencia, 2019).

Σύμφωνα με εκτιμήσεις, ο όγκος δεδομένων υγειονομικής περίθαλψης ανέρχονταν σε περισσότερα από 700 exabytes το 2017, από 153 εκατομμύρια που ήταν το 2013 και αναμένονταν να φθάσει το 2020, στα 2.314 exabytes (prnewswire.com, 2018). Για να κατανοήσουμε περίπου το μέγεθος των δεδομένων αυτών αρκεί να αναφέρουμε ότι σε μόλις 5 exabytes (1018 gigabytes) ανέρχεται το σύνολο όλων των λέξεων, που έχουν ειπωθεί από τον άνθρωπο, από τη στιγμή της εμφάνισής του στη γη ως σήμερα (Cottle et al., 2013). Εκτιμάται δε ότι ο διαθέσιμος αποθηκευτικός χώρος όλων των δεδομένων παγκοσμίως, θα επαρκεί για περίπου 985 exabytes το 2020, γεγονός που σημαίνει ότι απαιτείται τουλάχιστον 2,5 φορές περισσότερος αποθηκευτικός χώρος, για να «στεγάσει» τα δεδομένα από τον χώρο της υγείας (Desjardins, 2018).

Ο ρόλος της Πληροφορικής είναι να παρέχει τα κατάλληλα εργαλεία για να μετατρέπει δεδομένα σε πληροφορίες και πληροφορίες σε γνώση, βοηθώντας στην κατανόηση των βαθέων σχέσεων μέσα από τα δεδομένα, ανακτώντας και εξάγοντας υποκείμενα μοτίβα (patterns) μέσα από αυτά. Επιπλέον, η πληροφορική είναι ζωτικής σημασίας στην εκμετάλλευση της αποκτηθείσας γνώσης, για την υποστήριξη της περίθαλψης των ασθενών και τελικά την καθοδήγηση της ατομικής τους συμπεριφοράς.

Η εγκυρότητα των ΜΔΥ εξαρτάται από το βαθμό της ακρίβειάς τους κι αυτό επηρεάζει τη διατύπωση των διαγνώσεων των ασθενειών, σε διάφορες περιπτώσεις. Αρκετά συχνά, ασθενείς παρουσιάζουν μια εγγενή ποικιλομορφία, η οποία περιπλέκει τη δυνατότητα εξήγησης ή επικύρωσης των παραγόντων εκείνων, που συμβάλλουν διαφορετικά στον χαρακτηρισμό της νόσου τους. Τέτοιοι παράγοντες μπορούν να εξεταστούν και να μετρηθούν σε ένα μεταβλητό επίπεδο ακρίβειας και η συμπερίληψή τους σε καλά σχεδιασμένα μοντέλα, αντιπροσωπεύει την πραγματική προστιθέμενη αξία, που μπορούν να προσφέρουν τα ΜΔΥ στη γνώση της νόσου, αυξάνοντας την ερμηνεία διαφορετικών αλλά δυνητικά συσχετισμένων επιπέδων πληροφοριών.

Οι ασθενείς, ως «καταναλωτές» των προσωπικών τους δεδομένων υγείας, μπορούν να έχουν πλέον τον πλήρη έλεγχο και να αποθηκεύουν τα δεδομένα που αφορούν την υγεία τους, σε εφαρμογές της επιλογής τους, τοπικά ή δικτυακά στο Νέφος (Cloud). Επιτυχημένο παράδειγμα τέτοιας εφαρμογής αποτελεί το HealthKit της εταιρίας Apple. Εξάλλου, βάσει του ΓΚΠΔ, είναι υποχρεωμένοι πλέον οι υπεύθυνοι επεξεργασίας δεδομένων, να καθιστούν τα δεδομένα των ασθενών προσβάσιμα κατόπιν αίτησής τους, μεταξύ άλλων και σε εύχρηστη ηλεκτρονική μορφή.

Το HealthKit αποτελεί ένα ενοποιημένο αποθετήριο για ποικιλία δεδομένων, συμπεριλαμβανομένων ιατρικών αρχείων, με πληροφορίες όπως εργαστηριακά αποτελέσματα, ανοσοποιήσεις και φάρμακα, που έχουν ληφθεί απευθείας από φορείς παροχής υγειονομικών υπηρεσιών. Η εφαρμογή προσφέρει νέες κατηγορίες και τύπους αρχείων υγείας, που έχουν δημιουργηθεί για να παρέχουν εύκολη πρόσβαση σε τύπους εγγραφών υψηλού επιπέδου, χωρίς να αποτρέπεται η πρόσβαση στα υφιστάμενα ακατέργαστα δεδομένα. Χρησιμοποιεί νέες τεχνικές εξουσιοδότησης, που επιτρέπουν πρόσβαση σε δεδομένα καρτών υγείας για ανάλυση και παρουσίαση, κατά τρόπο που σέβεται το ιατρικό απόρρητο.

Τα ΜΔΥ φημίζονται για τον όγκο, την ακαταστασία και την πολυπλοκότητά τους. Για να καταστούν χρήσιμα, πρέπει να χαρτογραφούνται κατάλληλα και να υφίστανται προεπεξεργασία (preprocessing) πριν χρησιμοποιηθούν για εκπαίδευση, με μεθόδους Μηχανικής Μάθησης - Machine Learning (ML). Αυτό το βήμα είναι θεμελιώδες στη δημιουργία μοντέλων ML, επειδή η ακρίβεια του κάθε μοντέλου εξαρτάται σε μεγάλο βαθμό από την αξιοπιστία των δεδομένων, ως προς την κλινική τους πραγματικότητα.

Συνεπώς, ένα μεγάλο σύνολο δεδομένων (dataset), πρέπει να υποβληθεί σε προεπεξεργασία, όπου τα ακατέργαστα δεδομένα μετατρέπονται σε κατανοητή μορφή, με τη χρήση κατάλληλων τεχνικών εξόρυξης. Τα ακατέργαστα σύνολα δεδομένων είναι γενικά ελλιπή και προβληματικά, συνήθως τούς λείπουν κάποιες τιμές, εμφανίζονται με ακραίες τιμές και με μεταβλητές που περιέχουν «θόρυβο ή μη πληροφοριακά (noninformative) δεδομένα. Μπορεί επίσης να στερούνται παντελώς κάποιων σημαντικών μεταβλητών. Αυτά τα προβλήματα αντιμετωπίζονται χρησιμοποιώντας διαφορετικά εργαλεία και τεχνικές εξόρυξης δεδομένων κατά την προεπεξεργασία τους, η οποία μπορεί να αναφέρεται και ως «καθαρισμός» (cleaning) δεδομένων (Khan et al., 2019).

Για το χειρισμό των Big Data πολύ κατάλληλη θεωρείται η βαθιά μάθηση (deep learning), λόγω κυρίως της ικανότητάς της να αφομοιώνει μεγάλα σύνολα δεδομένων και να εξάγει συμπεράσματα για τις σύνθετες σχέσεις μεταξύ μεταβλητών, με τρόπο ευέλικτο και πρακτικό. Στα βαθιά νευρωνικά δίκτυα, πολλαπλά επίπεδα ενδιάμεσων μεταβλητών συνδέονται μεταξύ τους, έτσι ώστε οι έξοδοι ενός επιπέδου να χρησιμεύουν ως είσοδοι του επόμενου επιπέδου.

Αυτή η δομή διευκολύνει την ανάλυση των δεδομένα υψηλών διαστάσεων (high-dimensional data), που περιλαμβάνουν περισσότερες από 200 μεταβλητές (features). Αντίθετα, οι παραδοσιακές στατιστικές μέθοδοι, όπως η Γραμμική Παλινδρόμηση, περιλαμβάνουν μόνο ένα επίπεδο εισόδου-εξόδου και μπορούν να φιλοξενήσουν σχετικά μικρές ποσότητες μεταβολής (Ngiam and Khor, 2019).

Τα ΜΔΥ των ασθενών (στελεχών των ΕΔ εν προκειμένω) με μεγάλο πλήθος μεταβλητών, μπορούν να χρησιμοποιηθούν για να κτιστούν μοντέλα Τεχνητής Νοημοσύνης, τα οποία θα προσφέρουν πρόγνωση επί μελλοντικών συμβάντων, στη βάση του στατιστικού βάρους ιστορικών συσχετίσεων μεταξύ των δεδομένων υγείας τους. Αν συνδυαστούν με μεγάλα σύνολα δεδομένων γονιδιωματικής ή περιβαλλοντολογικής προέλευσης, μπορούν να οδηγήσουν στην πρόγνωση σοβαρών ασθενειών όπως ο καρκίνος, τόσο για μεμονωμένους ασθενείς όσο και για ομάδες ασθενών με συγκεκριμένα χαρακτηριστικά (Willems et al., 2019). Μια τέτοια δυνατότητα μπορεί να αποκαλύψει την ανάγκη να εκδηλωθούν συγκεκριμένες διορθωτικές ενδεχομένως ενέργειες, οι οποίες θα μπορέσουν να επιδράσουν σε όλους εκείνους τους περιβαλλοντικούς παράγοντες και σε όλη εκείνη τη συμπεριφορά των υποκειμένων των ομάδων-στόχων, που πιθανώς συνέβαλαν στην επαύξηση των κινδύνων για την υγεία τους.

Σύμφωνα με πρόσφατη ανασκόπηση της επιστημονικής βιβλιογραφίας και χαρτογράφηση των ερευνητικών τάσεων (Saheb και Izadi, 2019), η σύγκλιση των τεχνολογιών IoT και Big Data, οδηγεί σε ένα σύγχρονο πρότυπο για τα συστήματα υγειονομικής περίθαλψης, που διευκολύνει την έξυπνη διαχείριση των σχετικών με αυτήν διαδικασιών. Στο νέο έξυπνο πρότυπο που ονομάζεται IoTBDA (Internet of Things Big Data Analytics), οι δύο προαναφερθείσες τεχνολογίες λειτουργούν ως συμπληρωματικές, ώστε συνδεδεμένες μεταξύ τους, να δύνανται να προκαλέσουν σημαντικές αναταράξεις στη βιομηχανία της υγείας.

Η εννοιολογική προσέγγιση συνέχειας των δεδομένων, που απεικονίζεται στο τρίπτυχο «δεδομένα - πληροφορίες - γνώση», επαναπροσδιορίζεται πλέον και εκφράζεται από τον όρο «Κύκλος Εκμάθησης Συστήματος Ιατρικής Περίθαλψης» (Learning Healthcare System Cycle - LHSC), όπου η υγειονομική περίθαλψη στην πράξη και η επιστημονική έρευνα πρέπει να είναι μέρος της ίδιας συνεργατικής διαδικασίας. Η βασική καινοτομία αυτής της προσέγγισης είναι ότι η κλινική πρακτική και η έρευνα είναι συμπληρωματικοί παράγοντες, στη δημιουργία δεδομένων και γνώσεων.

Η ενσωμάτωση ετερογενών πληροφοριών από διαφορετικές πηγές, γεγονός που συνεπάγεται την ανάγκη εναρμόνισης πριν από την ορθή αξιολόγηση του προγνωστικού δυναμικού τους καθώς και η εξαγωγή της κλινικής αξίας, από μεθόδους αξιοποίησης δομημένων και μη δομημένων δεδομένων, είναι δύο από τα μεγάλα προβλήματα που παραμένουν στο χώρο των ΜΔΥ (Carobianco, 2020).

Η χρήση των ΜΔΥ είναι εξίσου σημαντική στις διάφορες φάσεις του κύκλου LHSC, δηλαδή από την έρευνα έως τη λήψη αποφάσεων, με βάση τα δεδομένα. Ο κύκλος LHSC βασίζεται πράγματι σε αυτές τις δύο συμπληρωματικές δράσεις, δηλαδή στην εκμετάλλευση των ιατρικών δεδομένων που παράγονται για ερευνητικούς σκοπούς και στην ανάπτυξη νέων συστημάτων που αξιοποιούν τα ΜΔΥ, ώστε να συνδράμουν σε λήψη κλινικών αποφάσεων (Dagliati et al., 2018).

6.2 Ατομικός Ηλεκτρονικός Φάκελος Υγείας (ΑΗΦΥ)

Ογκώδης, ποικίλες και ταχείες ιατρικές πληροφορίες ασθενών, διακινούνται, διαμοιράζονται και αποθηκεύονται μεταξύ ιατρικών φορέων, συνηθέστερα μέσω του Ηλεκτρονικού Φακέλου Υγείας (ΗΦΥ). Ο ΗΦΥ, όπου αποθηκεύονται τα δεδομένα υγείας, είναι χρήσιμος για την καταγραφή πλήρους ιατρικού ιστορικού του ασθενούς και για την εξαγωγή συμπερασμάτων, τα οποία προκύπτουν από τη διενέργεια συγκρίσεων και συσχετίσεων παλαιότερων ασθενειών του. Επίσης, επαρκώς συμπληρωμένοι ΗΦΥ, αποτελούν χρησιμότερη πηγή αναζήτησης προτύπων και τάσεων, κρυμμένων στα δεδομένα των ασθενών, μέσω τεχνικών εξόρυξης δεδομένων (data mining). Η αποκωδικοποίηση τέτοιων προτύπων και τάσεων μπορεί να οδηγήσει σε πρόγνωση και έγκαιρη διάγνωση ασθενειών.

Η πρώτη προσέγγιση του ορισμού του ιατρικού φακέλου είναι η προσέγγιση που υιοθετήθηκε από το preStandard ENV 136061 της Ευρωπαϊκής Επιτροπής Τυποποίησης (CEN). Σύμφωνα με το πρότυπο: «Ο Ιατρικός Φάκελος είναι η «αποθήκη» όλων των πληροφοριών που αφορούν το ιατρικό ιστορικό του ασθενούς. Αποτελεί επομένως τη βάση της διάγνωσης και της θεραπευτικής αντιμετώπισης του ασθενούς, αλλά και τη βάση επιδημιολογικών ερευνών, κλινικών ερευνών αποθήκευσης πληροφοριών και μελλοντικής αξιοποίησης των. Επιπλέον, παρέχει πληροφορίες διοικητικής, οικονομικής και στατιστικής φύσεως, καθώς και ποσοτικού και ποιοτικού ελέγχου των στοιχείων αυτών».

Ο Ν.4600/2019 προβλέπει την δημιουργία και τήρηση πλέον και στη χώρα μας του Ατομικού Ηλεκτρονικού Φακέλου Υγείας (ΑΗΦΥ), ο οποίος αποσκοπεί στην προάσπιση, την προστασία και την προαγωγή της υγείας του πληθυσμού, μέσω του προγραμματισμού και της υλοποίησης πολιτικών δημόσιας υγείας, στη διασφάλιση της καθολικής και ισότιμης πρόσβασης στην παροχή ποιοτικά και ποσοτικά επαρκών υπηρεσιών υγειονομικής φροντίδας από το Εθνικό Σύστημα Υγείας, στη διασφάλιση των πόρων που διατίθενται για την υγειονομική περίθαλψη, τον έλεγχο των δαπανών και την αποτελεσματική χρηματοδότηση της υγειονομικής περίθαλψης, καθώς επίσης και στη ρύθμιση της λειτουργίας και στην άσκηση εποπτείας στους φορείς υγειονομικής φροντίδας του ιδιωτικού τομέα.

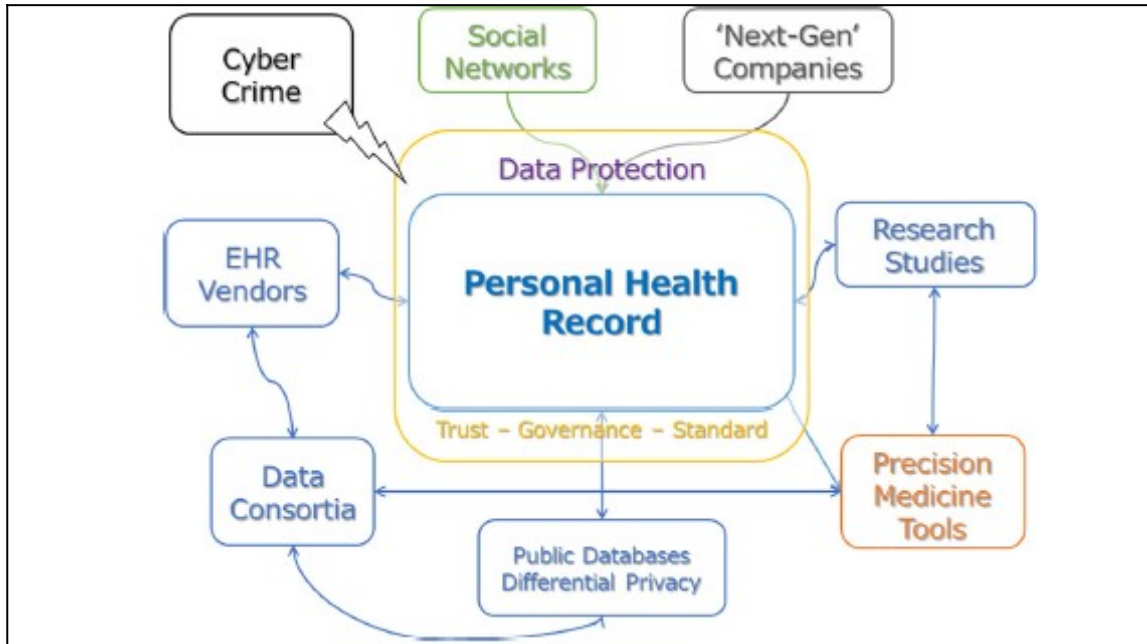
Η επικράτηση των Big Data και της ανάλυσής τους, προσφέρουν την ευκαιρία αλλαγής και αξιοποίησης της δυναμικής των ΑΗΦΥ, παρέχοντας καλύτερα εργαλεία και περισσότερες πληροφορίες στον θεράποντα ιατρό, ως προς τη λήψη αποφάσεων και συνεπώς καλύτερη ποιότητα ιατρικών υπηρεσιών (Stopa et al., 2019).

Σύμφωνα με τον Κώδικα Ιατρικής Δεοντολογίας (Ν.3418/2005) και συγκεκριμένα σύμφωνα με το άρθρο 14, η τήρηση ιατρικού αρχείου προβλέπεται για μία δεκαετία, από την τελευταία επίσκεψη του ασθενούς στα ιδιωτικά ιατρεία και τις λοιπές μονάδες πρωτοβάθμιας φροντίδας υγείας του ιδιωτικού τομέα και για μία εικοσαετία, από την τελευταία επίσκεψη του ασθενή, σε κάθε άλλη περίπτωση.

Για να γίνει κατανοητό το μέγεθος των ιατρικών δεδομένων που πιθανώς θα πρέπει να αποθηκεύονται σε έναν ΑΗΦΥ, αναφέρεται ότι σύμφωνα με εκτιμήσεις ειδικών, κάθε ασθενής προσθέτει περίπου 4 megabytes δεδομένων απλού κειμένου και 76 megabytes απεικονιστικών εξετάσεων στον ΑΗΦΥ του, ανά έτος. Σε μόλις 4 χρόνια, αυτή η διαρκής εναπόθεση δεδομένων μπορεί να πάρει τρομακτικές διαστάσεις, αφού ένας μόνο ασθενής θα χρειάζεται περίπου 320 megabyte αποθηκευτικού χώρου, για δεδομένα κειμένου και απεικονιστικών εξετάσεων. Επομένως, οι πάροχοι υγειονομικής περίθαλψης, πρέπει να διασφαλίσουν ότι έχουν αρκετό χώρο για αυτά τα πρόσθετα δεδομένα ασθενών, διατηρώντας παράλληλα και τα προηγούμενα δεδομένα του ιατρικού ιστορικού των ασθενών τους. Με τον παραπάνω ρυθμό, για έναν υποθετικό αριθμό 100 ασθενών και για χρονική διάρκεια 10 ετών, θα χρειαστούν περίπου 80 gigabytes (80 MB x 100 ασθενείς x 10 έτη) δεδομένων υγείας (Brown, 2015).

Ο προαναφερόμενος όγκος των ιατρικών δεδομένων ενός ασθενούς, της τάξης των 80 megabytes ανά έτος, μεταφράζεται σε περίπου 20.000 σελίδες κειμένου. Στα περίπου 20-40 λεπτά, που κατά μέσο όρο διαρκεί η επίσκεψη ενός ασθενούς στον ιατρό του, είναι ανθρωπίνως αδύνατο να προσπελαστούν τα ιατρικά δεδομένα του ΑΗΦΥ του. Με τις κατάλληλες όμως τεχνικές Μηχανικής Μάθησης, αυτός ο όγκος δεδομένων μεταφράζεται σε χρήσιμη και προσπελάσιμη γνώση (Waring et al., 2020).

Συναφές με την έννοια του ΑΗΦΥ είναι το «είδωλο υγείας» (health avatar), όπως αυτό προφητικά περιγράφηκε από τους Buchan et al., το 2009. Πρόκειται για μια εικονική αναπαράσταση ενός ατόμου, που περιλαμβάνει όλες τις σχετικές με την υγεία του πληροφορίες του και έξυπνους τρόπους διαχείρισης και πρόβλεψης της μελλοντικής της κατάστασης. Επικεντρώνεται στα ατομικά αρχεία υγείας του υποκειμένου και μπορεί να ενσωματωθεί στους φορείς υγειονομικής περίθαλψης, εμπορικής διακυβέρνησης και έρευνας.



Σχ. 2 Σχηματική απεικόνιση του «είδωλο υγείας», με όλες τις σχετικές πληροφορίες για την υγεία του υποκειμένου και έξυπνους τρόπους διαχείρισης/πρόβλεψης της μελλοντικής του κατάστασης υγείας (Πηγή: Prosperi et al., 2018).

Συνήθως οι ΑΗΦΥ δεν διασυνδέονται με εργαλεία διάγνωσης ή βελτιστοποίησης της θεραπείας, κι έτσι ο θεράπων ιατρός μπορεί μεν να λάβει ή να μεταφέρει αποτελέσματα εργαστηριακών εξετάσεων του ασθενούς μέσω δικτύου, στη συνέχεια όμως, οι διαγνώσεις του γίνονται με τον παραδοσιακό τρόπο, με βάση τα δεδομένα του μέσου πληθυσμού. Για το λόγο αυτό, σε σχετική μελέτη προτείνεται ένα μοντέλο σύγχρονου «είδωλο υγείας», στην εποχή των ΜΔΥ των ασθενών (Prosperi et al., 2018).

Το «είδωλο υγείας» θα πρέπει να συνδέει όλους τους νέους τύπους δεδομένων που σχετίζονται με την υγεία, όπως στοιχεία γονιδιωματικής, από διάφορες πηγές -omics, κινητές, φορητές ή περιβαλλοντολογικές πηγές. Έτσι θα συλλέγονται πληροφορίες κι από άλλους τομείς που επηρεάζουν την υγεία, όχι μόνο από τον τομέα της ιατροφαρμακευτικής περίθαλψης. Γενικότερα, το «είδωλο υγείας» πρέπει να είναι σε θέση να συνδεθεί και να εκμεταλλευτεί πληροφορίες, που δεν σχετίζονται με τον ΑΗΦΥ και που ενδεχομένως είναι χρήσιμες για την αξιολόγηση της υγείας του ατόμου, ακόμη και προερχόμενες από εξαιρετικά μη δομημένες πηγές, όπως τα μέσα κοινωνικής δικτύωσης, στα οποία ενδεχομένως συνδέεται και συμμετέχει. Για παράδειγμα, ένα εργαλείο AI θα μπορούσε να επεξεργαστεί εικόνες από αναρτήσεις του υποκειμένου στο Instagram, στο Facebook ή στο Twitter, για να εξακριβώσει τις διατροφικές συνήθειές του.

Αυτές οι πληροφορίες μπορούν στη συνέχεια να χρησιμοποιηθούν για τη συμπλήρωση ενός ερωτηματολογίου τροφίμων, κωδικοποιημένου σε κάποιο τύπο δομημένης πληροφορίας και να αποθηκευτούν στον ΑΗΦΥ του. Επιπλέον, πληροφορίες σε επίπεδο περιβάλλοντος που σχετίζονται με το άτομο (όπως για παράδειγμα μέσω της εξακρίβωσης της κατοικίας του ή της γεωγραφικής του θέσης), θα μπορούσαν να συμπληρώσουν πεδία του ΑΗΦΥ, αποθηκεύοντας πληροφορίες όπως η έκθεσή του σε αλλεργιογόνα και ρύπους. Ασφαλώς, πρόσβαση σε τέτοιου είδους πληροφορίες πρέπει να λαμβάνει υπόψη τους νομικούς περιορισμούς της ιδιωτικότητας και της προστασίας των προσωπικών δεδομένων του υποκειμένου.

Στο προτεινόμενο μοντέλο, το «είδωλο υγείας» θα διαθέτει σύνδεση με εξατομικευμένα εργαλεία πρόβλεψης της μελλοντικής κατάστασης της υγείας του υποκειμένου, ενώ στο πλαίσιο κανόνων δεοντολογίας και ρητής συναίνεσης του ατόμου, θα μπορεί να τροφοδοτεί με πληροφορίες, διάφορα ερευνητικά έργα, για τη δημιουργία νέων ακριβέστερων εργαλείων Ιατρικής Ακριβείας (Prosperi et al., 2018).

6.3 Συστήματα Υποστήριξης Κλινικών Αποφάσεων (CDSS)

Η χρήση υπολογιστών ως εργαλείων ασφάλειας και έρευνας για την παρακολούθηση, καταγραφή και αυτοματοποίηση πληροφοριών ασθενών, προτάθηκε ήδη από το τη δεκαετία του 1960. Κατά τις επόμενες δεκαετίες, έγιναν αρκετές προσπάθειες για την ενσωμάτωση συστημάτων πληροφορικής στο ιατρικό σύστημα, χρησιμοποιώντας Συστήματα Υποστήριξης Κλινικών Αποφάσεων (Clinical Decision Support Systems - CDSS).

Τυπικά ένα CDSS περιλαμβάνει: 1) Είσοδο από Μαζικά Δεδομένα Υγείας, προερχόμενα από μετρήσεις, κωδικούς διάγνωσης (π.χ. κατά ICD-10) ή εργαστηριακά αποτελέσματα, 2) Επεξεργασία με αλγορίθμους ΑΙ, συμπεριλαμβανομένων μεθόδων χειρισμού χαμένων ή ελλιπών δεδομένων, κατασκευής μεταβλητών από πρωτογενή δεδομένα και 3) Έξοδο ως υποστήριξη στη λήψη κλινικής απόφασης, που πρέπει να είναι αυτόνομη, χρηστική, αξιόπιστη κ.λπ..

Μεταξύ των διαφόρων χαρακτηριστικών των CDSS, δύο θεωρούνται ιδιαίτερα απαραίτητα, ήτοι η αξιοποίηση γνώσεων σχετικά με ασθένειες, μέσω της οποίας διατυπώνονται προβλέψεις κινδύνου και η ολοκλήρωση της γνώσης μεταξύ των διαφόρων κυρίων δρώντων σε ένα νοσοκομείο. Η συνάφεια μεταξύ κτήσης και ενσωμάτωσης γνώσεων που μπορούν να χρησιμοποιηθούν από ένα CDSS, οδηγεί σε μια κύρια προτεραιότητα, πώς δηλαδή θα αξιολογηθεί η κλινική χρηστικότητα μοντέλων πρόβλεψης που εφαρμόζονται μέσω του CDSS (Carobianco, 2020).

Η συστηματική ανασκόπηση τέτοιων συστημάτων, μας οδηγεί στον προσδιορισμό τεσσάρων (4) κυρίως χαρακτηριστικών, ως ανεξάρτητων προγνωστικών παραγόντων εξελιγμένης κλινικής πρακτικής. Αυτά τα χαρακτηριστικά αναγνωρίστηκαν ανεξάρτητα κι από άλλες έρευνες και συνοψίζονται στα εξής: 1) αυτόματη παροχή υποστήριξης αποφάσεων, ως μέρος της ροής εργασίας του κλινικού ιατρού 2) παροχή συστάσεων και όχι απλής αξιολόγησης της κλινικής εικόνας 3) χρήση της υποστήριξης αποφάσεων στον τόπο και χρόνο της λήψης της απόφασης και 4) υποστήριξη αποφάσεων μέσω υπολογιστή.

Καθώς η ταχύτητα επεξεργασίας και η διαθεσιμότητα των ΜΔΥ έχουν αυξηθεί δραματικά, έχουμε εισέλθει σε μια εποχή τεράστιων δυνατοτήτων, στην οποία τα CDSS μπορούν να είναι πιο ισχυρά στους αλγόριθμους τους, πιο προσαρμόσιμα σε νέα δεδομένα, ενσωματωμένα πιο απρόσκοπτα στις υπάρχουσες τεχνολογίες και πιο αποτελεσματικά στη βελτίωση των αποτελεσμάτων των ασθενών.

Οι ΑΗΦΥ αποτελούν πραγματικό θησαυρό πληροφοριών για ένα Σύστημα Υποστήριξης Κλινικών Αποφάσεων (CDSS), αλλά χωρίς επαρκή έρευνα και κατάλληλες συσχετίσεις, τα δεδομένα αυτά ενδέχεται να οδηγήσουν σε αμφισβητήσιμα αποτελέσματα (Carobianco, 2020). Η εκμετάλλευση των δεδομένων από τους ΑΗΦΥ ήταν ως τώρα περιορισμένη, χωρίς να είναι δυνατή η αξιοποίηση ενσωμάτωσης των κατευθυντήριων οδηγιών που βασίζονται σε αποδείξεις (Evidence Based Guidelines - EBG), στη ροή εργασίας του θεράποντος ιατρού.

Με τις εξελίξεις όμως στα δεδομένα πληροφορικής και στους αλγόριθμους ΑΙ, είναι τώρα δυνατός ο συνδυασμός ΜΔΥ και ταχύτερης επεξεργασίας τους, σε εργαλεία υποστήριξης κλινικών αποφάσεων. Η πιο αποτελεσματική μέθοδος ενσωμάτωσης εργαλείων υποστήριξης αποφάσεων που στηρίζονται σε αποδείξεις (Evidence-Based) στη ροή εργασίας του θεράποντος ιατρού, είναι μέσω του Ατομικού Ηλεκτρονικού Φακέλου Υγείας (ΑΗΦΥ). Έτσι, κρίνεται πολύ χρήσιμη η συμπερίληψη, στους ΑΗΦΥ, των κατευθυντήριων οδηγιών που βασίζονται σε αποδείξεις (EBG), για την αποτελεσματική υποστήριξη μιας κλινικής απόφασης (Stopa et al., 2019).

Εξάλλου, υπάρχουν μοντέλα πρόβλεψης τα οποία μπορούν να χρησιμοποιηθούν στα CDSS, για στοχευμένη διενέργεια θεραπειών ή αποφυγή της. Έτσι, μπορεί να προβλεφθεί αν η θεραπεία πρέπει να εφαρμοστεί σε άτομα για τα οποία υπάρχει πολύ υψηλή πιθανότητα οφέλους, ή αν πρέπει να αποφευχθεί, για τα άτομα που υπάρχει μεγάλη πιθανότητα βλάβης. Τα μοντέλα αυτά διενεργούν πρόβλεψη σε επίπεδο ασθενούς (Patient-Level Prediction - PLP) και παρέχουν τα μέσα, ώστε να προβλεφθεί η πιθανότητα να βιώσει κάποιος ασθενής ένα συγκεκριμένο αποτέλεσμα (ευεργετικό ή επιβλαβές), με βάση ένα σύνολο προβλέψεων, όπως τα δημογραφικά του, οι συννοσηρότητες ή άλλοι παράγοντες.

Επειδή κάθε PLP μοντέλο τείνει να εστιάζει σε μια μεμονωμένη πτυχή (κίνδυνο/πιθανότητα δυαδικού αποτελέσματος), ένα μεμονωμένο PLP δεν μπορεί να συνεισφέρει αποτελεσματικά ως εργαλείο CDSS. Ωστόσο, πληροφορίες που προέρχονται από πολλά μοντέλα PLP ταυτόχρονα και παρέχουν ενημέρωση για θεραπευτικούς κινδύνους και οφέλη, μπορούν να χρησιμοποιηθούν επιτυχώς ως εργαλείο CDSS. Η εκπαίδευση των μοντέλων πρόβλεψης με ολοκληρωμένα σύνολα δεδομένων, που περιλαμβάνουν δεδομένα των ασθενών από διοικητικές πηγές, από τον ΑΗΦΥ τους ή δεδομένα του τρόπου ζωής τους, θα κάνει αυτά τα μοντέλα πρόβλεψης ακόμη αποτελεσματικότερα στο μέλλον. (Johnston et al., 2019).

Τα Συστήματα Υποστήριξης Κλινικών Αποφάσεων ενδέχεται να εμπίπτουν στο πεδίο εφαρμογής των κανονισμών, οι οποίοι διέπουν τα ιατρο-τεχνολογικά προϊόντα. Με άλλα λόγια, αν θέλουμε πραγματικά να χρησιμοποιήσουμε οποιοδήποτε αποτέλεσμα προκύπτει από το ένα τέτοιο σύστημα, χωρίς τη μεσολάβηση κάποιου κλινικού ιατρού (ο οποίος διαφορετικά πρέπει να αναλάβει και την ευθύνη για την τελική κλινική απόφαση), τέτοια συστήματα πρέπει να πιστοποιηθούν σαν «Λογισμικό ως Ιατρική Συσκευή» (Software as a Medical Device - SaMD) σύμφωνα με το ισχύοντες κανονισμούς, π.χ κατά ISO 13485:2016 (Coronato, 2018).

6.4 Τεχνητή Νοημοσύνη (Artificial Intelligence)

Η εταιρεία Gartner ορίζει την Τεχνητή Νοημοσύνη (Artificial Intelligence - AI) ως την εφαρμογή προηγμένων τεχνικών ανάλυσης και λογικής, συμπεριλαμβανομένης της Μηχανικής Μάθησης (Machine Learning - ML), για την ερμηνεία γεγονότων, την υποστήριξη και αυτοματοποίηση αποφάσεων και την ανάληψη δράσης. Οι κοινοί ορισμοί της AI επικεντρώνονται στην αυτοματοποίηση, με αποτέλεσμα να μην αποσαφηνίζονται πλήρως όλες οι ευκαιρίες, που προσφέρονται στους επικεφαλής των τμημάτων Πληροφορικής και επιχειρήσεων. Η AI είναι η τεχνολογία που εξομοιώνει την ανθρώπινη συμπεριφορά, τυπικά μαθαίνοντας από αυτή (Gartner, 2019).

Τα τελευταία χρόνια επενδύονται τεράστια χρηματικά ποσά στην AI, στον τομέα της υγειονομικής περίθαλψης, προκειμένου να αντιμετωπιστούν τρεις βασικές προκλήσεις: η καλή υγεία των ασθενών, η ποιότητα της προσφερόμενης σε αυτούς υγειονομικής περίθαλψης και το κόστος της. Μερικές φορές υπάρχει η αντίληψη ότι στόχος της AI είναι να αντικαταστήσει τον άνθρωπο με τεχνητούς λειτουργούς. Αντί αυτού όμως, η AI παρέχει έναν τρόπο επέκτασης και αύξησης της ανθρώπινης δημιουργικότητας, της σκέψης και της ικανότητας προς λήψη αποφάσεων. Στην κλινική πρακτική, για παράδειγμα, βλέπουμε εξαιρετικά ευφυή συστήματα (Intelligent Systems) να υποστηρίζουν τους γιατρούς, προς βελτιστοποίηση της λήψης των αποφάσεών τους, παρέχοντάς τους ευκολότερη και ταχύτερη πρόσβαση σε πληροφορίες, οι οποίες είναι πιο κατάλληλες ή κρίσιμες, αποκρύπτοντάς τους παράλληλα μη συναφή δεδομένα ή γεγονότα (Coronato et al., 2020).

Σύμφωνα με εκτίμηση του Forbes, τα επόμενα χρόνια η AI θα αναπτυχθεί σε όλες τις πτυχές της υγειονομικής περίθαλψης και θα προσφέρει πολύ καλύτερες υπηρεσίες προς τον πελάτη-λήπτη της. Ενώ η ικανότητα του ανθρώπου να αναλύει και να φτάνει σε συμπεράσματα είναι ανώτερη από εκείνη οποιουδήποτε άλλου είδους στον πλανήτη, εξακολουθεί να είναι περιορισμένη, από την άποψη του όγκου των πληροφοριών που μπορεί να επεξεργαστεί πολύ γρήγορα. Η AI επιταχύνει αυτή τη διαδικασία σε πολύ μεγάλο βαθμό και πολύ πιο αποτελεσματικά από ό, τι είναι ανθρώπινως δυνατό. Το σύστημα Watson της IBM, για παράδειγμα, μπορεί να διαβάσει 40 εκατομμύρια έγγραφα σε 15 δευτερόλεπτα (Forbes, 2019).

Αν πρόκειται να διδάξουμε έναν υπολογιστή πώς να παίρνει αποφάσεις, πρέπει πρώτα να σκεφτούμε πώς παίρνει αποφάσεις ένας άνθρωπος. Υπάρχουν δύο βασικοί τρόποι λήψης των περισσότερων αποφάσεων από έναν άνθρωπο: χρησιμοποιώντας συλλογισμούς – λογική και χρησιμοποιώντας την εμπειρία του. Και οι δύο μπορούν να υλοποιηθούν από υπολογιστές και απαντούν στο όνομα Τεχνητή Νοημοσύνη (Serrano, 2020).

Η ΑΙ είναι επομένως το όνομα που δίνεται στη διαδικασία, με την οποία ένας υπολογιστής λαμβάνει αποφάσεις, μιμούμενος τον άνθρωπο. Εν συντομία, οι δύο προαναφερόμενοι βασικοί τρόποι λήψης απόφασης αποτελούν ΑΙ, ενώ ο δεύτερος τρόπος κατά τον οποίο δίνεται έμφαση στη χρησιμοποίηση της προηγούμενης εμπειρίας (προηγούμενων δεδομένων), αποτελεί ένα υποσύνολο της ΑΙ, τη Μηχανική Μάθηση (ML). Για παράδειγμα, ας θεωρήσουμε το πρόβλημα διάγνωσης μιας συγκεκριμένης ίωσης από έναν υπολογιστή-ιατρό, σε δεδομένο ιατρικό περιστατικό, το οποίο εμφανίζει συμπτώματα ίωσης. Υπάρχουν δύο βασικές προσεγγίσεις για να γίνει η διάγνωση. Σε μια προσέγγιση που χρησιμοποιεί συλλογισμούς παρόμοια με την ανθρώπινη λογική, εξετάζονται όλα τα συμπτώματα του περιστατικού, αναλύονται, συνδυάζονται με όλες τις πιθανές ιώσεις που τα περικλείουν και το σύστημα οδηγείται στην επιλογή, ως διάγνωση, εκείνης της ίωσης, που συγκεντρώνει όλα ή τα περισσότερα από τα αναφερόμενα συμπτώματα.

Μια άλλη προσέγγιση όμως θα μπορούσε να οδηγήσει στην επιλογή της ίωσης, μέσα από την παρακολούθηση πλήθους περιστατικών, που παρουσιάζουν τα ίδια συμπτώματα για παρόμοια χρονικά διαστήματα, προ της επιχειρούμενης διάγνωσης, καταγράφοντας ποια ίωση είχε τα αναφερόμενα συμπτώματα και ανακαλύπτοντας μοτίβα (patterns) συσχέτισης, που οδηγούν στη διάγνωση. Με αυτή την προσέγγιση, ο υπολογιστής καταλήγει στην απόφαση επιλογής της ίωσης, πολύ γρηγορότερα και βρίσκει κάποιο συγκεκριμένο μοτίβο διάγνωσης. Προφανώς, η πρώτη προσέγγιση χρησιμοποιεί λογική και συλλογισμούς, ενώ η δεύτερη χρησιμοποιεί προηγούμενη εμπειρία, με τη μορφή δεδομένων που είχαν εισαχθεί το σύστημα. Επομένως, και οι δύο προσεγγίσεις αποτελούν μορφές ΑΙ, ενώ η δεύτερη αποτελεί ML.

Η ΑΙ μπορεί να συνεισφέρει στη μείωση του διοικητικού φόρτου των φορέων υγειονομικής περίθαλψης, αφήνοντας έτσι περισσότερο πολύτιμο χρόνο στη διάθεση του ιατρικού και νοσηλευτικού προσωπικού, για την φροντίδα των ασθενών. Μαζί με άλλες ψηφιακές δυνατότητες και εργαλεία, τα σύγχρονα συστήματα υγείας μετασχηματίζονται και ενισχύονται δραστικά, με απώτερο σκοπό τη βελτίωση της υγείας και της ποιότητας των παρεχόμενων υγειονομικών υπηρεσιών στους ασθενείς. Απαραίτητη προϋπόθεση για αυτό τον μετασχηματισμό είναι η αποδοχή και η εμπιστοσύνη, από το ιατρικό, νοσηλευτικό προσωπικό αλλά και από τους ίδιους τους ασθενείς τους, προς αυτές τις νέες δυνατότητες, καθώς και ο σταδιακός τους «προσωποκεντρικός ψηφιακός γραμματισμός» (person-centered digital literacy), δηλαδή η απόκτηση από μέρους τους τέτοιων ψηφιακών δυνατοτήτων, που τους επιτρέπουν να ζουν, να μαθαίνουν, να εργάζονται, να συμμετέχουν και να διαπρέπουν, μέσα σε μια ψηφιακή κοινωνία (Szegner, 2019).

Σημαντική πρόκληση στον χώρο της ΑΙ αποτελεί η δυνατότητα να εντοπιστεί ο υπεύθυνος για ζημιές που προκαλούνται από μια συσκευή ή υπηρεσία, η οποία λειτουργεί με ΑΙ, για παράδειγμα σε ένα ατύχημα που περιλαμβάνει αυτοκινούμενο ιατρικό εργαλείο. Πρέπει να καλυφθεί η ζημιά από τον ιδιοκτήτη (νοσοκομείο), τον κατασκευαστή του εργαλείου ή τον προγραμματιστή; Εάν ο παραγωγός του εν λόγω εργαλείου είναι απολύτως απαλλαγμένος από την λογοδοσία, ενδέχεται να μην έχει κίνητρο να παρέχει ένα καλό προϊόν ή υπηρεσία και θα μπορούσε να βλάψει την εμπιστοσύνη των ανθρώπων προς την τεχνολογία. Οι σχετικοί κανονισμοί θα μπορούσαν να ρυθμίσουν τέτοια θέματα, όμως από την άλλη πλευρά θα μπορούσαν να είναι πολύ αυστηροί, εμποδίζοντας την καινοτομία (European Parliament News, 2020).

Σε αντίθεση με τις ιατρικές συσκευές, ένα μοναδικό χαρακτηριστικό των εργαλείων ΑΙ είναι η ικανότητά τους να βελτιώνονται συνεχώς με νέα δεδομένα. Αυτή η διαδικασία ονομάζεται σταδιακή (incremental) μάθηση, στην οποία τα αποτελέσματα από ένα εκπαιδευμένο σύστημα ΑΙ ενσωματώνονται σε έναν κλειστό βρόχο ανατροφοδότησης δεδομένων και χρησιμοποιούνται για να τελειοποιήσουν την προγνωστική ακρίβεια του συστήματος, μέσω επαναληπτικής επανεκπαίδευσης του μοντέλου. Αυτό το χαρακτηριστικό διακρίνει τα εκπαιδευόμενα νευρωνικά δίκτυα από τα αμετάβλητα συστήματα μέτρησης (scoring systems) ή από το τυποποιημένο λογισμικό (Ngiam and Khor, 2019).

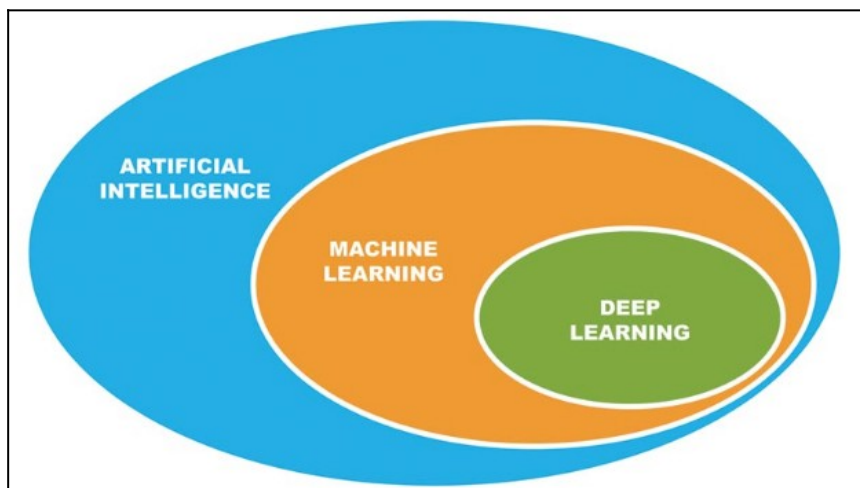
6.5 Μηχανική Μάθηση (Machine Learning)

Σύμφωνα με την εταιρεία SAS, πρωτοπόρο στον τομέα της Ανάλυσης Δεδομένων, η Μηχανική Μάθηση (Machine Learning - ML) είναι μια μέθοδος Ανάλυσης Δεδομένων, που αυτοματοποιεί την ανάπτυξη αναλυτικών μοντέλων. Είναι ένας κλάδος της Τεχνητής Νοημοσύνης (Artificial Intelligence – AI), που βασίζεται στην ιδέα ότι τα συστήματα μπορούν να μάθουν από τα δεδομένα, να προσδιορίσουν τα μοτίβα (patterns) και να λάβουν αποφάσεις, με ελάχιστη ανθρώπινη παρέμβαση. Ενώ η Τεχνητή Νοημοσύνη είναι η ευρύτερη επιστήμη της μίμησης των ανθρώπινων ικανοτήτων, η Μηχανική Μάθηση είναι ένα συγκεκριμένο υποσύνολο της Τεχνητής Νοημοσύνης, που εκπαιδεύει μια μηχανή στο πώς να μαθαίνει (SAS, 2019).

Προσπαθώντας να δώσουμε έναν πολύ γενικό ορισμό, θα λέγαμε ότι η ML είναι μια διαδικασία κατά την οποία ο υπολογιστής λύνει προβλήματα και λαμβάνει αποφάσεις, με παρόμοιο τρόπο που κάνουν και οι άνθρωποι. Η ML μπορεί να βοηθήσει, σχεδόν σε οποιοδήποτε τομέα ανθρώπινης δραστηριότητας, όπου υπάρχει μια εργασία που απαιτεί επανάληψη, εξέταση δεδομένων και συλλογή συμπερασμάτων. Ειδικά τα τελευταία χρόνια, που η υπολογιστική ισχύς έχει αυξηθεί τόσο γρήγορα και τα δεδομένα συλλέγονται και μπορούν να υποστούν επεξεργασία σχεδόν οπουδήποτε, η ML μπορεί να χρησιμοποιηθεί για να βελτιώσει, με τον έναν ή τον άλλο τρόπο, την ανθρώπινη ζωή.

Κάνοντας πιο συγκεκριμένο τον παραπάνω γενικό ορισμό, λέγοντας ML, αναφερόμαστε σε υπολογιστές που λαμβάνουν αποφάσεις με βάση την εμπειρία. Με τον ίδιο τρόπο που οι άνθρωποι λαμβάνουν αποφάσεις βάσει προηγούμενων εμπειριών, οι υπολογιστές μπορούν να λαμβάνουν αποφάσεις, βάσει προηγούμενων δεδομένων. Επομένως, η ML είναι η διαδικασία κατά την οποία ο άνθρωπος διδάσκει τον υπολογιστή πώς να κάνει κάτι, με βάση την εμπειρία κυρίως, και όχι με βάση κάποιες οδηγίες. Οι κανόνες που χρησιμοποιούν οι υπολογιστές για τη λήψη αποφάσεων ονομάζονται μοντέλα (Serrano, 2020).

Από τεχνικής άποψης η ML είναι ουσιαστικά μια αλληλεπίδραση μεταξύ μεγάλων συνόλων δεδομένων με μια συγκεκριμένη κατηγορία μεθόδων μάθησης, που ονομάζεται βαθιά μάθηση (deep learning) ή μάθηση με βαθιά νευρωνικά δίκτυα (deep neural networks). Τα νευρωνικά δίκτυα μπορούν να εκπαιδευτούν, ώστε να είναι εξαιρετικά ακριβή στην αναζήτηση πολύπλοκων μοτίβων, ανάμεσα σε Μαζικά Δεδομένα. Αυτά τα δίκτυα μάθησης μπορούν επίσης να επανεκπαιδευθούν με συγκεκριμένου πληθυσμού σύνολα δεδομένων, και να χρησιμοποιηθούν για πολλαπλές εφαρμογές υγειονομικής περίθαλψης.



Σχ. 3 Σχέση Τεχνητής Νοημοσύνης, Μηχανικής Μάθησης και Βαθιάς Μάθησης (Πηγή: nowmag.gr)

Οι μέθοδοι εκπαίδευσης ενός νευρωνικού δικτύου γενικά περιλαμβάνουν την εποπτευόμενη μάθηση (supervised learning), τη μη εποπτευόμενη μάθηση (unsupervised learning) και υβριδικές προσεγγίσεις των δύο προηγούμενων κατηγοριών. Εκτός της βαθιάς μάθησης υπάρχουν και διάφορες άλλες μέθοδοι ML, όπως οι μηχανές διανυσμάτων υποστήριξης (support vector machines), τα δέντρα αποφάσεων (decision trees) και τα Bayesian δίκτυα, που θα μπορούσαν να αντιμετωπίσουν επαρκώς εργασίες συναφείς με την υγειονομική περίθαλψη. Ωστόσο η βαθιά μάθηση κρίνεται ως η πιο κατάλληλη, για τη διαχείριση δεδομένων με πολλές μεταβλητές.

Καθώς η ML βελτιώνεται, έχει τη δυνατότητα μέσω της εφαρμογής νευρωνικών δικτύων να αναγνωρίζει χαρακτηριστικά εικόνας, χωρίς προ-εκπαίδευση (pre-training), τα οποία μπορούν να αφομοιώσουν διαφορετικά σύνολα κλινικών δεδομένων. Οι προκύπτοντες αλγόριθμοι μπορούν στη συνέχεια να εφαρμοστούν σε παρόμοιες, νέες κλινικές πληροφορίες, για την πρόβλεψη μεμονωμένων περιστατικών ασθενών, με βάση προηγούμενες μεγάλες ομάδες ασθενών. Εναλλακτικά, παρόμοιες τεχνικές μπορούν να εφαρμοστούν σε εικόνες, για τον προσδιορισμό υπο-πληθυσμών, που αλλιώς θα ήταν πολύ περίπλοκο να αναγνωριστούν.

Επιπλέον, η ML μπορεί να εντοπίζει μη αναγνωρισμένα, μοναδικά χαρακτηριστικά εικόνας ή συνδυασμό χαρακτηριστικών, που σχετίζονται με την εξέλιξη της νόσου. Για παράδειγμα, ένα υποσύνολο ασθενών με απώλεια μνήμης, που πιθανώς θα εξελιχθεί σε άνοια, μπορεί να έχουν χαρακτηριστικά ανιχνεύσιμα, πριν από την ανάπτυξη των συμπτωμάτων. Αυτή η προσέγγιση επιτρέπει την εξέταση μεγάλου όγκου πληθυσμού με προοπτική κλινικής παρακολούθησης και ταυτοποίηση των πιο σχετικών ιχνών εικόνας, αντί της ανάλυσης αναδρομικών δεδομένων μικρού όγκου, σε ασθενείς που έχουν ήδη αναπτύξει συμπτωματική εκφυλιστική εγκεφαλική νόσο (Hulsen et al., 2019).

Επανερχόμενοι στο παράδειγμα της διάγνωσης μιας συγκεκριμένης ίωσης από ένα σύστημα υπολογιστή-ιατρό, με βάση τα συμπτώματα ενός ιατρικού περιστατικού, είναι σκόπιμο να εξετάσουμε τον τρόπο λήψης απόφασης ανθρώπου και μηχανής, για να δούμε την εφαρμογή τους στην ML. Ο άνθρωπος γενικά, όταν πρόκειται να προβεί στη λήψη μιας απόφασης, ακολουθεί το τρίπτυχο «Θυμάμαι - Σχηματίζω Κανόνα – Προβλέπω», δηλαδή τα εξής τρία στάδια συλλογισμού: 1) θυμάται προηγούμενες παρόμοιες καταστάσεις, 2) σχηματίζει έναν κανόνα και 3) χρησιμοποιεί τον προαναφερόμενο κανόνα για να προβλέψει τι θα γίνει, αν λάβει τη συγκεκριμένη απόφαση (Serrano, 2020).

Έτσι αν ο άνθρωπος που θα λάβει την απόφαση είναι ένας παιδίατρος, ο οποίος ενημερώνεται τηλεφωνικά από τη μητέρα, ότι το παιδί της παρουσιάζει για τις 3 τελευταίες ημέρες συμπτώματα όπως πυρετό, ρίγη, αδιαθεσία, μυαλγίες, άλγος στο κατώτερο θωρακικό τοίχωμα, πονοκεφάλους, ρινική συμφόρηση και περιστασιακά ναυτία, τότε η διαδικασία για να προβεί σε διάγνωση, θα είναι περίπου η εξής: 1) θυμάται με βάση τις ιατρικές γνώσεις του, ότι τέτοια συμπτώματα παρουσιάζουν συνήθως ασθενείς με εποχική γρίπη, 2) σχηματίζει έναν κανόνα, κατά τον οποίο τις περισσότερες φορές που ένα παιδί έχει τέτοια συμπτώματα, πάσχει από εποχική γρίπη και 3) προβλέπει (διαγιγνώσκει) ότι το παιδί πάσχει από εποχική γρίπη. Μπορεί βεβαίως η διάγνωσή του να είναι λανθασμένη, προσπάθησε όμως να κάνει μια ακριβή, κατά το δυνατόν, διάγνωση – πρόβλεψη.

Στην περίπτωση της μηχανής (υπολογιστή), στόχος της ML είναι να κάνει τον υπολογιστή να σκεφθεί όπως ο άνθρωπος. Στο προηγούμενο λοιπόν παράδειγμα, το σύστημα υπολογιστής-ιατρός λειτουργεί και πάλι υπό το τρίπτυχο «Θυμάμαι - Σχηματίζω Κανόνα – Προβλέπω», ως εξής: 1) θυμάται, πραγματοποιώντας αναζήτηση σε τεράστια σύνολα δεδομένων υγείας, 2) σχηματίζει έναν κανόνα, ελέγχοντας έναν τεράστιο αριθμό κανόνων - συσχετίσεων και αποφασίζοντας ποιος κανόνας ταιριάζει περισσότερο στα συγκεκριμένα δεδομένα και 3) προβλέπει, χρησιμοποιώντας τον προαναφερόμενο κανόνα, για διαγνώσεις σε μελλοντικά δεδομένα.

Όπως κάθε εργαλείο, ένας αλγόριθμος ML πρέπει να είναι κατάλληλα σχεδιασμένος, ώστε να είναι αποτελεσματικός. Στο πλαίσιο της υγειονομικής περίθαλψης, το κλινικό πρόβλημα πρέπει να είναι ταυτόχρονα οδηγός και σημείο αναφοράς για τις εφαρμογές ML. Με το κλινικό πρόβλημα στο επίκεντρο, η ικανότητα της ML να αφομοιώνει και να αναλύει μεγάλα και ποικίλα σύνολα δεδομένων, αποδεικνύεται ανεκτίμητης αξίας βοήθεια για τους κλινικούς ιατρούς, ενόψει της λήψης αποφάσεων, για την περίθαλψη των ασθενών τους.

Για να είναι ένα εργαλείο ML αποτελεσματικό ως προς την παροχή υγειονομικής περίθαλψης, θα πρέπει να ενσωματωθεί σε μια ενιαία πλατφόρμα. Έτσι παρέχεται κεντρικός έλεγχος σε λειτουργίες όπως η προεπεξεργασία δεδομένων, η διαχείριση δεδομένων, οι ρυθμιστικές απαιτήσεις και η επιχειρησιακή αλληλεπίδραση με υφιστάμενα συστήματα Ατομικού Ηλεκτρονικού Φακέλου Υγείας (ΑΗΦΥ).

Σημαντική πρόκληση, σε οποιοδήποτε πρόβλημα ML στον χώρο της υγειονομικής περίθαλψης, είναι η συγκέντρωση υψηλής ποιότητας, αντιπροσωπευτικών και ποικιλόμορφων συνόλων δεδομένων (datasets) υγείας. Στην ιδανική περίπτωση, το μοντέλο ML θα εκπαιδευτεί με δεδομένα που ταιριάζουν ακριβώς με τη μορφή και την ποιότητα των δεδομένων, που θα χρησιμοποιηθούν σε μεταγενέστερο σημείο. Σε κλινικό πλαίσιο, τα δεδομένα αυτά προέρχονται από τους ΑΗΦΥ και συνήθως είναι προβληματικά, αναξιόπιστα και επιρρεπή σε ανεπιθύμητη μεταβλητότητα.

Τα παραπάνω προβλήματα υφίστανται κυρίως επειδή οι ΑΗΦΥ σχεδιάστηκαν για χρέωση και κωδικοποίηση και όχι απαραίτητα για την Ανάλυση των Δεδομένων Υγείας (Health Data Analytics) που περιέχουν ή για βελτίωση της ποιότητας της υγειονομικής περίθαλψης. Επιπλέον, πολλά συστήματα υγείας χρησιμοποιούν συστήματα ΑΗΦΥ, που είναι μερικώς διαλειτουργικά, καθιστώντας δύσκολη την απόκτηση πλήρους ιατρικού ιστορικού των ασθενών. Έτσι η κατάρτιση μοντέλων ML από αυτή την εξαιρετικά «θορυβώδη» πηγή δεδομένων, σε κάποιες περιπτώσεις μπορεί να οδηγήσει ακόμη και σε ανεπιθύμητα αποτελέσματα (Waring et al., 2020).

Μετά από επαρκή εκπαίδευση, οι τεχνικές ML μπορούν να παρέχουν προκαταρκτική εξέταση (pre-screening) εικόνων, προς αναζήτηση περιστατικών με μεγάλη πιθανότητα ασθένειας και επιτρέποντας την ιεράρχησή τους. Απεικονιστικές εξετάσεις, όπως η μαστογραφία, θα μπορούσαν να υποβληθούν σε προ-ανάγνωση (pre-reading) με τη χρήση ML, για τον εντοπισμό των λίγων θετικών περιπτώσεων μεταξύ των πολλών φυσιολογικών μελετών, που επιτρέπουν ταχεία αναγνώριση και θεραπεία. Η προκαταρκτική εξέταση περίπλοκων περιπτώσεων υψηλής οξύτητας, επιτρέπει επίσης την εστιασμένη προσέγγιση για προσδιορισμό και επισκόπηση, κατά προτεραιότητα, περιοχών που προκαλούν μεγάλη ανησυχία. Ο ποσοτικός προσδιορισμός των δομών μπορεί να ενσωματωθεί σε μια εικόνα, όπως το μέγεθος του κακοήθους όγκου, η παρακολούθηση της ανάπτυξης ή η απόκριση στη θεραπεία, για τη διαχείριση της φαρμακευτικής θεραπείας (Hulsen et al., 2019).

Η εκπαίδευση των εργαλείων ML για κλινική εφαρμογή είναι πολύ διαφορετική από την εκπαίδευση αντίστοιχων εργαλείων έρευνας. Τα περισσότερα εργαλεία ML βασίζονται σε εποπτευόμενες (supervised) μεθόδους μάθησης, στις οποίες ταξινομούνται τα δεδομένα σε προκαθορισμένες κατηγορίες. Η ακρίβεια της κλινικής αποτελεσματικότητας των εργαλείων κλινικής ML, προσεγγίζει εκείνη των ρυθμιζόμενων ιατρικών συσκευών (Ngiam and Khor, 2019).

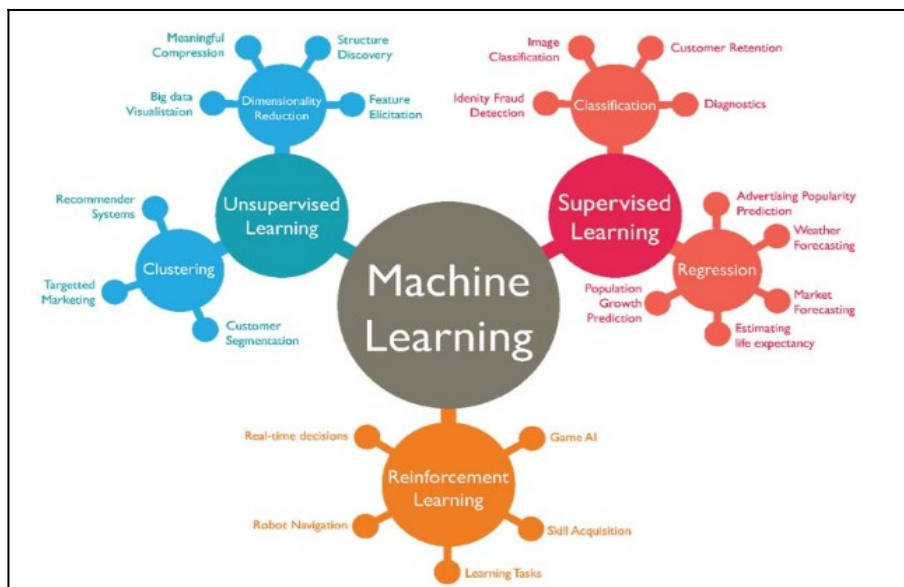
Εν προκειμένω, ο στρατιωτικός ιατρός, έχοντας στα χέρια του πλήθος ιατρικών εξετάσεων του στελέχους-ασθενούς, όπως αυτές συγκεντρώθηκαν από τις υποχρεωτικές ετήσιες και λοιπές έκτακτες κατά περίπτωση υγειονομικές εξετάσεις του, μπορεί με τη βοήθεια κατάλληλων αλγορίθμων ML να διαγνώσει έγκαιρα μεταβολές δεικτών, που σηματοδοτούν τάσεις εμφάνισης κάποιας ασθένειας και να υποβάλλει τον ασθενή του σε θεραπεία ή να τον θέσει υπό στενή παρακολούθηση.

Πρέπει να διευκρινισθεί ότι στην παρούσα εργασία, γίνεται λόγος για στελέχη κι όχι γενικότερα για στρατευσίμους ή προσωπικό των ΕΔ, λόγω της υποχρεωτικότητας υποβολής των πρώτων σε υγειονομικές εξετάσεις ετησίως, κι επομένως, λόγω της δυνατότητας συγκέντρωσης και ανάλυσης Μαζικών Δεδομένων Υγείας (ΜΔΥ). Εκ των πραγμάτων, τα δεδομένα υγείας των κληρωτών, που υπηρετούν μήνες θητείας, δεν επιτρέπουν τη συγκέντρωση και επεξεργασία ΜΔΥ ή τουλάχιστον τέτοιων ποσοτήτων που να μπορούν να υποβληθούν σε Προγνωστική Ανάλυση. Εντούτοις, οι αλγόριθμοι ML μπορούν υπό προϋποθέσεις να επεκταθούν, αναλόγως του εύρους των συγκεντρωθέντων ιατρικών δεδομένων, ενδεχομένως και του ιστορικού, από τον πολιτικό προ της στράτευσης βίο, ακόμη και για στρατευσίμους (κληρωτούς).

Στην περίπτωση αυτή, ιδιαίτερης σημασίας θα ήταν ο έγκαιρος εντοπισμός περιπτώσεων κληρωτών που παρουσιάζουν αυτοκτονικές τάσεις, με κατάλληλη ενσωμάτωση και επεξεργασία των υγειονομικών δεδομένων που υπάρχουν στον ΑΗΦΥ τους, προ της κατάταξής τους στο στράτευμα. Σε πρόσφατη έρευνα (Van Mens et al., 2020), χρησιμοποιήθηκαν αλγόριθμοι ML για τον εντοπισμό αυτοκτονικών τάσεων, από τα ιατρικά δεδομένα, τα οποία συλλέγονται από Γενικούς Ιατρούς.

Η εν λόγω έρευνα πραγματοποιήθηκε με χρήση ιατρικών δεδομένων, προερχομένων από μια Ολλανδική βάση δεδομένων πρωτοβάθμιας περίθαλψης (Nivel Primary Care Database), με 1,5 εκ. ασθενείς. Επιλέχθηκαν N=574 περιπτώσεις που έκαναν τελικώς απόπειρα και N=207.308 ασθενείς με ψυχολογική ευπάθεια. Αποδείχτηκε έτσι, ότι οι τεχνικές ML μπορούν να χρησιμοποιηθούν ως συμπληρωματικό βοήθημα στην ταυτοποίηση και τη διαστρωμάτωση ασθενών, οι οποίοι διατρέχουν κίνδυνο, λόγω παρατηρούμενης αυτοκτονικής συμπεριφοράς. Επίσης, βρέθηκε ότι υπάρχει χρονικό περιθώριο αντίδρασης, μήνες ή εβδομάδες πριν από την απόπειρα αυτοκτονίας, ώστε να αντιληφθεί ο Γενικός Ιατρός, με τη χρήση κατάλληλων αλγορίθμων ML, τον κίνδυνο που διατρέχει ο μελλοντικός αυτόχειρας.

Έτσι, προτάθηκε η δημιουργία αυτόματης ειδοποίησης του Ιατρού, καθώς διαπιστώθηκε ότι το 93% ζήτησε τη γνώμη του Ιατρού ένα χρόνο πριν από την απόπειρα, πάνω από 50% είχαν εγγραφεί ένα μήνα πριν από την απόπειρα, ενώ το 33% είχε εγγραφεί μία εβδομάδα πριν την απόπειρα. Η βελτίωση της τεχνικής που χρησιμοποιήθηκε στην έρευνα, μπορεί ασφαλώς να βελτιωθεί περαιτέρω, με χρήση δεδομένων από την εκπαίδευση και την απασχόληση.



Σχ. 4 Σχηματική αναπαράσταση των κατηγοριών Μηχανικής Μάθησης (Πηγή: nowmag.gr).

Η προσέγγιση της Μηχανικής Μάθησης υποβοηθούμενης από ΑΙ, χρησιμοποιείται ευρέως στην Ιατρική Ακριβείας (Precision Medicine). Το σύστημα «μαθαίνει» από ένα υποσύνολο δεδομένων (training set) χρησιμοποιώντας στατιστικά ή διάφορα μοντέλα, που βασίζονται σε αλγόριθμους και στη συνέχεια επικυρώνει το μοντέλο πρόβλεψης, σε ένα άλλο υποσύνολο δεδομένων δοκιμής (testing set), χρησιμοποιώντας μεγάλη υπολογιστική ισχύ. Το μέγεθος του δείγματος και η σχέση μεταξύ μεταβλητών και σταθερών είναι δύο πολύ σημαντικοί παράγοντες, κατά την επιλογή κατάλληλων αλγορίθμων ML (Khan et al., 2019).

Απαραίτητη προϋπόθεση αποτελεί η διατήρηση των ιατρικών δεδομένων σε ψηφιακή μορφή και ιδανικά στον ΑΗΦΥ του κάθε ασθενούς, που πλέον καθίσταται υποχρεωτικός με νόμο του κράτους (N.4600/2019). Επίσης, η εκπαίδευση των εργαλείων ML πρέπει να γίνεται με ικανό πλήθος πραγματικών (όχι «συνθετικών») δεδομένων ασθενών και με χρήση κατευθυντήριων οδηγιών (guidelines) έμπειρων ιατρών, διαφορετικά είναι πιθανό να οδηγηθούμε σε εξαγωγή ανακριβών και λανθασμένων συμπερασμάτων (Ross and Swetlitz, 2018). Τα κυριότερα μοντέλα ML ομαδοποιούνται σε διαφορετικούς τύπους, ανάλογα με τον τρόπο με τον οποίο λειτουργούν, όπως σχηματικά φαίνεται στο Σχ. 4. Οι κυριότερες οικογένειες μοντέλων ML είναι: η Επιτηρούμενη (Supervised), η Μη Επιτηρούμενη (Unsupervised) και η Ενισχυτική (Reinforcement).

Ας θεωρήσουμε έναν πίνακα με ιατρικά δεδομένα ασθενών, στελεχών των ΕΔ. Κάθε γραμμή του πίνακα αποτελεί ένα σημείο δεδομένων (data point), δηλαδή μια εγγραφή, η οποία αναφέρεται σε έναν συγκεκριμένο ασθενή.

Κάθε ασθενής περιγράφεται από συγκεκριμένα χαρακτηριστικά (features), όπως η ημερομηνία γέννησής του, ο κλάδος των ΕΔ όπου υπηρετεί, η αρτηριακή του πίεση, η θερμομέτρησή του, η τιμή του σακχάρου του κ.λπ. Επομένως, τα χαρακτηριστικά (features) που χαρακτηρίζουν καθέναν από τους ασθενείς του πίνακα είναι οι επικεφαλίδες των αντίστοιχων στηλών του. Κάποια από αυτά τα χαρακτηριστικά έχουν ιδιαίτερη σημασία και καλούνται ετικέτες (labels). Κανονικά, όταν προσπαθούμε να προβλέψουμε ένα χαρακτηριστικό με βάση άλλα χαρακτηριστικά, τότε αυτό το χαρακτηριστικό αποτελεί ετικέτα (label). Αν για παράδειγμα προσπαθούμε να προβλέψουμε αν ένας ασθενής πάσχει από κάποια συγκεκριμένη ασθένεια, με βάση τα συμπτώματά του και άλλες συναφείς πληροφορίες, τότε αυτό το χαρακτηριστικό είναι μια ετικέτα. Επομένως, τα δεδομένα μπορούμε να κατηγοριοποιήσουμε σε δύο μεγάλες κατηγορίες, σε αυτά με ετικέτα (labeled) και σε εκείνα χωρίς ετικέτα (unlabeled).

Το σύνολο αλγορίθμων στους οποίους χρησιμοποιείται ένα σύνολο δεδομένων με ετικέτα (labeled dataset) ονομάζεται Επιτηρούμενη (Supervised) Μάθηση, ενώ το σύνολο αλγορίθμων στους οποίους χρησιμοποιείται ένα σύνολο δεδομένων χωρίς ετικέτα (unlabeled dataset), ονομάζεται Μη Επιτηρούμενη (Unsupervised) Μάθηση.

Στον παρακάτω πίνακα (πηγή: data.world), εμφανίζονται ανωνυμοποιημένα δεδομένα ασθενών, όπου καθένας τους περιγράφεται από συγκεκριμένα χαρακτηριστικά (features), ως επικεφαλίδες των αντίστοιχων στηλών, όπως η ηλικία, το φύλο, ο τύπος προκάρδιου άλγους, η αρτηριακή πίεση, η χοληστερίνη κ.λπ. Κάθε γραμμή του πίνακα είναι ένα σημείο δεδομένων (data point), δηλαδή μια εγγραφή, η οποία αναφέρεται σε έναν συγκεκριμένο ασθενή. Αν υποθέσουμε ότι προσπαθούμε να προβλέψουμε το χαρακτηριστικό της τελευταίας στήλης «Καρδιοπάθεια» (Heart Disease) με βάση τα υπόλοιπα 13 χαρακτηριστικά, τότε αυτό το χαρακτηριστικό αποτελεί ετικέτα (label).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Age	Sex	Chest pain type	BP	Cholesterol	FBS over 120	EKG results	Max HR	Exercise angina	ST depression	Slope of ST	No vessels fluro	Thallium	Heart Disease
2	70	1	4	130	322	0	2	109	0	2.4	2	3	3	Presence
3	67	0	3	115	564	0	2	160	0	1.6	2	0	7	Absence
4	57	1	2	124	261	0	0	141	0	0.3	1	0	7	Presence
5	64	1	4	128	263	0	0	105	1	0.2	2	1	7	Absence
6	74	0	2	120	269	0	2	121	1	0.2	1	1	3	Absence
7	65	1	4	120	177	0	0	140	0	0.4	1	0	7	Absence
8	56	1	3	130	256	1	2	142	1	0.6	2	1	6	Presence
9	59	1	4	110	239	0	2	142	1	1.2	2	1	7	Presence
10	60	1	4	140	293	0	2	170	0	1.2	2	2	7	Presence
11	63	0	4	150	407	0	2	154	0	4	2	3	7	Presence
12	59	1	4	135	234	0	0	161	0	0.5	2	0	7	Absence
13	53	1	4	142	226	0	2	111	1	0	1	0	7	Absence
14	44	1	3	140	235	0	2	183	0	0	1	0	3	Absence
15	61	1	1	134	234	0	0	145	0	2.6	2	2	3	Presence
16	57	0	4	128	303	0	2	159	0	0	0	1	3	Absence
17	71	0	4	112	149	0	0	125	0	1.6	2	0	3	Absence
18	46	1	4	140	311	0	0	120	1	1.8	2	2	7	Presence
19	53	1	4	140	203	1	2	155	1	3.1	3	0	7	Presence
20	64	1	1	110	211	0	2	144	1	1.8	2	0	3	Absence
21	40	1	1	140	199	0	0	178	1	1.4	1	0	7	Absence
22	67	1	4	120	229	0	2	129	1	2.6	2	2	7	Presence
23	48	1	2	130	245	0	2	180	0	0.2	2	0	3	Absence
24	43	1	4	115	303	0	0	181	0	1.2	2	0	3	Absence
25	47	1	4	112	204	0	0	143	0	0.1	1	0	3	Absence
26	54	0	2	132	288	1	2	159	1	0	1	1	3	Absence
27	48	0	3	130	275	0	0	139	0	0.2	1	0	3	Absence
28	46	0	4	138	243	0	2	152	1	0	2	0	3	Absence
29	51	0	3	120	295	0	2	157	0	0.6	1	0	3	Absence
30	58	1	3	112	230	0	2	165	0	2.5	2	1	7	Presence
31	71	0	3	110	265	1	2	130	0	0	1	1	3	Absence
32	57	1	3	128	229	0	2	150	0	0.4	2	1	7	Presence

Σχ. 5 Πίνακας με ανωνυμοποιημένα δεδομένα ασθενών, για πρόβλεψη της στήλης N «Καρδιοπάθεια» (Πηγή: data.world).

Στην Ενισχυτική (Reinforcement) Μάθηση, ένα πρόγραμμα υπολογιστή αλληλεπιδρά με ένα δυναμικό περιβάλλον, στο οποίο πρέπει να επιτευχθεί ένας συγκεκριμένος στόχος (όπως ο χειρισμός ενός ιατρικού μηχανήματος), χωρίς κάποιος δάσκαλος να του λέει ρητά, αν έχει φτάσει κοντά στο στόχο του. Σε αντίθεση με την Επιτηρούμενη ML, αποτελεί μια προσέγγιση που βασίζεται στη μάθηση, την κατευθυνόμενη από το στόχο (on goal-directed learning). Η μάθηση πραγματοποιείται μέσω αλληλεπίδρασης με τα γύρω περιβάλλοντα, παρατηρώντας αλλαγές της κατάστασης.

Πρόσφατα, η ημι-Επιτηρούμενη (Semi-Supervised) ML έχει εισαχθεί ως υβριδική προσέγγιση μεταξύ της Επιτηρούμενης και της μη Επιτηρούμενης ML, κατά την οποία ο αλγόριθμος εκπαιδεύεται βάσει ενός συνδυασμού δεδομένων με ή χωρίς ετικέτα και είναι κατάλληλος για σενάρια, όπου υπάρχει έλλειψη αποτελεσμάτων σε συγκεκριμένα περιβάλλοντα (Jiang et al., 2017).

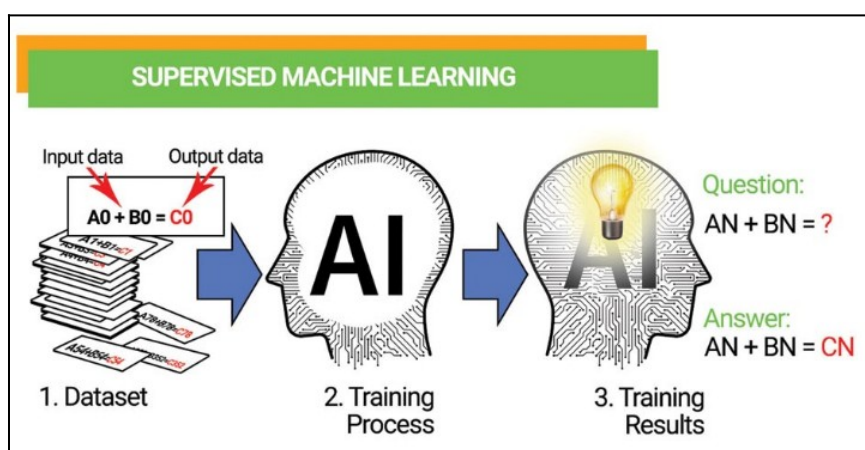
Με τις δυνατότητες της ML, οι εφαρμογές υγειονομικής περίθαλψης είναι απεριόριστες. Ορισμένες από τις εφαρμογές που αναπτύσσονται, επί του παρόντος, εφαρμόζονται ήδη στην κλινική πρακτική (Coronato et al., 2020). Τέτοια παραδείγματα αποτελούν η παροχή βοήθειας σε γιατρούς και ακτινολόγους για να κάνουν ακριβείς διαγνώσεις (IBM Watson Health), η διενέργεια προβλέψεων για το ποιες θεραπευτικές αγωγές είναι πιθανότερο να λειτουργούν ως πιο αποτελεσματικά φάρμακα (Atomwise), καθώς και η εξόρυξη ιατρικών αρχείων δεδομένων, για τη βελτίωση της παροχής υπηρεσιών υγειονομικής περίθαλψης (Google DeepMind Health).

6.5.1 Επιτηρούμενη Μηχανική Μάθηση (Supervised Machine Learning)

Η Επιτηρούμενη (Supervised) Μηχανική Μάθηση είναι ο τύπος ML, που βρίσκεται στις περισσότερες εφαρμογές σήμερα, συμπεριλαμβανομένων εφαρμογών αναγνώρισης εικόνας, επεξεργασίας κειμένου, συστημάτων υποβολής προτάσεων κ.λπ. Αποτελεί έναν τύπο Προγνωστικής (Predictive) Μηχανικής Μάθησης, στην οποία τα δεδομένα συνοδεύονται από ετικέτες, όπου η ετικέτα είναι ο «στόχος» που μας ενδιαφέρει να προβλέψουμε (Serrano, 2020).

Στην περίπτωση του πίνακα των ασθενών που παρουσιάστηκε παραπάνω, το μοντέλο ML θα χρησιμοποιήσει προηγούμενα δεδομένα, για να προβλέψει την ετικέτα κάποιας νέας εγγραφής (data point) ασθενούς. Δηλαδή, αν εισαχθεί μια νέα εγγραφή ασθενούς στον πίνακά μας, χωρίς την ετικέτα «Καρδιοπάθεια» (Heart Disease), το μοντέλο ML θα προσπαθήσει να προβλέψει αν ο νέος ασθενής (data point) παρουσιάζει ή όχι «Καρδιοπάθεια», με βάση προηγούμενα δεδομένα (π.χ. συμπτώματα ή αποτελέσματα ιατρικών εξετάσεων).

Όπως προαναφέρθηκε, η μηχανή λειτουργεί για τη λήψη απόφασης όπως και ο άνθρωπος, δηλαδή υπό το τρίπτυχο «Θυμάμαι - Σχηματίζω Κανόνα – Προβλέπω». Έτσι περίπου λειτουργεί και το μοντέλο Επιτηρούμενης ML. Δηλαδή, το μοντέλο «θυμάται» ένα σύνολο δεδομένων όπως συμπτώματα, ιατρικές μετρήσεις, αποτελέσματα απεικονιστικών εξετάσεων κ.λπ., «σχηματίζει» ένα μοντέλο ή έναν κανόνα, σύμφωνα με τον οποίο ο ασθενής «πάσχει από» την x ασθένεια, όταν παρουσιάζει τα y συμπτώματα και έχει τα z αποτελέσματα εξετάσεων. Όταν εισάγεται μια νέα εγγραφή ασθενούς (data point) στο μοντέλο, αυτό «προβλέπει» σχετικά με την ετικέτα π.χ. «πάσχει από» την συγκεκριμένη ασθένεια, που πρέπει να αποδοθεί στη νέα εγγραφή. Στόχος λοιπόν του μοντέλου είναι να «μάθει» έναν γενικό κανόνα, προκειμένου να αντιστοιχίσει τις εισόδους με τα αποτελέσματα.



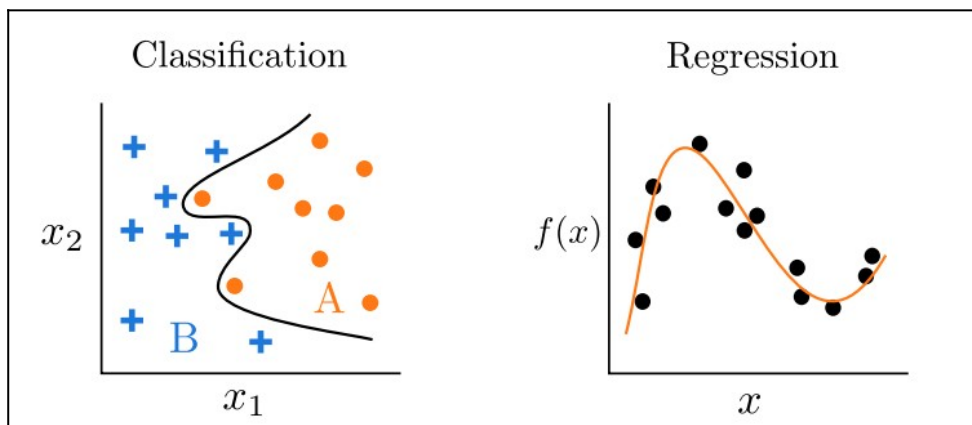
Σχ. 6 Σχηματική αναπαράσταση της Επιτηρούμενης Μηχανικής Μάθησης (Πηγή: nowmag.gr).

Στο παράδειγμα του πίνακα που παρουσιάστηκε παραπάνω, υπάρχουν δύο τύποι συνόλων δεδομένων (datasets), ένας τύπος στον οποίο οι ετικέτες είναι αριθμοί, όπως για παράδειγμα η αρτηριακή πίεση (BP) και ένας δεύτερος τύπος, στον οποίο οι ετικέτες είναι καταστάσεις (states) ή τάξεις (classes), όπως για παράδειγμα η ετικέτα «Καρδιοπάθεια», όπου η τιμή της μπορεί να είναι η κατάσταση «Παρουσία» (Presence) ή «Απουσία» (Absence) της καρδιοπάθειας.

Έτσι προκύπτουν και δύο κατηγορίες μοντέλων Επιτηρούμενης ML, ως εξής: 1) Παλινδρόμησης (Regression), όπου τα μοντέλα προβλέπουν αριθμούς (όπως η αρτηριακή πίεση) και 2) Ταξινόμησης (Classification), όπου τα μοντέλα προβλέπουν καταστάσεις (όπως η παρουσία ή απουσία ασθένειας από την οποία πάσχει ή δεν πάσχει ο ασθενής).

Η έξοδος (output) ενός μοντέλου Παλινδρόμησης (Regression) λέγεται συνεχής (continuous), καθώς η πρόβλεψη μπορεί να έχει οποιαδήποτε πραγματική τιμή, που επιλέγεται από ένα συνεχές διάστημα. Η έξοδος ενός μοντέλου Ταξινόμησης (Classification) λέγεται διακριτή (discrete), καθώς η πρόβλεψη μπορεί να είναι μια τιμή από μια πεπερασμένη λίστα. Αξιοσημείωτο είναι το γεγονός ότι η έξοδος μπορεί να αναφέρεται σε περισσότερες από δύο καταστάσεις. Αν είχαμε περισσότερες καταστάσεις, για παράδειγμα ένα μοντέλο που προβλέπει εάν η ασθένεια από την οποία πάσχει ο ασθενής είναι a, b, ή c, μπορούμε να χρησιμοποιήσουμε το διακριτό (discrete) μοντέλο. Αυτά τα μοντέλα ονομάζονται διακριτά μοντέλα πολλαπλών παραλλαγών (multivariate discrete). Υπάρχουν ταξινομητές με πολλές καταστάσεις (states), αλλά πρέπει πάντα ο αριθμός τους να είναι πεπερασμένος (Serrano, 2020).

Παράδειγμα μοντέλου Παλινδρόμησης (Regression) θα μπορούσε να είναι η πρόβλεψη της τιμής ενός γενόσημου φαρμάκου, έναντι του αντίστοιχου πρωτοτύπου (κατά κανόνα χαμηλότερη). Στο μοντέλο αυτό, κάθε data point είναι ένα σκεύασμα με την ίδια δραστική ουσία, με τα ίδια χαρακτηριστικά δοσολογίας, αποτελεσματικότητας, ποιότητας και τρόπου χορήγησης, με το εγκεκριμένο πρωτότυπο φάρμακο. Η ετικέτα κάθε γενόσημου είναι η τιμή του. Σκοπός του μοντέλου είναι κάθε φορά που ένα νέο γενόσημο (data point) είναι έτοιμο προς κυκλοφορία στην αγορά, να διενεργείται πρόβλεψη της ετικέτας του, δηλαδή της τιμής του.



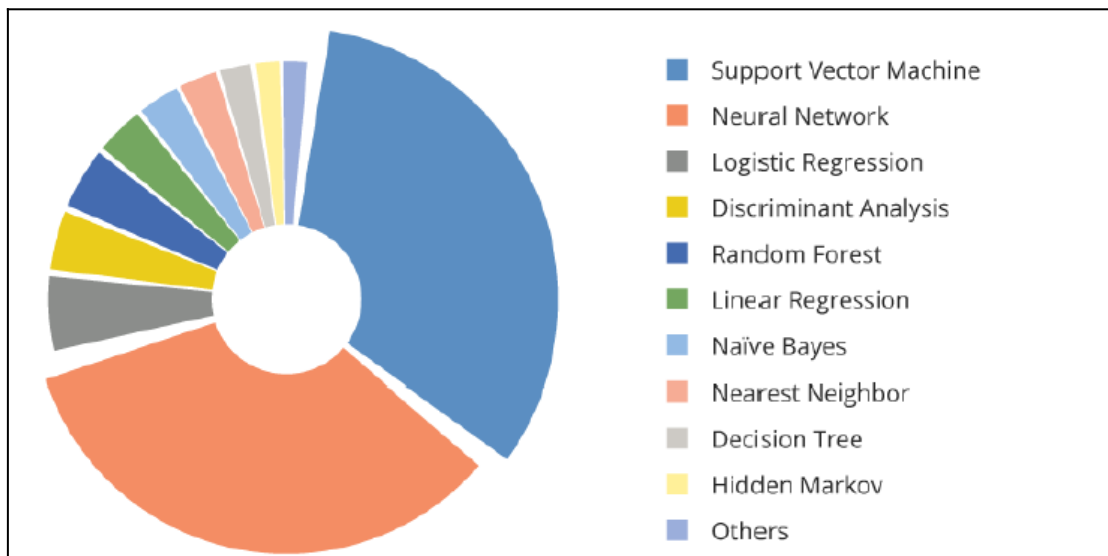
Σχ. 7 Οι δυο σημαντικότερες κατηγορίες μοντέλων Επιτηρούμενης ML (Πηγή: Beck and Kurz, 2020).

Τα μοντέλα Παλινδρόμησης προβλέπουν έναν αριθμό μέσα από τα χαρακτηριστικά (features). Στο παράδειγμα της πρόβλεψης της τιμής του γενόσημου φαρμάκου, τα χαρακτηριστικά του μπορεί να είναι η δραστική ουσία, η δοσολογία, η αποτελεσματικότητα, οι παρενέργειες κ.λπ. Η πιο συνηθισμένη μέθοδος Παλινδρόμησης είναι η Γραμμική Παλινδρόμηση (Linear Regression), κατά την οποία χρησιμοποιούνται γραμμικές συναρτήσεις, για τη διενέργεια προβλέψεων, που βασίζονται στα χαρακτηριστικά (features).

Παράδειγμα μοντέλου Ταξινόμησης (Classification) θα μπορούσε να είναι ο έλεγχος της καταλληλότητας, ως προς τη συνέχιση προσφοράς των υπηρεσιών του στο στράτευμα, ενός στελέχους που νόσησε ή τραυματίστηκε σοβαρά, σωματικά ή ψυχικά. Στο μοντέλο αυτό, κάθε data point είναι ένας ασθενής-στέλεχος των ΕΔ που υπέφερε από μια ασθένεια ή έναν τραυματισμό. Η ετικέτα για κάθε data point είναι είτε «ικανός» είτε «μη ικανός» προς ανάληψη υπηρεσίας.

Στόχος (target) του μοντέλου είναι να προβλέπει, κάθε φορά που ένα νέο data point εισέρχεται σε αυτό, ποια είναι η ετικέτα καταλληλότητάς του, δηλαδή αν είναι «ικανός» ή «μη ικανός» να αναλάβει εκ νέου υπηρεσία. Τα μοντέλα Ταξινόμησης μπορούν να προβλέπουν μια κατάσταση, από ένα πεπερασμένο σύνολο καταστάσεων. Τα πιο συνηθισμένα κάνουν προβλέψεις παρέχοντας ως απάντηση ένα «ναι» ή «όχι» (ικανός, μη ικανός), αλλά υπάρχουν μοντέλα που χρησιμοποιούν μεγαλύτερο σύνολο καταστάσεων.

Στο παράδειγμα της πρόβλεψης της ασθένειας από την οποία πάσχει το στέλεχος, η πρόβλεψη γίνεται με χρήση των χαρακτηριστικών (features), όπως συμπτώματα, αποτελέσματα εξετάσεων, διαγνωστικών ελέγχων κ.λπ. Συνηθισμένες μέθοδοι Ταξινόμησης είναι η Λογιστική Παλινδρόμηση (Logistic Regression), τα Δέντρα Αποφάσεων (Decision Trees), οι Ταξινομητές naïve Bayes, οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines) και τα Νευρωνικά Δίκτυα (Neural Networks).



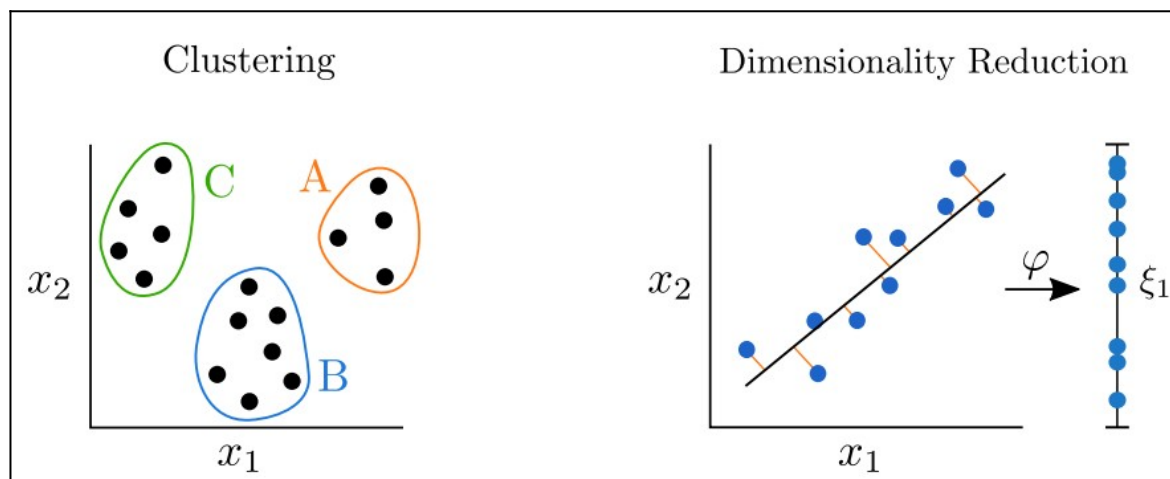
Σχ. 8 Κυριότερες Μέθοδοι Επιτηρούμενης Μηχανικής Μάθησης (Πηγή: Jiang et al., 2017).

6.5.2 Μη Επιτηρούμενη Μηχανική Μάθηση (Unsupervised Machine Learning)

Μια άλλη μορφή ML, η οποία χρησιμοποιείται αρκετά συχνά, είναι η μη Επιτηρούμενη (Unsupervised) Μηχανική Μάθηση. Διαφέρει από την Επιτηρούμενη (Supervised) στο ό,τι τα δεδομένα δεν έχουν ετικέτες, δηλαδή έχουν μόνο χαρακτηριστικά (features) και δεν υπάρχει ο στόχος πρόβλεψης (target). Στο παράδειγμα των γενόσημων φαρμάκων που προαναφέρθηκε, αν στο σύνολο των δεδομένων δεν αναφέρεται η χρηματική αξία του φαρμάκου, τότε πρόκειται για ένα σύνολο δεδομένων χωρίς ετικέτες.

Ωστόσο, ακόμη και με τέτοια δεδομένα, τα οποία στερούνται ετικέτας, υπάρχουν αρκετές δυνατότητες εξόρυξης, μέσα από σύνολα δεδομένων, με κατάλληλες τεχνικές ML. Έτσι στο παράδειγμα του συνόλου δεδομένων με ασθενείς στελέχη των ΕΔ, αν δεν υπάρχουν ετικέτες, έχουμε μεν το σύνολο των ασθενών, δεν γνωρίζουμε όμως, από ποια συγκεκριμένη ασθένεια πάσχει ο καθένας. Το μοντέλο μπορεί να μάς πληροφορήσει για το αν δύο ή περισσότεροι ασθενείς του παρουσιάζουν παρόμοια κλινικά συμπτώματα ή διαφοροποιούνται από άλλους ασθενείς του συνόλου δεδομένων (dataset). Έτσι, είναι εφικτή η ομαδοποίηση των ασθενών σε διακριτές ομάδες, με βάση την κοινή για παράδειγμα συμπτωματολογία τους, έστω κι αν δεν γνωρίζουμε ακόμη, την κοινή ασθένεια από την οποία πάσχουν. Ακόμη κι αν το σύνολο δεδομένων περιέχει ετικέτες, θα μπορούσαν να χρησιμοποιηθούν τεχνικές μη Επιτηρούμενης ML, προκειμένου να πετύχουμε προεργασία και αποτελεσματικότερη εφαρμογή των τεχνικών της Επιτηρούμενης ML. Χωρίς λοιπόν να παρέχεται κάποια εμπειρία στον αλγόριθμο μάθησης, πρέπει να μπορέσει να βρει τη δομή των δεδομένων εισόδου. Η μη Επιτηρούμενη ML μπορεί να είναι αυτοσκοπός (ανακαλύπτοντας κρυμμένα μοτίβα σε δεδομένα) ή μέσο για ένα τέλος (χαρακτηριστικό της μάθησης).

Οι δύο κυριότεροι κλάδοι της μη Επιτηρούμενης ML είναι η Συσταδοποίηση (Clustering), δηλαδή η ομαδοποίηση των δεδομένων με βάση τις ομοιότητες που παρουσιάζουν και η Μείωση Διαστασιμότητας (Dimensionality Reduction), δηλαδή η απλοποίηση των δεδομένων, αντιστοιχίζοντάς τα σε ένα χώρο λιγότερων διαστάσεων, περιγράφοντάς τα δηλαδή με λιγότερα χαρακτηριστικά, χωρίς όμως να χάνουν πολύ από την γενικότητά τους.



Σχ. 9 Οι δύο κυριότεροι κλάδοι της μη Επιτηρούμενης ML (Πηγή: Beck and Kurz, 2020).

Στην περίπτωση του παραδείγματος του συνόλου δεδομένων των γενόσημων φαρμάκων χωρίς χρηματική αξία, μπορούμε να τα ομαδοποιήσουμε με βάση τη δραστική τους ουσία, τα χαρακτηριστικά δόσολογίας τους, την αποτελεσματικότητά τους, την ποιότητά τους ή ακόμη και με βάση κάποιον συνδυασμό των παραγόντων αυτών. Η ομαδοποίηση μεγάλων συνόλων δεδομένων, με βάση ποικίλα χαρακτηριστικά τους γνωρίσματα, μπορεί να γίνει ταχύτατα από τους σύγχρονους ηλεκτρονικούς υπολογιστές.

Στο παραπάνω σύνολο δεδομένων των γενόσημων φαρμάκων, μπορεί να εφαρμοστεί και η δεύτερη τεχνική της μη Επιτηρούμενης ML, δηλαδή η Μείωση Διαστασιμότητας. Έστω ότι στο σύνολο αυτό, στόχος είναι η πρόβλεψη της τιμής ενός γενόσημου φαρμάκου και τα διαθέσιμα χαρακτηριστικά κάθε φαρμάκου είναι: 1) η δραστική του ουσία, 2) η δόσολογία του, 3) η αποτελεσματικότητά του, 4) η ποιότητά του και 5) ο τρόπος χορήγησής του. Αν προσπαθήσουμε να απλοποιήσουμε αυτό το σύνολο δεδομένων με λιγότερες στήλες (χαρακτηριστικά), χωρίς ωστόσο να χαθεί η αξιοπιστία της παρεχόμενης από αυτά πληροφορίας, θα διαπιστώσουμε ότι είναι δυνατή η ομαδοποίηση των χαρακτηριστικών γνωρισμάτων σε μικρότερες και γενικότερες κατηγορίες. Έτσι διαπιστώνουμε ότι τα χαρακτηριστικά 1, 3 και 4 σχετίζονται με την ποιότητα του φαρμάκου, ενώ τα χαρακτηριστικά 2 και 5, σχετίζονται με τον τρόπο λήψης του. Έτσι τα πέντε προαναφερόμενα χαρακτηριστικά μπορούν να συρρικνωθούν και να αποτελέσουν δύο γενικότερες ομάδες χαρακτηριστικών, δηλαδή την «Ποιότητα» και τον «Τρόπο Λήψης» του γενόσημου φαρμάκου.

Ο αριθμός των στηλών, οι οποίες αναφέρονται στα προαναφερόμενα χαρακτηριστικά, αποτελούν τη διάσταση (dimension) του αντίστοιχου συνόλου δεδομένων των γενόσημων φαρμάκων, για αυτό και η μείωση των χαρακτηριστικών, άρα των στηλών, αποτελεί μείωση της διαστασιμότητας του συνόλου. Οι αλγόριθμοι Μείωσης Διαστασιμότητας μπορούν να ανακαλύπτουν τρόπους να ομαδοποιούν τα χαρακτηριστικά γνωρίσματα των συνόλων δεδομένων, με όσο το δυνατόν λιγότερη απώλεια πληροφορίας και διατηρώντας τα όσο το δυνατόν πιο ανέπαφα, ενώ παράλληλα μπορούν να τα απλοποιούν, για ευκολότερη επεξεργασία και αποθήκευση (Serrano, 2020).

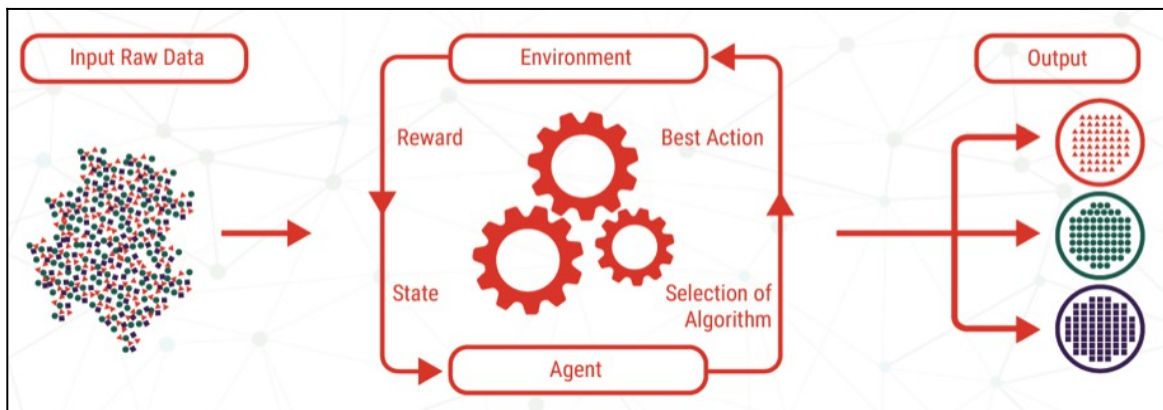
Από τα παραπάνω προκύπτει ότι οι δύο βασικοί κλάδοι της μη Επιτηρούμενης ML δεν διαφέρουν πολύ μεταξύ τους. Σε ένα σύνολο δεδομένων, δηλαδή σε ένα πίνακα με δεδομένα, καθεμιά από τις γραμμές του αποτελεί κι ένα σημείο δεδομένων (data point), ενώ κάθε στήλη αποτελεί χαρακτηριστικό (feature). Έτσι, με την Συσταδοποίηση επιτυγχάνεται ομαδοποίηση των γραμμών του πίνακα, ενώ με τη Μείωση Διαστασιμότητας επιτυγχάνεται ομαδοποίηση των στηλών του πίνακα. Με την παραγοντοποίηση πίνακα (matrix factorization), είναι δυνατή η ταυτόχρονη συρρίκνωση γραμμών και στηλών του πίνακα. Χρησιμοποιώντας γραμμική άλγεβρα, είναι δυνατόν να εκφραστεί ο μεγάλος πίνακας του συνόλου δεδομένων σε δύο μικρότερους πίνακες. Στην περίπτωση ενός μεγάλου πίνακα, όπου κάθε γραμμή είναι ένας ασθενής-λήπτης κάποιου γενόσημου φαρμάκου, κάθε στήλη είναι ένα γενόσημο φάρμακο και κάθε εισαγωγή στον πίνακα είναι η αξιολόγηση, που κάθε ασθενής-λήπτης έδωσε στο αντίστοιχο φάρμακο.

Με την παραγοντοποίηση πίνακα μπορεί να προβλεφθεί ο βαθμός αξιολόγησης που ο ασθενής θα δώσει στο γενόσημο φάρμακο, με βάση χαρακτηριστικά όπως ο τύπος του φαρμάκου, οι δραστικές του ουσίες, οι παρενέργειές του κ.λπ.

6.5.3 Ενισχυτική Μηχανική Μάθηση (Reinforcement Machine Learning)

Η Ενισχυτική ML είναι ένας διαφορετικός τύπος μηχανικής μάθησης, στον οποίο ενώ δεν παρέχονται δεδομένα, το πρόβλημα πρέπει τελικά να λυθεί χωρίς αυτά. Αντί δεδομένων, παρέχεται ένα περιβάλλον (environment ή controlled system) και ένας πράκτορας (agent ή controller), που υποτίθεται ότι πρέπει να πλοηγηθεί σε αυτό το περιβάλλον. Ο πράκτορας έχει έναν στόχο ή ένα σύνολο στόχων. Το περιβάλλον περιλαμβάνει ανταμοιβές (rewards) και ποινές (punishments), οι οποίες καθοδηγούν τον πράκτορα στο να λαμβάνει τις σωστές αποφάσεις, για να επιτύχει τον στόχο ή το σύνολο των στόχων του (Serrano, 2020).

Ειδικότερα, ο πράκτορας επιλέγει ενέργειες ή σήματα ελέγχου και το περιβάλλον αλλάζει κατά συνέπεια την κατάστασή του, επιστρέφοντας ανταμοιβές ή ποινές, δηλαδή ειδικές αριθμητικές τιμές, είτε θετικές είτε αρνητικές. Ο στόχος του πράκτορα είναι η μεγιστοποίηση των ανταμοιβών που λαμβάνονται με την πάροδο του χρόνου. Μια εργασία είναι μια πλήρης προδιαγραφή ενός περιβάλλοντος, το οποίο περιλαμβάνει τον τρόπο παραγωγής της ανταμοιβής. Ένα σύστημα Ενισχυτικής ML περιλαμβάνει: 1) την πολιτική, 2) το σήμα ανταμοιβής, 3) τη συνάρτηση τιμής και 4) το μοντέλο.



Σχ. 10 Σχηματική αναπαράσταση της Ενισχυτικής ML (Πηγή: Van Loon, 2020).

Η πολιτική είναι ένα σύνολο κανόνων διέγερσης-απόκρισης, οι οποίοι αντιστοιχίζουν κάθε κατάσταση του περιβάλλοντος σε μια λίστα ενεργειών, από τις οποίες ο πράκτορας μπορεί ελεύθερα να επιλέξει μία. Στην απλούστερη περίπτωση, η πολιτική μπορεί να εφαρμοστεί από έναν πίνακα αναζήτησης (look up table). Το σήμα ανταμοιβής κωδικοποιεί τον στόχο για τον πράκτορα. Καθώς το περιβάλλον ανταποκρίνεται σε ενέργειες που αλλάζουν την κατάστασή του και παράγουν ανταμοιβή, ο πράκτορας προσπαθεί να μεγιστοποιήσει το άθροισμα των ανταμοιβών, που συλλέγονται μετά την αλληλεπίδραση. Η ανταμοιβή αντιπροσωπεύει έτσι τον κύριο μηχανισμό τροποποίησης της πολιτικής.

Σε περίπτωση που μια ενέργεια που επιλέγεται από τον πράκτορα (η οποία ακολουθεί μια συγκεκριμένη πολιτική), οδηγεί σε κακή ανταμοιβή δηλαδή σε ποινή, τότε ο πράκτορας μπορεί να αποφασίσει να επιλέξει μια άλλη ενέργεια, για την ίδια κατάσταση περιβάλλοντος στο μέλλον, ώστε να εξερευνήσει διαφορετικές δυνατότητες. Ένα σήμα ανταμοιβής, γενικά, μπορεί να είναι μια στοχαστική συνάρτηση της κατάστασης του περιβάλλοντος και της δράσης, που επιλέγεται από τον πράκτορα. Αν και το σήμα ανταμοιβής είναι μια ένδειξη για το πόσο καλή ήταν η επιλογή της τελευταίας δράσης, είναι γενικά καλύτερο να ακολουθούνται στρατηγικές, που οδηγούν σε μεγαλύτερο όφελος μεσοπρόθεσμα και μακροπρόθεσμα. Η συνάρτηση τιμής σε ένα πρόβλημα Ενισχυτικής ML ορίζει «τι είναι καλό μακροπρόθεσμα». Συγκεκριμένα, η τιμή μιας συγκεκριμένης κατάστασης είναι το άθροισμα των ανταμοιβών, που ο πράκτορας υποτίθεται ότι συλλέγει με την πάροδο του χρόνου, απομακρυνόμενος από αυτήν την κατάσταση.

Επομένως, οι ενέργειες επιλέγονται εξετάζοντας τις υψηλότερες τιμές και όχι τις υψηλότερες ανταμοιβές. Ο υπολογισμός τους ωστόσο είναι δυσκολότερος, επειδή οι ανταμοιβές λαμβάνονται αμέσως από το περιβάλλον, ενώ οι τιμές πρέπει να προβλέπονται από τον πράκτορα, εξετάζοντας τις αλληλεπιδράσεις που έχει βιώσει στο παρελθόν με το περιβάλλον. Η αποτελεσματική εκτίμηση τιμής είναι το κεντρικό πρόβλημα σχεδόν όλων των αλγορίθμων Ενισχυτικής ML.

Το τέταρτο στοιχείο ενός συστήματος Ενισχυτικής ML είναι το μοντέλο, το οποίο περιγράφει τη λειτουργία του περιβάλλοντος. Για παράδειγμα, κοιτάζοντας μια συγκεκριμένη κατάσταση και έχοντας επιλέξει μία από τις ενέργειες που είναι δυνατό να επιλεχθούν σε αυτήν την κατάσταση, το μοντέλο μπορεί να αξιοποιηθεί για να προβλέψει τη μελλοντική κατάσταση και την προκύπτουσα επιβράβευση.

Οι μέθοδοι που χρησιμοποιούν μοντέλα για την επίλυση προβλημάτων Ενισχυτικής ML είναι γνωστές ως μέθοδοι βάσει μοντέλου (model-based methods). Αντίθετα, οι μέθοδοι χωρίς μοντέλα (model-free method) είναι αυτές που βασίζονται σε ρητή δοκιμή και σφάλμα, για να μάθουν τη βέλτιστη πολιτική.

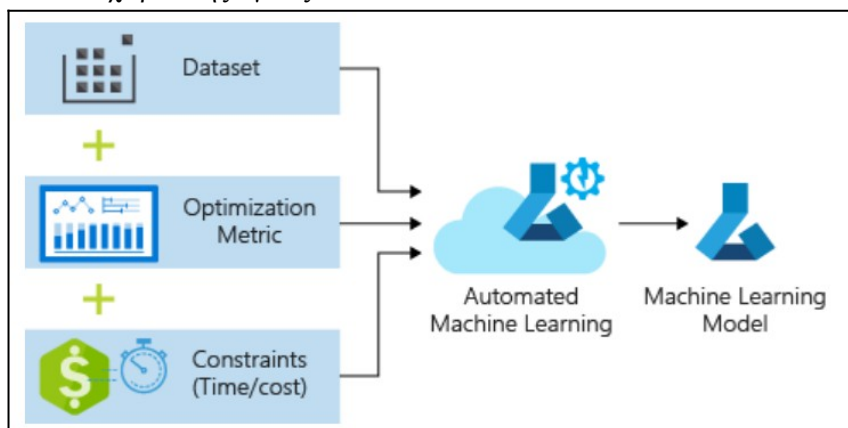
Στον τομέα της υγειονομικής περίθαλψης, η Ενισχυτική ML χρησιμοποιείται στην Ιατρική Ακριβείας, σε διαγνωστικά συστήματα, σε συστήματα ελέγχου, σε συστήματα διαλόγου (chat-bots), σε συστήματα διαχείρισης υγείας κ.λπ. Αν εξάλλου, η Ενισχυτική ML συνδυαστεί με τα Μαζικά Δεδομένα Υγείας (ΜΔΥ), παρουσιάζονται μοναδικές ευκαιρίες για βελτιστοποίηση των θεραπειών στην υγειονομική περίθαλψη, ειδικά για τις επαναλαμβανόμενες. Ωστόσο, για να αξιοποιηθεί αυτή η δυνατότητα, πρέπει να επιδεικνύεται προσοχή και η δέουσα επιμέλεια, κατά την εφαρμογή και την αξιολόγησή τους (Coronato et al., 2020).

6.5.4 Αυτοματοποιημένη Μηχανική Μάθηση (Automated Machine Learning)

Η Αυτοματοποιημένη Μηχανική Μάθηση (AutoML) είναι η διαδικασία αυτοματοποίησης της διαδικασίας εφαρμογής της ML σε πραγματικά προβλήματα, η οποία περιλαμβάνει τα κατά σειρά βήματα, που ξεκινούν από το ακατέργαστο σύνολο δεδομένων και καταλήγουν στο αναπτυγμένο μοντέλο ML. Ο υψηλός βαθμός αυτοματοποίησης στην AutoML, επιτρέπει σε μη ειδικούς να χρησιμοποιούν μοντέλα και τεχνικές ML, χωρίς να απαιτείται πρώτα να γίνουν ειδικοί στον συγκεκριμένο τομέα.

Παρόλο που οι ερευνητές στον χώρο της υγείας είναι εξοικειωμένοι με τα κλινικά δεδομένα, εξακολουθούν συχνά να στερούνται τεχνογνωσίας, απαραίτητης για την εφαρμογή των τεχνικών ML, σε πηγές Μαζικών Δεδομένων Υγείας (ΜΔΥ). Οι ερευνητές συνεργάζονται μεν με ειδικούς επιστήμονες δεδομένων, αλλά η διαδραστική διαδικασία μεταξύ τους, απαιτεί γενικά πολύ χρόνο και προσπάθεια και από τα δύο μέρη.

Τα δεδομένα και η ανθρώπινη εμπειρογνωμοσύνη γενικά, δεν είναι άμεσα διαθέσιμα, ειδικά στο περιβάλλον της υγειονομικής περίθαλψης. Επομένως, είναι δύσκολο να επινοηθούν και να αναπτυχθούν λύσεις ML, καθώς η όλη προσπάθεια ξεκινά με μια μακρά διαδικασία παροχής των κατάλληλων δεδομένων, συνεχίζεται με την ανεύρεση των σωστών συνεργατών και συνεπάγεται συνεχή εναλλαγή μεταξύ εμπειρογνομόνων ML και εμπειρογνομόνων του χώρου της υγείας.



Σχ. 11 Σχηματική αναπαράσταση AutoML (Πηγή: Microsoft Azure Machine Learning AutoML, 2020).

Επομένως, η αυτοματοποίηση ορισμένων τμημάτων, τα οποία απαιτούν την παρέμβαση της ανθρώπινης εμπειρογνομosύνης, θα επέτρεπε στον κλάδο της υγειονομικής περίθαλψης να δημιουργήσει, να επικυρώσει και να αναπτύξει ταχύτερα λύσεις ML, και επομένως να αποκομίσει ευκολότερα τα οφέλη από τη βελτίωση της ποιότητας της προσφερόμενης υγειονομικής περίθαλψης στους ασθενείς. Με κίνητρο αυτόν ακριβώς τον στόχο, η AutoML έχει αναδειχθεί σε ένα νέο ταχέως εξελισσόμενο ερευνητικό πεδίο, για την αυτόματη βελτιστοποίηση τμημάτων της γραμμής παραγωγής της (pipeline).

Τυπικά στοιχεία ενός προβλήματος, στη γραμμή παραγωγής της AutoML είναι: 1) η προετοιμασία των δεδομένων, δηλαδή η φόρτωση και ο καθαρισμός των δεδομένων για χρήση τους στο σύστημα, καθώς και η εφαρμογή τυχόν μετασχηματισμών, κανονικοποιήσεων ή κωδικοποιήσεων, 2) η επιλογή χαρακτηριστικών (features) που θα χρησιμοποιηθούν για τη δημιουργία του μοντέλου ή ενδεχομένως τη χρήση γνώσης για την κατασκευή νέων χαρακτηριστικών προς βελτίωση του μοντέλου ML και 3) μια επαναληπτική διαδικασία στην οποία κάποιος ειδικός δημιουργεί, εκπαιδεύει, βελτιστοποιεί, επικυρώνει και επιλέγει έναν δεδομένο αλγόριθμο ML, που θα χρησιμοποιηθεί για ένα συγκεκριμένο πρόβλημα (Waring et al., 2020).

Κάθε μοντέλο ML έχει δύο τύπους παραμέτρων: υπερπαραμέτρους (hyperparameters) που ο σχεδιαστής του μοντέλου πρέπει να ορίσει χειροκίνητα, πριν από την εκπαίδευσή του, και τις κανονικές παραμέτρους, που βελτιστοποιούνται κατά την εκπαίδευση του μοντέλου. Αυτές οι υπερπαραμέτροι είναι ρυθμίσεις που ελέγχουν τη συμπεριφορά του αλγορίθμου ML, συχνά με τρόπο που είναι ιδιαίτερα ειδικός για αυτόν τον αλγόριθμο. Η βασικότερη αποστολή της AutoML είναι να ρυθμίσει αυτόματα αυτές τις υπερπαραμέτρους, ώστε να βελτιστοποιήσει την απόδοση του μοντέλου. Η απόδοση των περισσότερων μεθόδων ML μπορεί να εξαρτάται σε μεγάλο βαθμό από αυτές τις ρυθμίσεις, και ως εκ τούτου είναι μία από τις πιο σημαντικές εργασίες στην ML (Hoos et al., 2014).

Προκειμένου να καταστεί η ML πιο προσιτή σε μη τεχνικούς επαγγελματίες, οι ερευνητές της πληροφορικής έχουν προτείνει διάφορες μεθόδους αυτόματης επιλογής υπερπαραμέτρων. Αυτές οι μέθοδοι προσπαθούν να βρουν γρήγορα έναν βέλτιστο ή τουλάχιστον έναν αποτελεσματικό συνδυασμό τιμών υπερπαραμέτρων, που μεγιστοποιεί κάποια μέτρηση απόδοσης, για τη δεδομένη εργασία ML (hyperparameter optimization). Σε αυτές τις μεθόδους δίνεται συχνά ένα συγκεκριμένο όριο υπολογιστικών πόρων (constraints), όπως περιορισμένος χρόνος αναζήτησης ή περιορισμένη χρήση μνήμης. Η χρήση μιας αυτόματης μεθόδου επιλογής υπερπαραμέτρων μπορεί να μειώσει σημαντικά τον φόρτο, για όσους δημιουργούν λύσεις ML και πολλοί αλγόριθμοι επιλογής έχουν αποδειχθεί ότι βρίσκουν τιμές υπερπαραμέτρων, που είναι εξίσου καλές ή καλύτερες από τη χειροκίνητη ρύθμιση από ειδικούς ML (Komer et al., 2014).

Σκοπός των προσπαθειών στον χώρο της AutoML είναι να δημιουργηθεί ένα «μαύρο κουτί», το οποίο ελαχιστοποιεί τις απαιτήσεις ανθρώπινης εμπειρογνομosύνης, στην εφαρμογή της ML σε ένα ευρύ φάσμα προβλημάτων, δίνοντας λύσεις σε περίπτωση πιθανής έλλειψης επιστημόνων στον τομέα των δεδομένων και ενισχύοντας όσους κατέχουν τη σχετική γνώση. Ωστόσο, οι μέχρι σήμερα προσπάθειες εφαρμογής AutoML καταδεικνύουν αρκετές δυσκολίες στον χειρισμό του όγκου και της ποικιλίας των Μαζικών Δεδομένων, σε βιοϊατρικά περιβάλλοντα (Waring et al., 2020).

6.6 Υπολογιστική Νέφους και Άκρου (Cloud and Edge Computing)

Τα ιατρικά δεδομένα είναι περίπλοκα και λιγότερο «εύχρηστα» σε σύγκριση με αυτά που παρέχονται σε μεγάλες εμπορικές εταιρείες και συνεπώς απαιτείται επεξεργασία, προκειμένου να περιέλθουν σε μια εύκολα χρησιμοποιήσιμη μορφή. Τα ΜΔΥ παρέχουν δυνατότητα συνδυασμού ιατρικών πληροφοριών από πολλές πηγές, τόσο εσωτερικές όσο και εξωτερικές (όπως κλινικές ή εργαστηριακές δοκιμές, απεικονίσεις, γενετική, κ.λπ.), με αποτέλεσμα την παροχή «νοσημότητας», που δεν προέρχεται από καμία από τις παραπάνω πηγές.

Ωστόσο, η τεχνική υποδομή που επιτρέπει την κίνηση, τον χειρισμό και τη διαχείριση τους, δεν είναι πάντοτε άμεσα διαθέσιμη, αφού τα κεντρικά αποθετήρια αυτών των δεδομένων, υπάρχουν συνήθως σε υποδομές εντός νοσοκομείων.

Αυτές οι πληροφοριακές υποδομές είχαν σχεδιαστεί και χτιστεί στην προ-ΜΔΥ εποχή, χωρίς να επιτρέπεται η ανάλυση κι ο συνδυασμός τους με άλλα σύνολα δεδομένων και χωρίς να είναι δυνατή η κεντρική κοινή τους χρήση, η οποία θα ευνοούσε την ταχύτητα ή τον όγκο για εκμετάλλευση των μεθόδων των ΜΔΥ (Hulsen et al., 2019).

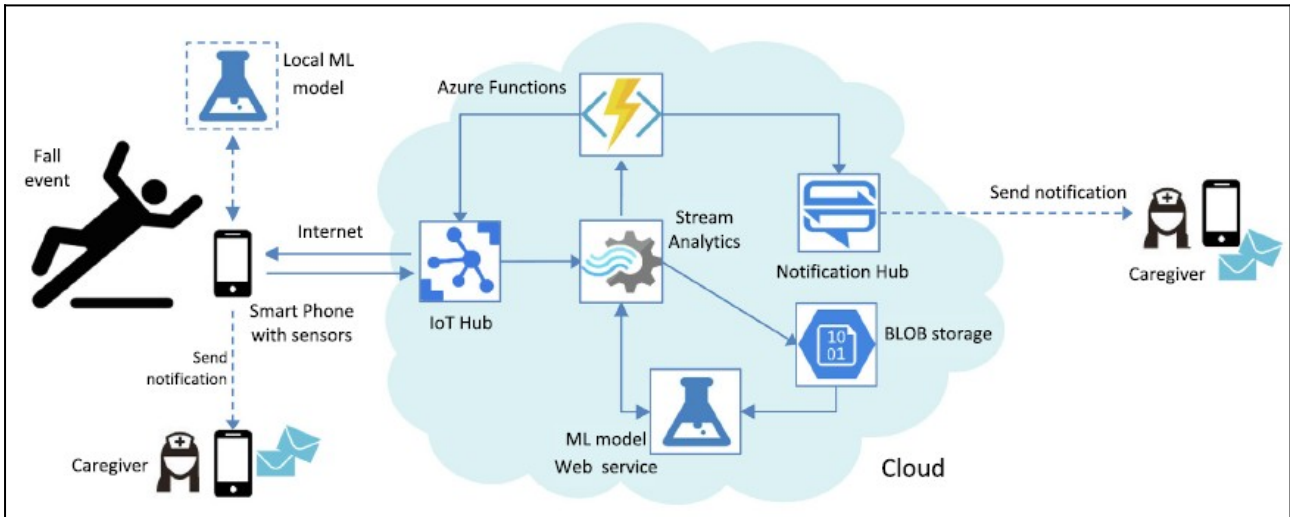
Τα παραπάνω μειονεκτήματα και τους περιορισμούς έρχονται να καλύψουν νέες τεχνολογίες, όπως η Υπολογιστική Νέφος και Άκρου. Η υπηρεσία Azure της Microsoft ορίζει την Υπολογιστική Νέφος (Cloud Computing) ως την παροχή υπηρεσιών πληροφορικής, περιλαμβανομένων διακομιστών, αποθηκευτικών χώρων, βάσεων δεδομένων, δικτύωσης, λογισμικού, αναλύσεων και νοημοσύνης, μέσω του Διαδικτύου («του νέφους»), για παροχή ταχύτερης καινοτομίας, ευέλικτων πόρων και οικονομίας κλίμακας. Οι χρήστες πληρώνουν τυπικά μόνο για τις υπηρεσίες νέφους που χρησιμοποιούν, μειώνοντας τα λειτουργικά τους κόστη, χρησιμοποιώντας αποτελεσματικά την υποδομή τους και προσαρμόζοντας την επιχείρησή τους, ανάλογα με τις ανάγκες.

Ήδη οι υπηρεσίες νέφους χρησιμοποιούνται ολοένα και περισσότερο και στον τομέα των ιατρικών δεδομένων, με ζητούμενο ασφαλώς την ασφάλεια της αποθήκευσής τους. Σε προτεινόμενο σύστημα από τους Kumar et al., (2018) η τεχνολογία του Cloud Computing συγκλίνει και συνεργάζεται με αυτές του IoT και των Big Data, σε τρεις φάσεις, με σκοπό την πρόγνωση σοβαρών ασθενειών.

Η πρώτη φάση χρησιμοποιείται για τη συλλογή των απαραίτητων δεδομένων από συσκευές IoT, από το UCI Repository (Dua and Graff, 2019) και από τα διαθέσιμα ιατρικά αρχεία. Η δεύτερη φάση χρησιμοποιείται για την ασφαλή αποθήκευση των ιατρικών αρχείων, στη βάση δεδομένων του νέφους. Η τρίτη φάση είναι υπεύθυνη για την πρόβλεψη κάποιας σοβαρής νόσου και τη διάγνωσή της (π.χ. διαβήτη). Σε αυτή τη φάση, υπάρχει ένα module που ονομάζεται αναλυτής σοβαρότητας (severity analyser), που χρησιμοποιείται για τον εντοπισμό της σοβαρότητας της ασθένειας, με χρήση κανόνων fuzzy, σύμφωνα με τις πληροφορίες που παρέχονται από τα ιατρικά αρχεία.

Οι προσπάθειες αποκέντρωσης των δεδομένων, που συλλέγονται και αποθηκεύονται στο νέφος, οδήγησε στην ανάπτυξη της τεχνολογίας Edge Computing, κατά την οποία ένας ηλεκτρονικός υπολογιστής τοποθετείται κοντά σε δρομολογητές (routers) και πύλες (gateways), μεταξύ νέφους και τελικού χρήστη μιας ηλεκτρονικής συσκευής οποιουδήποτε τύπου, η οποία είναι διασυνδεδεμένη στο IoT. Μια τέτοια διάταξη μειώνει σημαντικά το απαιτούμενο εύρος ζώνης και τις καθυστερήσεις στην επικοινωνία, ενώ παράλληλα βοηθά στη διατήρηση περισσότερων δεδομένων σε ιδιωτική σφαίρα, καθώς μεγάλο μέρος των υπολογισμών μπορεί να εκτελείται τοπικά. Έτσι, τα δεδομένα που παράγονται από συσκευές IoT, επεξεργάζονται πιο κοντά στην πηγή δημιουργίας τους, αντί να αποστέλλονται μέσα από μεγάλες διαδρομές, στα κέντρα δεδομένων ή στο νέφος. Με την επεξεργασία να διεξάγεται πιο κοντά στο «άκρο» (edge) του δικτύου, οι ενδιαφερόμενοι φορείς (π.χ. παροχής ιατροφαρμακευτικής περίθαλψης) είναι σε θέση να αναλύουν σημαντικά Μαζικά Δεδομένα Υγείας, σε σχεδόν πραγματικό χρόνο, ενώ οι συσκευές IoT μπορούν να τα διακινούν ταχύτερα και αποτελεσματικότερα. Τα δεδομένα επιλέγονται σε τοπικό επίπεδο, ώστε κάποια από αυτά να επεξεργάζονται επί τόπου, μειώνοντας έτσι την κυκλοφορία backhaul στον κεντρικό αποθηκευτικό χώρο, αλλά και την εξάρτηση από τα διαθέσιμα δεδομένα δικτύου.

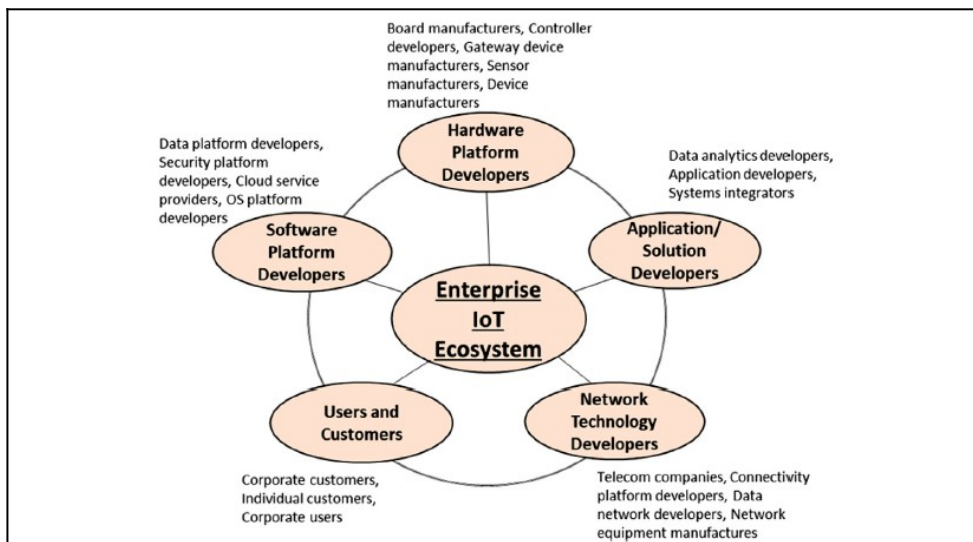
Σε πρόσφατη έρευνά τους οι Mrozek et al. (2020) παρουσίασαν μια επεκτάσιμη αρχιτεκτονική παρακολούθησης και ανίχνευσης περιστατικών πτώσης, καθώς και επικύρωση μοντέλων Μηχανικής Μάθησης, για την διαδικασία εντοπισμού τους. Για μια τέτοια απομακρυσμένη επιτήρηση ευπαθών ομάδων μεγάλης κλίμακας, απαιτούνται τοπικές συσκευές IoT, οι οποίες να μπορούν να εντοπίζουν επικίνδυνες καταστάσεις, αλλά και κέντρα δεδομένων που παρέχουν σημαντικό χώρο αποθήκευσης, με αυξημένες υπολογιστικές δυνατότητες. Η αρχιτεκτονική συστήματος Whoops που προτάθηκε, λειτουργεί σε περιβάλλον νέφους Azure, ανιχνεύοντας πτώσεις από κινητή συσκευή, κοντά στο άτομο που παρακολουθείται. Αν και η εφαρμογή προορίζεται για περιστατικά παρακολούθησης πτώσης ηλικιωμένων ατόμων, η ιδέα αυτή θα ήταν χρήσιμη για εντοπισμό πτώσης μαχητών στο πεδίο της μάχης και έγκαιρη ενημέρωση των προϊσταμένων τους ή του ιατρικού προσωπικού, για τις απαραίτητες ενέργειες (εντοπισμό, περισυλλογή, αναπλήρωση, πρώτες βοήθειες κ.λπ.)



Σχ. 12 Αρχιτεκτονική συστήματος Whoops για τον εντοπισμό περιστατικών πτώσεων, με εναλλακτική θέση συσκευών ανίχνευσης, που βασίζονται σε Μηχανική Μάθηση (Πηγή: Mrozek et al., 2020).

6.7 Διαδίκτυο των Πραγμάτων Υγείας (IoHT ή Internet of Healthcare Things)

Τον Ιούνιο του 2009, ο Kevin Ashton, αυτός που προφητικά, δέκα χρόνια νωρίτερα, είχε πρώτος χρησιμοποιήσει τον όρο «Internet of things» διευκρίνισε σε ένα άρθρο του (Ashton, 2009), τι εννοούσε με εκείνον τον όρο. Μεταξύ άλλων ανέφερε ότι «... αν είχαμε υπολογιστές που γνώριζαν όλα όσα θα έπρεπε να γνωρίζουν – χρησιμοποιώντας τα δεδομένα που συνέλεξαν χωρίς τη βοήθειά μας – θα μπορούσαμε να εντοπίζουμε και να μετράμε τα πάντα, και να μειώσουμε δραστικά τα απόβλητα, τις ζημιές και τα κόστη. Θα γνωρίζαμε τότε τα πράγματα χρειάζονται αντικατάσταση, επιδιόρθωση ή επανάκτηση, καθώς κι αν είναι φρέσκα ή έχουν χάσει την αξία τους...».



Σχ. 13 Το επιχειρηματικό IoT οικοσύστημα με τους πέντε κύριους εμπλεκόμενους φορείς (stakeholders) (Πηγή: I. Lee, 2019).

Το Διαδίκτυο των Πραγμάτων (IoT ή Internet of Things), σύμφωνα με την εταιρεία Oracle, περιγράφει το δίκτυο των φυσικών αντικειμένων («πραγμάτων»), που ενσωματώνουν αισθητήρες, λογισμικό και άλλες τεχνολογίες, με σκοπό τη σύνδεση και ανταλλαγή δεδομένων με άλλες συσκευές και συστήματα, μέσω του Διαδικτύου. Αυτές οι συσκευές ποικίλουν από τα συνηθισμένα οικιακά αντικείμενα έως τα εξελιγμένα βιομηχανικά εργαλεία.

Οι ειδικοί αναμένουν ότι ο αριθμός αυτός θα φτάσει τις 10 δισεκατομμύρια συνδεδεμένες συσκευές IoT το 2020 και τις 22 δισεκατομμύρια, μέχρι το 2025. Το IoT αναφέρεται σε ένα δίκτυο που το αποτελούν φυσικά αντικείμενα, όπως οι οικιακές συσκευές, τα ιατρο-τεχνολογικά προϊόντα, τα οχήματα και οι ετικέτες RFID (αναγνώριση ραδιοσυχνότητας). Μέσα από αυτό το δίκτυο, η επικοινωνία μεταξύ αυτών των αντικειμένων-πραγμάτων, λαμβάνει χώρα χωρίς να απαιτείται ανθρώπινη παρέμβαση, προκειμένου να κλείσει ο κύκλος (loop) της επικοινωνίας και της ανταλλαγής δεδομένων (Thibaud et al., 2018).

Σύμφωνα με έρευνα του I. Lee (2019), ένα επιχειρηματικό IoT οικοσύστημα περιλαμβάνει πέντε κύριους εμπλεκόμενους φορείς (stakeholders): 1) τους δημιουργούς της πλατφόρμας λογισμικού, 2) τους κατασκευαστές της πλατφόρμας του εξοπλισμού, 3) τους τεχνολόγους δικτυακών υποδομών, 4) τους προγραμματιστές των εφαρμογών/λύσεων και 5) τους πελάτες-χρήστες. Όλοι μαζί συνεργατικά και συνδυαστικά συνεισφέρουν στην καινοτομία, επεκτείνουν τις αγορές, ευνοούν τον ανταγωνισμό στη βιομηχανία, με τελικό όφελος τόσο υπέρ των επιχειρήσεων, όσο και υπέρ των πελατών-χρηστών.

Προσαρμοζόμενο ένα τέτοιο οικοσύστημα IoT στον χώρο της υγείας, περιλαμβάνει τους κατασκευαστές των συσκευών που φιλοξενούν τα λειτουργικά συστήματα και το λογισμικό, τους προγραμματιστές λειτουργικών συστημάτων (OS), των απαραίτητων για τη διαχείριση των δεδομένων υγείας και τον διαμοιρασμό τους μέσω των εφαρμογών/λύσεων των χρηστών, τους δημιουργούς τεχνολογιών δικτύου π.χ. παρόχους δικτύων 5G, τους προγραμματιστές εφαρμογών π.χ. φορητών συσκευών καταγραφής μετρήσεων κι ελέγχων (wearables), καθώς και τους ασθενείς ή λήπτες υπηρεσιών ιατροφαρμακευτικής περίθαλψης.

Η αρχιτεκτονική ενός επιχειρηματικού IoT περιλαμβάνει πέντε επίπεδα. Αυτά είναι το επίπεδο αντίληψης (perception layer), το επίπεδο δικτύου (network layer), το επίπεδο επεξεργασίας (processing layer), το επίπεδο εφαρμογής (application layer) και το επίπεδο διαχείρισης υπηρεσιών (service management layer).

Το επίπεδο αντίληψης ή αισθητήρα περιλαμβάνει διάφορες συσκευές όπως αισθητήρες, ετικέτες RFID και αναγνώστες, βιντεοκάμερες και έξυπνα τηλέφωνα, που χρησιμοποιούνται για τον εντοπισμό αντικειμένων και τοποθεσιών και τη συλλογή δεδομένων από το περιβάλλον. Συχνά στο επίπεδο αυτό παράγονται τεράστιες ποσότητες δεδομένων που πρέπει να συγκεντρωθούν και να αναλυθούν, ώστε να ανταποκριθούν σε διάφορα γεγονότα ενεργοποίησης (triggering) σε πραγματικό χρόνο.

Το επίπεδο δικτύου ή μετάδοσης, μεταδίδει δεδομένα που συλλέγονται από τους αισθητήρες προς το επίπεδο επεξεργασίας μέσω διαφόρων δικτύων. Λειτουργεί ως γέφυρα μεταξύ του επιπέδου αντίληψης και του επιπέδου εφαρμογής. Αυτό το επίπεδο παρέχει την υποδομή για την υποστήριξη ασύρματων ή ενσύρματων συνδέσεων μεταξύ συσκευών.

Το επίπεδο επεξεργασίας (middleware), φιλτράρει, αποθηκεύει, αναλύει και επεξεργάζεται δεδομένα, που μεταφέρονται από το επίπεδο δικτύου. Το επίπεδο επεξεργασίας περιέχει πλατφόρμες όπως διαχείριση βάσεων δεδομένων, αναλύσεις δεδομένων και υπολογιστική νέφος. Λόγω των Μάζικών Δεδομένων που παράγονται από τις συσκευές IoT, πολλές εφαρμογές IoT απαιτούν μαζική αποθήκευση δεδομένων, τεράστιες ταχύτητες επεξεργασίας για να επιτρέπουν τη λήψη αποφάσεων σε πραγματικό χρόνο και ευρυζωνικά δίκτυα υψηλής ταχύτητας, για τη ροή δεδομένων, ήχου ή βίντεο.

Το επίπεδο εφαρμογής αποτελείται από ένα σύνολο εφαρμογών/λύσεων, οι οποίες αλληλεπιδρούν με τους χρήστες, επιλύουν προβλήματα και μοιράζονται προβλήματα/λύσεις με άλλες εφαρμογές. Το επίπεδο εφαρμογής είναι επίσης υπεύθυνο για την ενσωμάτωση δεδομένων και πληροφοριών και την παρουσίασή τους στους χρήστες, σε φιλική προς αυτούς μορφή.

Στην κορυφή των πέντε επιπέδων αρχιτεκτονικής του επιχειρηματικού IoT βρίσκεται το επίπεδο διαχείρισης υπηρεσιών, που είναι υπεύθυνο για την επιλογή και την παράδοση των υπηρεσιών IoT της επιχείρησης. Το επίπεδο διαχείρισης υπηρεσιών αποτελεί σημείο εκκίνησης για την ανάπτυξη ενός συνόλου υπηρεσιών IoT. Η διαχείριση υπηρεσιών αναφέρεται στις δραστηριότητες και τις υποστηρικτικές διαδικασίες που εκτελούνται από έναν οργανισμό για τη σχεδίαση, την υλοποίηση, τη λειτουργία και τη βελτίωση των υπηρεσιών IoT, που χρησιμοποιούνται από τους πελάτες (I. Lee, 2019).

Εξάλλου, ένα επιχειρηματικό μοντέλο IoT, πρέπει να περιλαμβάνει τέσσερα βασικά δομικά στοιχεία, ως εξής: 1) Προτεινόμενη αξία, η οποία προκύπτει όταν νέες υπηρεσίες ή νέες αξίες γίνονται αντιληπτές από τους χρήστες/πελάτες, μέσα από την υιοθέτηση υπηρεσιών IoT. Για παράδειγμα, ένα έξυπνο δωμάτιο κλινικής ή νοσοκομείου με δυνατότητες IoT, θα βελτιώσει την άνεση και την περιθάλψη των νοσηλευομένων ασθενών, 2) Δραστηριότητες δικτύου, η βελτίωση των οποίων πρέπει να ενδιαφέρει τους συμμετέχοντες στο δίκτυο. Συμμετέχοντες στο δίκτυο είναι όλοι οι φορείς που τού προσθέτουν αξία, μέσω της ανάπτυξης και εξέλιξης των υπηρεσιών IoT, όπως κατασκευαστές συσκευών, προγραμματιστές, επιστήμονες δεδομένων κ.λπ., 3) Πόρους, όπως υπαλλήλους, τεχνολογίες, προϊόντα, υπηρεσίες, εγκαταστάσεις, εξοπλισμό, κανάλια μάρκετινγκ και μάρκα. Οι βιώσιμοι και αναπτυξιακοί αξιολογούνται από τις επιχειρήσεις και εξετάζονται, ως προς τις ευκαιρίες και τις απειλές, για την απόκτηση ανταγωνιστικού πλεονεκτήματος από την ανάπτυξη υπηρεσιών IoT, και 4) Βιωσιμότητα, που αποτελεί αναπόσπαστο μέρος του επιχειρηματικού μοντέλου IoT, σε μια προσπάθεια επίτευξης μακροπρόθεσμης εταιρικής ανάπτυξης και κερδοφορίας και ταυτόχρονα εκπλήρωσης περιβαλλοντικών και κοινωνικών ευθυνών. Για να είναι δε μακροπρόθεσμη, η αξία που δημιουργείται από την προσφορά των υπηρεσιών IoT στους χρήστες/πελάτες, πρέπει να είναι μεγαλύτερη από το κόστος παροχής των υπηρεσιών (I. Lee, 2019).

Για τη σωστή επιλογή μιας πλατφόρμας IoT, οι προγραμματιστές εφαρμογών πρέπει να λαμβάνουν υπόψη: 1) Τη δυνατότητα επεκτασιμότητας, για προσαρμογή στις γρήγορες αυξήσεις δεδομένων και στην αύξηση του αριθμού των συνδεδεμένων συσκευών, 2) Τις δυνατότητες έξυπνης υπολογιστικής άκρου (intelligent edge computing) για επέκταση της ισχύος της Υπολογιστικής Νέφους (cloud) σε κινητά και IoT συσκευές, 3) Τις δυνατότητες υποδομής νέφους, που θα ταιριάζει με τα υπάρχοντα συστήματα πληροφορικής, όπως το υβριδικό cloud, ώστε να δύναται να επωφεληθεί από την ευκολία προσβασιμότητας του ιδιωτικού cloud και τις δυνατότητες επέκτασης του δημόσιου cloud, 4) Πρόβλεψη περί πρόληψης και αποκατάστασης καταστροφών, όπως σχέδια αποκατάστασης καταστροφών, μέτρα ασφαλείας και σχέδια δημιουργίας αντιγράφων ασφαλείας δεδομένων και 5) Πρωτόκολλα και πρότυπα επικοινωνίας, με δυνατότητα υποστήριξης αναβαθμισμένων εκδόσεων αυτών των πρωτοκόλλων ή για εύκολη αντικατάστασή τους από νεότερα πρωτόκολλα (Savjani, 2017).

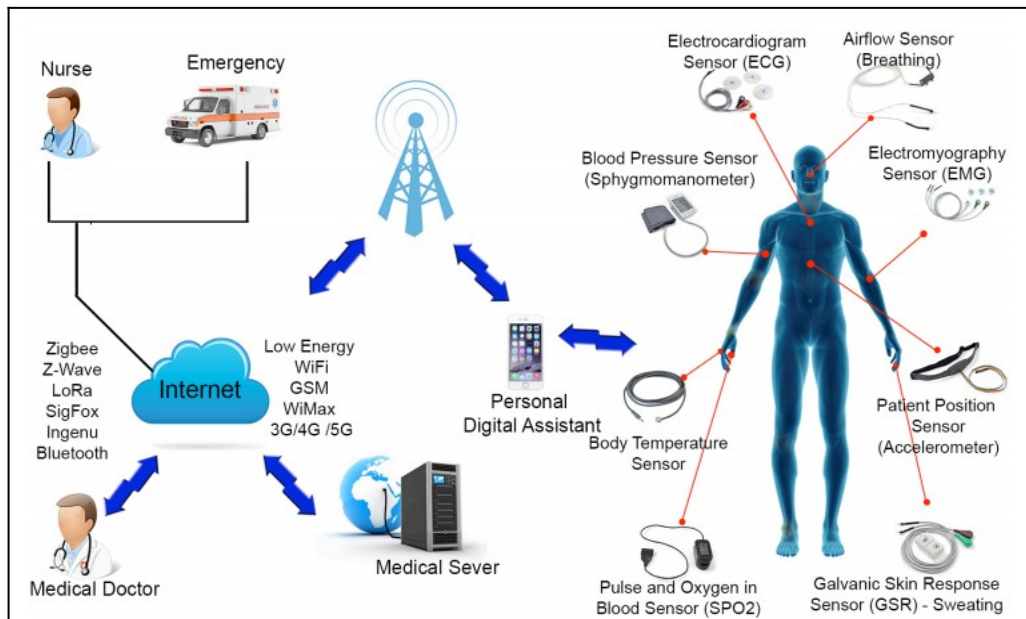
Το IoT επιτρέπει σενάρια, όπου έξυπνες συσκευές μπορούν να αλληλεπιδρούν με άλλα έξυπνα διασυνδεδεμένα αντικείμενα, προκειμένου να αποκτηθούν νέες γνώσεις και ενημέρωση, τόσο για τους χρήστες όσο και για το περιβάλλον, για την υποστήριξη λήψης αποφάσεων. Υπάρχουν διάφορες παραλλαγές του IoT στον τομέα της υγείας, καθεμία με τις ιδιαιτερότητές της.

Στο Διαδίκτυο των Πραγμάτων Υγείας (IoHT), οι υπάρχουσες λύσεις επιτρέπουν εντοπισμό, ανταλλαγή πληροφοριών σε πραγματικό χρόνο, απομακρυσμένη ή αυτόματη διαχείριση πόρων, υψηλότερη ποιότητα περίθαλψης, εξοικονόμηση χρόνου και ασφάλεια ασθενών. Χάρη στο IoHT, η διαχείριση των ιατρικών εγκαταστάσεων είναι πιο αποτελεσματική, με την αδιάκοπη πρόσβαση στον εξοπλισμό, τα δεδομένα και την ενημέρωση των ασθενών (Comarch, 2019). Όταν τα δεδομένα που συλλέγονται με τις συσκευές IoHT συνδυάζονται με συστήματα ηλεκτρονικού φακέλου υγείας (EHR), ανοίγεται μια νέα διάσταση και γεννιούνται πολλές δυνατότητες και χρήσεις (Technosoft, 2019).

Το IoHT διευκολύνει καθημερινές αλλά σημαντικές εργασίες, οι οποίες συντελούν στη βελτίωση της κατάστασης των ασθενών και μειώνει τον φόρτο των επαγγελματιών υγείας. Η εξ αποστάσεως παρακολούθηση των ασθενών, της εξέλιξης της θεραπείας τους και της φαρμακευτικής τους αγωγής, είναι μερικές από τις προσφερόμενες δυνατότητες του IoHT. Η παραδοσιακή αρχιτεκτονική εξ αποστάσεως παρακολούθησης των ασθενών, περιλαμβάνει το φυσικό επίπεδο που αποτελείται από το Ασύρματο Δίκτυο Περιοχής Σώματος (Wireless Body Area Network – WBAN) για τη λήψη των ιατρικών δεδομένων και την εκτέλεση κάποιων υπηρεσιών, το επίπεδο επικοινωνιών και δικτύου για τη μετάδοση των δεδομένων και το επίπεδο υπηρεσιών για την επεξεργασία νέφους (Thibaud et al., 2018).

Το Διαδίκτυο των Φορετών Πραγμάτων - Wearable Internet of Things (WIoT) χρησιμοποιεί δυνατότητες τηλε-υγείας (telehealth), με σκοπό τη δημιουργία και λειτουργία ενός οικοσυστήματος για αυτοματοποιημένες παρεμβάσεις.

Αξιοποιώντας αισθητήρες που φέρονται στο σώμα του χρήστη, το WIoT επιτρέπει την παρακολούθηση δεδομένων χρήσιμων για υποστήριξη των ατόμων στην καθημερινή τους ζωή (π.χ. συμπεριφορές, ευεξία, συνήθειες κ.λπ.) και συνδέει τους ασθενείς με ιατρικές υποδομές (Aceto et al., 2020)..



Σχ. 14 Τυπική απεικόνιση ενός WBAN σε ένα έξυπνο σύστημα παροχής υγειονομικής περίθαλψης εντός του IoT, με πράγματα, δίκτυα και εφαρμογές (Πηγή: Rodrigues et al., 2018)

Παράλληλα, σύγχρονες τεχνολογικές πλατφόρμες προσφέρουν αξιόπιστες και οικονομικές λύσεις, με δυνατότητα απομακρυσμένης παρακολούθησης των ασθενών στο σπίτι τους. Ο ασθενής φέρει τη φορητή συσκευή καταγραφής, η οποία στέλνει τα δεδομένα στο θεράποντα ιατρό ή νοσηλεύτη κι αυτός με τη σειρά του, έχοντας όλα τα στοιχεία στα χέρια του, συμβουλεύει τον ασθενή. Τέτοια συστήματα απευθύνονται περισσότερο σε ασθενείς με καρδιοπάθειες, πνευμονοπάθειες, υπέρταση και ζάχαρο, οι οποίοι χρειάζονται χρόνια παρακολούθηση. Ωστόσο, υπάρχει η δυνατότητα να χρησιμοποιηθούν και για την παρακολούθηση άλλων καταστάσεων, όπως οι μετεγχειρητικοί ή ψυχιατρικοί ασθενείς. Μέσω τέτοιων εφαρμογών, εξοικονομείται πολύτιμος χρόνος από επισκέψεις σε γιατρούς και νοσοκομεία, ενώ οι ασθενείς απολαμβάνουν μια σχεδόν αδιάλειπτη ιατρική παρακολούθηση.

Το Διαδίκτυο των Ιατρικών Πραγμάτων (IoMT) αναφέρεται σε εφαρμογές που αποτελούνται από εμφυτεύσιμες και φορητές συσκευές, συνδεδεμένες σε ένα προσωπικό έξυπνο τηλέφωνο (ή ρολόι) και στο Διαδίκτυο, επιτρέποντάς του να ενεργεί ως προσωπικός κόμβος (hub).

Το Διαδίκτυο των Νανο-Πραγμάτων (IoNT) αναφέρεται στην εφαρμογή του IoT στη νανοϊατρική, στην εφαρμογή πιο εξατομικευμένης παρακολούθησης, διάγνωσης και θεραπείας, για την εφαρμογή προληπτικής παρακολούθησης, προληπτικής υγείας, διαχείρισης χρόνιων νοσημάτων και παρακολούθησης.

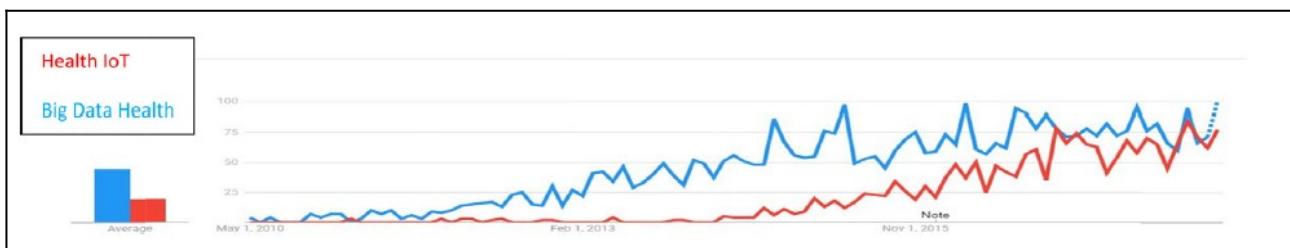
Το Διαδίκτυο των Κινητών Υπηρεσιών Υγείας (m-IoT) αναφέρεται σε ένα μοντέλο σύνδεσης μεταξύ δικτύων προσωπικής περιοχής (Personal Area Network – PAN), χαμηλής ισχύος και εξελισσόμενων δικτύων κινητής τηλεφωνίας, δίνοντας έμφαση στα υπάρχοντα ειδικά χαρακτηριστικά, που είναι απαραίτητα για την παγκόσμια κινητικότητα των συμμετεχόντων οντοτήτων (Aceto et al., 2020).

Σε άρθρο των Mani et al. (2020), συνοψίζονται τα κίνητρα, η σημασία και η συνεισφορά της τεχνολογίας IoT στον χώρο της υγείας, μέσω της παροχής συνεχούς επιτήρησης, της οικοδόμησης εμπιστοσύνης, της απομακρυσμένης παρακολούθησης των ασθενών, της μείωσης του κόστους των δαπανών υγείας, των ειδοποιήσεων έκτακτης ανάγκης, των γνώσης που παρέχουν τα δεδομένα, της απομακρυσμένης ιατρικής βοήθειας, της παρακολούθησης προσωπικού, της απογραφής και των ασθενών, καθώς και της διαχείρισης της φαρμακευτικής τους αγωγής.

Στην περίπτωση του στελεχών των ΕΔ, μπορεί να εφαρμοστεί η απομακρυσμένη παρακολούθηση με συσκευές ΙοΗΤ, κατά τη διάρκεια της εργασίας, της διαμονής τους στο σπίτι ή σε δραστηριότητες εκτός υπηρεσίας και κυρίως κατά τη συμμετοχή τους σε ασκήσεις ή σε πολεμικές επιχειρήσεις. Ειδικά στην τελευταία περίπτωση, ο εξ αποστάσεως έλεγχος των ζωτικών σημείων των μαχόμενων στελεχών (πίεση, σφίξεις, θερμοκρασία, κορεσμός) μπορεί να δώσει πολύ κρίσιμη πληροφόρηση στον επιχειρησιακό διοικητή, σχετικά με την κατάσταση της υγείας του προσωπικού του και να τον διευκολύνει στη λήψη απόφασης για τις ενδεδειγμένες ενέργειες (π.χ. αίτηση αναπλήρωσης, παραπομπή προς εξέταση, αλλαγή τρόπου ενεργείας, αυτοματοποιημένη αναφορά απωλειών υγείας κ.ά). Απαραίτητη προϋπόθεση και στην περίπτωση αυτή, αποτελεί η ασφαλής διακίνηση των δεδομένων υγείας, μέσω κρυπτασφαλισμένου ασύρματου ευρυζωνικού δικτύου υψηλών ταχυτήτων.

Σε όλες τις παραπάνω αναφερόμενες περιπτώσεις χρήσης ΙοΗΤ, όπου υποκείμενα συλλογής δεδομένων υγείας είναι τα στελέχη των ΕΔ, το αποτέλεσμα είναι η συσσώρευση μεγάλου όγκου δεδομένων (Big Data) υγείας, στους υπεύθυνους ή εκτελούντες την επεξεργασία, αφού το ΙοΗΤ αποτελεί έναν από τους τρόπους συλλογής τους.

Όπως είναι αυτονόητο, οι τεχνολογίες ΙοΗΤ και ΜΔΥ συγκλίνουν, με τη δημοτικότητά τους, όπως προκύπτει από αναζητήσεις στο google, να αυξάνει συνεχώς λόγω της σπουδαιότητας τους στον τομέα της υγείας. Το Σχ. απεικονίζει τη δημοτικότητα των δυο συγκλινουσών τεχνολογιών, με την πάροδο του χρόνου. Όπως φαίνεται, η ανοδική τάση του ΙοΗΤ ξεκίνησε από τα μέσα του 2014, ενώ τα ΜΔΥ κέρδιζαν σταδιακά δημοτικότητα από τα μέσα του 2012, με τη δημοτικότητα να αυξάνει μέχρι και σήμερα (Saheb και Izadi, 2019).



Σχ. 15 Γραφήματα τάσεων από έρευνες στο google, για Health IoT και Big Data Health (Πηγή: Saheb και Izadi, 2019).

6.8 Δίκτυο Κινητής Τηλεφωνίας 5G

Το δίκτυο κινητής τηλεφωνίας 5G αποτελεί την πέμπτη γενιά κυψελοειδούς τεχνολογίας. Έχει σχεδιαστεί για να αυξήσει την ταχύτητα, να μειώσει την καθυστέρηση και να βελτιώσει την ευελιξία των ασύρματων υπηρεσιών. Η τεχνολογία 5G έχει θεωρητική μέγιστη ταχύτητα 20 Gbps, ενώ η μέγιστη ταχύτητα 4G είναι μόνο 1 Gbps. Η υποσχόμενη χαμηλότερη καθυστέρηση μπορεί να βελτιώσει την απόδοση των επιχειρηματικών εφαρμογών καθώς και άλλες ψηφιακές εμπειρίες (CISCO, 2019).

Το δίκτυο 5G πρέπει να είναι σε θέση να ανταποκρίνεται δυναμικά στις απαιτήσεις συγκεκριμένων συσκευών και εφαρμογών, με συνέπειες τόσο στο συνολικό αρχιτεκτονικό σχεδιασμό του δικτύου, όσο και στις μεμονωμένες τεχνολογίες και τεχνικές που χρησιμοποιούνται για την παροχή των υπηρεσιών υγείας. Λόγω των χαρακτηριστικών του (μεγάλη ταχύτητα/χωρητικότητα - μικρές καθυστερήσεις), θα δίνει τη δυνατότητα απρόσκοπτης θέασης οπτικοακουστικού υλικού σε πολύ υψηλή ανάλυση, ώστε να μπορεί να χρησιμοποιηθεί πολύ αποτελεσματικά, σε εφαρμογές μετάδοσης ζωντανής εικόνας, σε συστήματα, τηλεϊατρικής, παρακολούθησης/επιτήρησης ασθενών από απόσταση, μέτρησης, καταγραφής και αποστολής των τιμών των ζωτικών τους σημείων κ.λπ.

Οι προσφερόμενες, μέσω του 5G δικτύου, επικοινωνίες παρουσιάζουν σημαντική αλλαγή από δίκτυα κινητής τηλεφωνίας που χρησιμοποιούνταν ως τώρα, παρέχοντας καθολικές συνδέσεις πολύ υψηλής ταχύτητας και απρόσκοπτη εμπειρία χρήσης για τον συνδρομητή (Boccardi et al., 2014).

Τα δίκτυα 5G στοχεύουν (Osseiran et al., 2014) στην παροχή 1000x υψηλότερου όγκου δεδομένων κινητής τηλεφωνίας ανά περιοχή, 100x μεγαλύτερου αριθμού συνδεδεμένων συσκευών, 100x υψηλότερου ρυθμού μετάδοσης χρήστη, 10x μεγαλύτερης διάρκειας ζωής μπαταρίας για μαζικές επικοινωνίες μηχανών χαμηλής ισχύος και 5x μειωμένης End-to-End καθυστέρησης (E2E latency). Αυτοί οι στόχοι θα επιτευχθούν με βασικές τεχνολογίες, όπως mmWaves, δίκτυα μικρών κυψελών, μαζική πολλαπλή είσοδο πολλαπλή έξοδο (Multiple Input Multiple Output - MIMO) και διαμόρφωση δέσμης (Andrews et al., 2014).

Χρησιμοποιώντας αυτές τις τεχνολογίες, το 5G θα υποστηρίζει τρεις κατηγορίες υπηρεσιών, δηλαδή βελτιωμένο εύρος ζώνης για κινητά (eMBB), πολύ αξιόπιστη και χαμηλής καθυστέρησης επικοινωνία (URLLC) και επικοινωνία τύπου μαζικής μηχανής (mMTC). Έτσι, το 5G θα επιτρέπει παροχή ψηφιακών υπηρεσιών υγείας, εν παντί τόπω και χρόνο. Τα δίκτυα 5G θα αναπτυχθούν παράλληλα με θεμελιώδεις τεχνολογίες όπως το Software Defined Networking (SDN), το Network Function Virtualization (NFV), το Multi-access Edge Computing (MEC) και το Network Slicing (NS). Οι τεχνολογίες SDN και NFV θα δώσουν τη δυνατότητα στα προγραμματιζόμενα δίκτυα 5G να υποστηρίζουν γρήγορη ανάπτυξη και ευέλικτη διαχείριση των προσφερόμενων υπηρεσιών 5G. Η τεχνολογία MEC θα επεκτείνει τη νοημοσύνη στο άκρο του ασύρματου δικτύου, μαζί με μεγαλύτερες δυνατότητες επεξεργασίας και αποθήκευσης. Η τεχνολογία NS θα δημιουργεί λογικά δίκτυα σε μια κοινή υποδομή, για να επιτρέπει διαφορετικούς τύπους υπηρεσιών στα δίκτυα 5G (Siriwardhana et al., 2020).

Το δίκτυο 5G είναι το επόμενο μεγάλο βήμα στον τρόπο με τον οποίο τα δίκτυα θα βελτιώσουν την ανθρώπινη ζωή και καθημερινότητα, μέσα από μια έκρηξη των εφαρμογών γύρω από το IoT, σχεδόν σε κάθε πεδίο της ανθρώπινης δραστηριότητας. Έτσι, ο χώρος της υγείας αναμένεται να σημειώσει εκθετική ανάπτυξη στη χρήση της τεχνολογίας IoT, καθιστώντας εφικτή την ακόμη μεγαλύτερη πρόσβαση στις προσφερόμενες υπηρεσίες του εν λόγω χώρου και δημιουργώντας τεράστιες ευκαιρίες για βελτιωμένη υγεία, πιο άνετη αποθεραπεία και καλύτερη ζωή. Αυτό θα επιτευχθεί ενισχύοντας τους ασθενείς, διευκολύνοντας την αυτοεξυπηρέτηση ευπαθών ομάδων και ηλικιωμένων, βελτιώνοντας την προσφερόμενη σε αυτούς ιατροφαρμακευτική φροντίδα και παρέχοντας επιπλέον εργαλεία για την καλύτερη υποστήριξη του θεράποντος ιατρού, στη λήψη αποφάσεων, με τη χρήση δεδομένων σε πραγματικό χρόνο.

Άλλα αναμενόμενα οφέλη από την ευρεία χρήση αυτής της τεχνολογίας είναι η μείωση του κόστους θεραπείας και παρακολούθησης περιστατικών χρόνιων παθήσεων ή αυτών που χρήζουν επείγουσας θεραπείας, η εξ αποστάσεως υποστήριξη από ειδικούς και οι χειρουργικές επεμβάσεις με τη χρήση ρομποτικής, η ενίσχυση της λειτουργίας μιας αδιάλειπτης εφοδιαστικής αλυσίδας υγείας, η εκτέλεση βοηθητικών εργασιών (καθαριότητας, τροφοδοσίας, συντήρησης), η οποία θα επιτρέπει στο ιατρικό και νοσηλευτικό προσωπικό να αφιερώνει περισσότερο χρόνο στους ασθενείς, η ορθολογικότερη κατανομή και εξοικονόμηση πόρων, εντός των νοσηλευτικών ιδρυμάτων (προσωπικό, μηχανήματα, εργαλεία, κλίνες), η συνεχής παρακολούθηση εμβολίων ή άλλων ευαίσθητων φαρμακευτικών σκευασμάτων, μέσω αισθητήρων και απευθείας απεικόνιση της κατάστασής τους, με εξ αποστάσεως έλεγχο της θερμοκρασίας κ.λπ.

6.9 Εφαρμογές Κινητής Υγείας (mHealth)

Η ραγδαία εξέλιξη της Τεχνολογίας της Πληροφορικής και των Επικοινωνιών ανέδειξε μια νέα μορφή ηλεκτρονικής υγείας, την «κινητή υγεία» (mobile health, mHealth), την οποία ο Παγκόσμιος Οργανισμός Υγείας (Π.Ο.Υ.) ορίζει ως «πρακτική ιατρικής και δημόσιας υγείας που υποστηρίζεται από κινητές συσκευές, όπως κινητά τηλέφωνα, συσκευές παρακολούθησης ασθενών, προσωπικούς ψηφιακούς βοηθούς και άλλες ασύρματες συσκευές».

Η mHealth παρέχει υπηρεσίες υγειονομικής περίθαλψης, υπερβαίνοντας τους γεωγραφικούς, τους χρονικούς και τους οργανωτικούς φραγμούς. Οι εφαρμογές (apps) της mHealth αντιμετωπίζουν σημαντικά προβλήματα των υπηρεσιών υγείας, όπως ο διαρκώς αυξανόμενος αριθμός ασθενών με χρόνια νοσήματα και το υψηλό κόστος περίθαλψής τους για τα συστήματα υγείας, ενδυναμώνοντας τους ασθενείς και παρέχοντας σε αυτούς και τις οικογένειές τους, τη δυνατότητα αυτοδιαχείρισης της νόσου τους.

Πολλές εφαρμογές της mHealth έχουν σχεδιαστεί αποκλειστικά για τους επαγγελματίες υγείας. Οι εν λόγω εφαρμογές χαρακτηρίζονται από ιδιαίτερη πολυπλοκότητα, που ορίζεται κυρίως από την υιοθέτηση ιατρικής ορολογίας, καθώς απευθύνονται σε άτομα τα οποία διαθέτουν τις απαραίτητες ιατρικές γνώσεις. Άλλοι τύποι εφαρμογών που είναι δημοφιλείς μεταξύ των φοιτητών Ιατρικής, είναι οι εφαρμογές μεταφοράς δεδομένων για ανταλλαγή αρχείων και οι εφαρμογές με τρισδιάστατες αναπαραστάσεις οργάνων και συστημάτων του ανθρώπινου οργανισμού.

Υπάρχουν εφαρμογές για τους χρόνια πάσχοντες, όπως για ασθενείς που πάσχουν από διαβήτη, καρδιαγγειακές παθήσεις και χρόνια αποφρακτική πνευμονοπάθεια (ΧΑΠ), που διατίθενται από την ηλεκτρονική πλατφόρμα Google Play. Οι συγκεκριμένες εφαρμογές επικεντρώνονται σε ένα ευρύ φάσμα παροχών υγείας, από απλή υπενθύμιση της λήψης της δόσης ενός φαρμάκου, με τη μορφή μιας ηχητικής ειδοποίησης ή γραπτού μηνύματος SMS, μέχρι σημαντικά πιο σύνθετες λειτουργίες, όπως η λήψη και η καταγραφή ζωτικών σημάτων σε πραγματικό χρόνο, με τη βοήθεια πρόσθετου ιατρο-τεχνολογικού εξοπλισμού, ο οποίος συνδέεται ενσύρματα ή ασύρματα στην έξυπνη συσκευή, ή η αποθήκευση ιατρικών δεδομένων (Σπυριδάκη et al., 2019).

Ωστόσο, λόγω του μεγάλου αριθμού διαθέσιμων εφαρμογών που συνεισφέρουν στην ευζωία (well-being) χωρίς εγγυήσεις για την ποιότητα και την αξιοπιστία τους, προκύπτουν ανησυχίες ως προς την ικανότητα των χρηστών τους, να εκτιμήσουν τη χρησιμότητά τους, περιορίζοντας έτσι την αποτελεσματική πρόσβασή τους, προς όφελος της δημόσιας υγείας. Η διασφάλιση της ποιότητας των δεδομένων είναι επίσης απαραίτητη για τη σύνδεση των εφαρμογών με τα ηλεκτρονικά μητρώα υγείας και για την αξιοποίηση στην κλινική πρακτική.

Τον Απρίλιο του 2014, η Ε.Ε. δημοσίευσε την «Πράσινη Βίβλο» για την κινητή υγεία (Green Paper on mobile health), δρομολογώντας παράλληλα μια δημόσια διαβούλευση για την ανάδειξη εμποδίων στην ανάπτυξη του mHealth, σε ευρωπαϊκό επίπεδο. Από τη διαβούλευση αυτή προέκυψαν χρήσιμα συμπεράσματα, ως προς την προστασία των δεδομένων που διακινούνται μέσω mHealth. Έτσι, τονίστηκε η σημασία της ιδιωτικότητας και της τήρησης αρχών ασφαλείας, προκειμένου να οικοδομηθεί η εμπιστοσύνη από πλευράς των χρηστών αυτής της τεχνολογίας. Ως πιο δημοφιλή μέτρα ασφαλείας προωθήθηκαν οι μηχανισμοί κρυπτογράφησης και ελέγχου ταυτότητας δεδομένων, ενώ αναγνωρίστηκε ότι τα δεδομένα υγείας, ως ευαίσθητα, πρέπει να είναι κρυπτογραφημένα τόσο «σε διαμετακόμιση» (in transit) όσο και «σε κατάσταση ηρεμίας» (at rest).

Η συλλογή, επεξεργασία και αποθήκευση προσωπικών δεδομένων που διακινούνται μέσω mHealth πρέπει να περιορίζεται στα απολύτως απαραίτητα για τους σκοπούς της συλλογής, όπως άλλωστε επιτάσσει και η αρχή της «ελαχιστοποίησης» του ΓΚΠΔ.

Η διαλειτουργικότητα και η ανάγκη χρήσης πρότυπων (π.χ. στον τρόπο επεξεργασίας των πληροφοριών) τέθηκε ως απαραίτητη προϋπόθεση, για την πλήρη αξιοποίηση των δυνατοτήτων των ΜΔΥ, τα οποία διακινούνται μέσω mHealth. Εντοπίστηκε η ανάγκη ύπαρξης σαφώς καθορισμένων δομών διακυβέρνησης για την προώθηση της εμπιστοσύνης του κοινού στη χρήση ΜΔΥ, καθώς και η ανάγκη δυνατότητας διαμοιρασμού δεδομένων, μεταξύ εφαρμογών mHealth και ατομικών ηλεκτρονικών φακέλων υγείας (ΑΗΦΥ). Ωστόσο, αναγνωρίστηκε πως η ανωνυμοποίηση πρέπει πάντα να είναι η προτιμώμενη επιλογή για τον μετριασμό των κινδύνων που διατρέχει η ιδιωτικότητα των πολιτών, ενώ είναι και επιτακτική η ανάγκη για την ύπαρξη μέτρων λογοδοσίας από τους υπεύθυνους ή εκτελούντες επεξεργασία ΜΔΥ και ελέγχου από τις αντίστοιχες αρχές.

7. Ανάλυση Δεδομένων

Οι δραματικές μειώσεις του κόστους υπολογισμών μέσω της πληροφορικής, καθώς και η υπερεπάρκεια δυνατοτήτων αποθήκευσης, επεξεργαστικής ισχύος και συνδεσιμότητας δικτύων, προσφέρουν πλέον τεράστιες ευκαιρίες βελτίωσης της ανθρώπινης υγείας. Ωστόσο, το υπάρχον εργατικό δυναμικό δεν είναι σε θέση να αξιοποιήσει τις προαναφερθείσες δυνατότητες, λόγω ανεπαρκούς γνώσης, εκπαίδευσης και έλλειψης εργαλείων, απαραίτητων για την επεξεργασία δεδομένων αυτής της πολυπλοκότητας και ποικιλομορφίας, ενώ ακόμη λιγότεροι είναι οι εξειδικευμένοι ερευνητές, για την ανάπτυξη και αξιολόγηση προσβάσιμων, εφαρμόσιμων εργαλείων, με οφέλη για την ανθρώπινη υγεία.

Οι πληροφορίες οι σχετικές με την υγεία, η μορφή δηλαδή στην οποία μεταπίπτουν τα δεδομένα υγείας μετά από την επεξεργασία τους, μπορούν να αποτελέσουν πολύ σημαντικούς πόρους, για τις βιομηχανίες υγειονομικής περίθαλψης. Τα οφέλη, που σχετίζονται με τα ιατρικά δεδομένα και την ανάλυσή τους από διάφορους οργανισμούς υγειονομικής περίθαλψης, έχουν συμβάλει στην συγκέντρωση αξιοσημείωτου ενδιαφέροντος προς το πεδίο της Ιατρικής Πληροφορικής και Ανάλυσης (Healthcare Informatics and Analytics - HCI&A). Το πεδίο αυτό προσελκύει ερευνητές από διάφορους επιστημονικούς κλάδους, που περιλαμβάνουν την επιστήμη των υπολογιστών, τις κοινωνικές επιστήμες, τις επιχειρήσεις, τη φυσική, τη βιολογία και την ιατρική.

Μέσω των υφιστάμενων τεχνολογιών, η επεξεργασία των δεδομένων εγείρει πολλές προκλήσεις, όχι μόνο λόγω του όγκου τους αλλά και της ετερογενούς φύσης τους, καθώς και λόγω της ταχύτητας διαχείρισής τους. Επομένως, τόσο η επιστήμη της Ιατρικής Πληροφορικής όσο και η συναφής υπάρχουσα βιβλιογραφία, δίνουν ιδιαίτερη βαρύτητα στην υπολογιστική προοπτική, δηλαδή στην αποδοτικότητα της επεξεργασίας, στην ανάκτηση και χειρισμό των πληροφοριών, καθώς και στις τεχνολογικές δυνατότητες του χώρου της υγειονομικής περίθαλψης (Pramanik et al., 2020).

Σε άρθρο τους οι Zhang και Zhou (2019), συνοψίζουν τις έξι βασικότερες προκλήσεις στην Ανάλυση Δεδομένων Υγείας (ΑΔΥ). Η πρώτη από αυτές είναι η «ετερογένεια» (“heterogeneity”), καθώς τα διαθέσιμα δεδομένα υγείας σήμερα είναι ένα μείγμα δομημένων, ημι-δομημένων και μη δομημένων δεδομένων. Πιθανή λύση αντιμετώπισης αυτής της ετερογένειας είναι η «συγχώνευσή» τους, η οποία ενσωματώνει τα ετερογενή δεδομένα σε προσπάθειες δημιουργίας αναλυτικών μοντέλων καλύτερης απόδοσης, σε σύγκριση με τα μοντέλα που χρησιμοποιούν δεδομένα μονής μορφής (single modality).

Δεύτερη μεγάλη πρόκληση είναι ο χειρισμός των «Μεγάλων» Δεδομένων Υγείας (ΜΔΥ). Η International Data Corporation (IDC) προέβλεψε ότι «η παγκόσμια βάση δεδομένων θα αυξηθεί από 33 zettabytes το 2018 σε 175 zettabytes έως το έτος 2025».

Το ίδιο έτος, υπολογίζεται ότι από τα παραπάνω δεδομένα, 90 ZB θα δημιουργηθούν σε συσκευές IoT, 49% θα αποθηκευτούν σε δημόσιες υποδομές Νέφους, ενώ το 30% αυτών θα χρησιμοποιούνται σε πραγματικό χρόνο (Patrizio, 2018).

Αυτές οι τεράστιες ποσότητες δεδομένων δημιουργούν με τη σειρά τους άλλες προκλήσεις, όπως έλλειψη τυποποίησης δεδομένων, ανησυχία για ζητήματα απορρήτου και ασφάλειας, περιορισμό ταχύτητας και δυνατότητα υποκλοπής κατά τη μεταφορά δεδομένων, αξιοπιστία αποθήκευσης δεδομένων κ.λπ. Κατά συνέπεια, επιβραδύνεται η διαχείριση δεδομένων υγείας και η χρήση υπαρχόντων μοντέλων ανάλυσής τους. Οι διακομιστές υψηλών ταχυτήτων με ενσωμάτωση καταναμημένων υπολογιστικών συστημάτων, αλγορίθμων ροής (streaming) ή υπολογιστικού νέφους, αποτελούν πιθανές λύσεις σε αυτήν την πρόκληση.

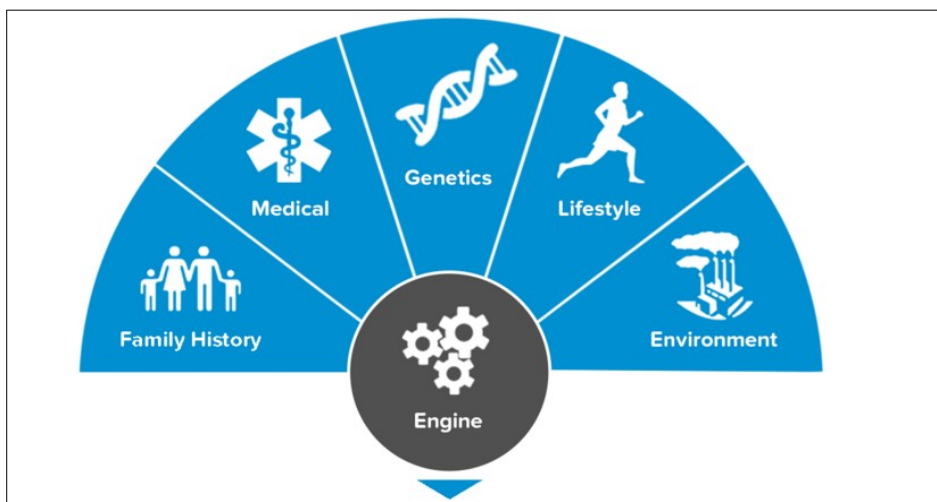
Τρίτη πρόκληση είναι η γενίκευση ενός σαφούς «ορισμού» για την ΑΔΥ, ένα διεπιστημονικό πεδίο που βασίζεται στα ΜΔΥ, στην ανάλυσή τους, στην ΑΙ κ.λπ. Απαιτείται επομένως ένας σαφής ορισμός, ώστε οι χρήστες να μπορούν να γνωρίζουν τι είναι ΑΔΥ και πώς να την προσεγγίσουν κατάλληλα, ενδεχομένως μέσω δημιουργία ενός ακριβούς πλαισίου μεθοδολογίας.

Τέταρτη μεγάλη πρόκληση είναι το «μικρό μέγεθος» των μελετών σειράς των ασθενών (patient cohorts). Αν και τα ιατρικά δεδομένα είναι συχνά ετερογενή και σημαντικά για ένα άτομο, το μέγεθος των μελετών σειράς των ασθενών, είναι συνήθως αρκετά μικρό σε σύγκριση με τις υγιείς ομάδες. Αυτή η ανισορροπία, μπορεί να προκαλέσει «υπερπροσαρμογή» (over-fitting) όχι μόνο στα κλασικά μοντέλα AI αλλά και στα σύγχρονα μοντέλα βαθιάς μάθησης (deep learning). Πιθανές λύσεις αποφυγής του over-fitting είναι η χρήση πινάκων κόστους (cost matrix), η υπερ-δειγματοληψία (oversampling), η ανάλυση ευαισθησίας κ.λπ.

Πέμπτη μεγάλη πρόκληση στην ΑΔΥ είναι η «κρίση αναπαραγωγιμότητας» (reproducibility crisis), η οποία αναφέρεται στην έκρηξη του αριθμού των δεδομένων που είναι αφενός διαθέσιμα στους ερευνητές, αλλά και στη συγκριτική έλλειψη αναλυτικών δεξιοτήτων, που είναι αφετέρου απαραίτητες, για την εξεύρεση σημασίας στα δεδομένα αυτά. Έτσι, ενώ η ικανότητά παραγωγής δεδομένων έχει αυξηθεί δραματικά, η ικανότητα κατανόησής τους, δεν έχει αναπτυχθεί με τον ίδιο ρυθμό. Για τον λόγο αυτό απαιτούνται αξιόπιστες τεχνικές επικύρωσης, που υπερβαίνουν τις τρέχουσες τεχνικές στατιστικής επικύρωσης, καθώς και ισχυρότερα πειραματικά σχέδια και αξιόπιστες στατιστικές (Peng, 2015).

Έκτη πρόκληση είναι το πρόβλημα του «απορρήτου» και της «ιδιοκτησίας», η εξασφάλιση των οποίων, εγγυάται την εμπιστευτικότητα των ιατρικών αρχείων των ασθενών από τους εργοδότες τους, την ασφαλιστική τους εταιρεία και την κοινωνία. Η ανάπτυξη νέων τεχνικών AI μπορεί να αυξήσει την απειλή κατά της ιδιωτικότητας. Ωστόσο, ο Γενικός Κανονισμός Προστασίας Δεδομένων (GDPR), υπό τις διατάξεις του οποίου εκπονείται η παρούσα εργασία, η υιοθέτηση ενός γενικά αυστηρότερου νομοθετικού πλαισίου προστασίας, καθώς και αξιόπιστες μέθοδοι κρυπτογράφησης, μπορούν να βοηθήσουν τους επιστήμονες και τους τεχνικούς, στην προστασία της ιδιωτικής ζωής των ιατρικών δεδομένων.

Υπάρχουν τουλάχιστον 5 τομείς πηγών δεδομένων που μπορούν να επηρεάσουν την ανθρώπινη υγεία και πρέπει να αναλυθούν, ώστε να οδηγήσουν σε χρήσιμα συμπεράσματα πρόγνωσης ή διάγνωσης της νόσου. Αυτά είναι: 1) τα δεδομένα ιατρικής ή υγειονομικής περίθαλψης, 2) τα δεδομένα γενετικής, γονιδιοματικής και ολοκληρωμένης βιολογίας, 3) τα δεδομένα συμπεριφοράς, 4) τα δεδομένα κοινωνικών παραγόντων και 5) τα δεδομένα φυσικού και κοινωνικού περιβάλλοντος, όπου ζουν και δραστηριοποιούνται τα υποκείμενα.



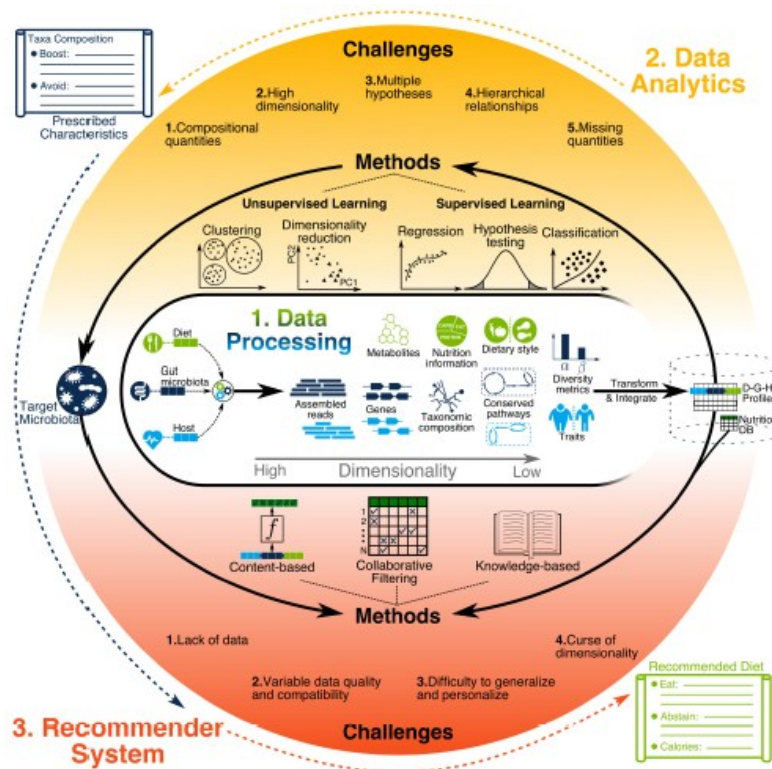
Σχ.16 Τομείς πηγών δεδομένων υγείας προς ανάλυση (<https://hitconsultant.net/>)

Παρουσιάζεται επομένως αυξημένο ενδιαφέρον, όχι μόνο για τα δεδομένα γενετικής-γονιδιοματικής, αλλά και για αυτά της συμπεριφορικής ιατρικής και για τον ρόλο των κοινωνικών και περιβαλλοντικών παραγόντων στην υγεία των ατόμων. Αφενός, η ενοποίηση δεδομένων υγείας που αφορούν μεμονωμένα άτομα (υποκείμενα), περιγράφει τον πληθυσμό. Τα δεδομένα υγείας του πληθυσμού αφετέρου, όταν συνδέονται με ένα συγκεκριμένο άτομο, περιγράφουν την έκθεση του ατόμου και τον πρόσθετο κίνδυνο που διατρέχει να εμφανίσει κάποια νόσο. Η ανάλυση των παραπάνω δεδομένων υγείας διευρύνει τις δυνατότητες εξατομικευμένης παροχής ιατρικής φροντίδας, προσαρμοσμένης στο κάθε άτομο ξεχωριστά, με την χρήση Ιατρικής Ακριβείας (Pearson et al., 2020).

Καθώς όλο και περισσότεροι επιστήμονες της βιοϊατρικής, τόσο σε ακαδημαϊκό όσο και σε επίπεδο βιοτεχνολογίας, χρησιμοποιούν τεχνολογίες υψηλής απόδοσης, όπως απεικονίσεις μικροσκοπίου και σύνθετους ελέγχους, σημειώνεται ραγδαία αύξηση στον όγκο και την ποιότητα των εργαστηριακών δεδομένων που παράγονται. Σε συνδυασμό με διάφορες εξελισσόμενες τεχνικές ML, καθίσταται πλέον όλο και περισσότερο εφικτό να αντληθεί βιολογική εικόνα από αυτά τα παραγόμενα Μαζικά Δεδομένα Υγείας, για να γίνουν κατανοητοί οι μηχανισμοί της νόσου, να εντοπισθούν νέες θεραπευτικές στρατηγικές και να βελτιωθούν τα διαγνωστικά εργαλεία, για κλινική εφαρμογή (Toh et al., 2019).

Στις σύνθετες ασθένειες είναι γνωστή η πολυπλοκότητα της πρόγνωσης της πορείας του ασθενούς. Η δυσκολία καθορισμού παραγόντων κινδύνου, η σημασία της πρόβλεψης για την πορεία της ασθένειας, ο ρόλος που διαδραματίζουν οι συννοσηρότητες, αποτελούν παραδείγματα προβλημάτων που απαιτούν την υποστήριξη δεικτών απόδοσης δεδομένων (data performance indicators), κι αποτελούν σημαντικό μέρος μιας προσέγγισης βασισμένης στα δεδομένα (data-centric). Αυτοί οι δείκτες θα πρέπει να μπορούν να αξιολογούν: 1) τη διαγνωστική ακρίβεια της ασθένειας 2) την προγενέστερη παρέμβαση, ώστε να είναι δυνατή η προληπτική χρήση δεδομένων 3) τη στοχευμένη θεραπεία, για εκμετάλλευση γνώσης από τύπους δεδομένων πολλαπλών πηγών αλληλεπίδρασης και 4) την αυξημένη αποτελεσματικότητα των φαρμάκων, για έλεγχο της γνώσης, σε σχέση με τα υπάρχοντα φαρμακευτικά αποθέματα.

Το παράδειγμα της Ιατρικής Ακρίβειας καταδεικνύει ότι τόσο στον γενικό πληθυσμό, όσο και σε εξατομικευμένες περιπτώσεις ασθενών, πρέπει να αντιμετωπιστούν παρόμοιες προκλήσεις σε σχέση με τα δεδομένα υγείας, όπως η ανάγκη συγκέντρωσης, ολοκλήρωσης, μοντελοποίησης και ερμηνείας των δεδομένων από διάφορες πηγές πληροφοριών και κλίμακες, που επηρεάζουν δυνητικά τις ασθένειες, από την έναρξη μέχρι και την εξέλιξή τους. Σε πολλές περιπτώσεις, τα προβλήματα στα δεδομένα απαιτούν υπολογιστική επεξεργασία μέσω λύσεων, χωρίς τις οποίες θα παρέμεναν δυσεπίλυτα (Carobianco, 2020).



Σχ. 17 Σχηματική αναπαράσταση επεξεργασίας και ανάλυσης δεδομένων στο προτεινόμενο σύστημα για την επίτευξη μικροβιοϊατρικού στόχου (target microbiota) (Πηγή: Eetemadi et al., 2020).

Εξάλλου, έχει διαπιστωθεί επιστημονικά ότι το μικροβίωμα του ανθρώπινου εντέρου, δηλαδή το σύνολο των γονιδίων των μικροβίων της εντερικής μικροχλωρίδας ενός ατόμου (Gordon, 2012), έχει δυνατότητα προσαρμογής στη διατροφή και σε άλλους περιβαλλοντικούς παράγοντες. Το γεγονός αυτό καθιστά το μικροβίωμα έναν εξαιρετικό στόχο για παρεμβάσεις, οι οποίες σχετίζονται με τη διατροφή, για τη βελτίωση της ανθρώπινης υγείας. Στο σημείο αυτό ακριβώς παρεμβαίνει η επιστήμη των υπολογιστών και της Ανάλυσης Δεδομένων, και επιταχύνει την σχετική έρευνα, χρησιμοποιώντας κατάλληλα εργαλεία, που συντελούν στην κατάρτιση αποτελεσματικής δίαιτας και λήψης συμπληρωμάτων διατροφής, που προσαρμόζονται σε συγκεκριμένες πληθυσμιακές ομάδες ή ακόμη και σε μεμονωμένα άτομα (Eetemadi et al., 2020).

Στην έρευνα των Eetemadi et al. (2020), μέθοδοι ανάλυσης δεδομένων χρησιμοποιούνται για να ανακαλύψουν τα χαρακτηριστικά της εντερικής μικροχλωρίδας, που έχει συνταγογραφηθεί για την επίτευξη των στόχων υγείας των ατόμων. Η επεξεργασία δεδομένων δημιουργεί διαφορετικούς τύπους πληροφοριών με διαφορετικά επίπεδα ανάλυσης και διαστάσεων.

Οι πληροφορίες πρέπει να μετασχηματιστούν και να ενσωματωθούν σε όλους τους χρήστες για τη δημιουργία μιας συλλογής. Οι μέθοδοι του προτεινόμενου συστήματος χρησιμοποιούν τη συλλογή, για να βρουν τις κατάλληλες διατροφικές συστάσεις, ώστε να βοηθήσουν τα άτομα να επιτύχουν την εντερική μικροχλωρίδα στόχο.

7.1 Προγνωστική Ανάλυση (Predictive Analytics)

Η Προγνωστική Ανάλυση (Predictive Analytics) είναι μια κατηγορία Ανάλυσης Δεδομένων (Data Analytics), που στοχεύει στη διενέργεια προβλέψεων για μελλοντικά αποτελέσματα, βάσει ιστορικών δεδομένων και τεχνικών ανάλυσης, όπως η στατιστική μοντελοποίηση (Statistical Modelling) και η Μηχανική Μάθηση (Machine Learning). Η επιστήμη της Προγνωστικής Ανάλυσης μπορεί να παράγει μελλοντικές προβλέψεις με σημαντικό βαθμό ακρίβειας. Με τη βοήθεια εξελιγμένων εργαλείων και μοντέλων πρόβλεψης, κάθε οργανισμός μπορεί πλέον να χρησιμοποιεί παλαιότερα και πιο πρόσφατα δεδομένα, για αξιόπιστες προβλέψεις τάσεων και συμπεριφορών στο μέλλον (CIO, 2019).

Ο σκοπός της Προγνωστικής Ανάλυσης είναι η διενέργεια προβλέψεων για μελλοντικά γεγονότα, με εφαρμογή σε πολλούς επιστημονικούς τομείς. Στον τομέα της υγείας για παράδειγμα, οι επαγγελματίες υγείας που μελετούν τις καρδιαγγειακές παθήσεις, μπορεί να ενδιαφέρονται για τον προσδιορισμό της εμφάνισης μεταβολών σε ορισμένους δείκτες του αίματος του ασθενούς, οι οποίοι προμηνύουν καρδιακή προσβολή.

Η σημασία της πρόληψης στην Ιατρική Ακρίβειας (Precision Medicine) έχει αυξήσει τη βαρύτητα της εξαγωγής συμπερασμάτων (inferences) μέσα από ένα ευρύ φάσμα ΜΔΥ, προκαλώντας την ανάγκη αναθεώρησης της μεθοδολογικής προσέγγισης, ώστε να αξιοποιείται πλέον η προληπτική προσαρμοστική εξαγωγή συμπερασμάτων (anticipative adaptive inference).

Αυτή η προσέγγιση αντικατοπτρίζει τις σημαντικές βελτιώσεις που επιτεύχθηκαν στον έγκαιρο εντοπισμό σημείων, ενδείξεων ή συμπτωμάτων ασθένειας, διαμέσου τεχνολογίας φορητών (wearables), κινητών (mobile) ή αισθητήρων (sensors), που εκμεταλλεύονται τη δυναμική φύση πολλών παραγόντων και παρέχουν ενημέρωση, σχετικά με την κατάσταση της υγείας του ασθενούς (Carobianco, 2020).

Σε πολλές περιπτώσεις, ένας προβλεπόμενος κίνδυνος υγείας δεν μπορεί να συνδυαστεί με μια παρέμβαση που αποδεδειγμένα θα επιφέρει ευεργετικό αποτέλεσμα. Αυτό μπορεί να οδηγήσει σε μια κατάσταση στην οποία η απόπειρα παρέμβασης κάνει περισσότερο κακό παρά καλό στον ασθενή. Ενώ λοιπόν η πρόβλεψη μπορεί να χρησιμοποιηθεί, προς το παρόν, για να προτείνει την ανάγκη για προληπτική παρέμβαση με βάση το κίνδυνο που διατρέχει ο ασθενής, καθώς εξελίσσονται οι μέθοδοι εξαγωγής συμπερασμάτων, μπορούν να χρησιμοποιηθούν στο μέλλον, για την επιλογή των παρεμβάσεων (Pearson et al., 2020)

Οι σύνθετες ασθένειες αποτελούν εξαιρετικά ετερογενείς δυναμικές διεργασίες, των οποίων τα αίτια, η πορεία της εξέλιξης, η θεραπεία και η απόκριση στη θεραπεία, καθορίζουν ένα πλήθος πιθανών θεραπευτικών οδών που θα ακολουθηθούν. Επομένως, η εξατομίκευση της θεραπευτικής αγωγής που απαιτείται, ανταποκρίνεται αποφασιστικά στην ανάγκη πρώιμων παρεμβάσεων, κατά προτίμηση στο στάδιο αναγνώρισης συγκεκριμένων μοριακών μοτίβων. Συνοπτικά, είναι κρίσιμο για τον εντοπισμό της έναρξης μιας νόσου, να επιδιώκεται συντόμευση της θεραπευτικής οδού, που υπολογίζεται ως το χρονικό παράθυρο μεταξύ μοριακής ενεργοποίησης και κλινικής παρέμβασης (Carobianco, 2020).

Συχνά, η θεραπεία σε πρώιμα στάδια οδηγεί σε σημαντικά μειωμένο κίνδυνο εξέλιξης της νόσου και η ανάπτυξη μοριακών και κλινικών βιοδεικτών (biomarkers) σε πρώιμα στάδια της νόσου, συνδέονται με άριστες πιθανότητες παρεμβατικής επιτυχίας. Οι βιοδείκτες είναι μετρήσιμα βιομόρια (δηλαδή, γονίδια, πρωτεΐνες και μεταβολίτες) που μπορούν να βοηθήσουν στην ταξινόμηση ή στον εντοπισμό ασθενών, μέσα από τον υγιή πληθυσμό. Αποτελούν ουσίες που παράγονται από φυσιολογικά ή καρκινικά κύτταρα και εντοπίζονται σε βιολογικά υγρά (αίμα, ούρα, πτύελα) ή/και σε ιστούς. Έτσι, σε δεδομένο σύνολο ασθενών, ένας βιοδείκτης μπορεί να αναγνωρίσει άτομα με κακοήγη έναντι καλοήθους όγκου. Επιπλέον, ένας βιοδείκτης θα μπορούσε να τα αναγνωρίσει ασθενείς με καρκίνο, οι οποίοι ανταποκρίνονται καλύτερα σε μια συγκεκριμένη θεραπεία (Khan et al., 2019).

Για την ανακάλυψη βιοδεικτών, η χρήση δικτύων αποτελεί μια ιδιαίτερα αποτελεσματική προσέγγιση, όταν υπάρχουν πληροφορίες σχετικές με δείγματα, προερχόμενες από πολλές πηγές. Για κάθε τύπο δεδομένων κατασκευάζεται ένα μεμονωμένο δίκτυο ομοιότητας δειγμάτων, χρησιμοποιώντας την Ευκλείδεια απόσταση, όπου οι κόμβοι του δικτύου είναι οι ασθενείς και οι άκρες του είναι, οι κατά ζεύγη, ομοιότητές τους. Τα δίκτυα για κάθε τύπο δεδομένων συνδυάζονται με επαναληπτική ενημέρωση των άκρων τους, με βάση πληροφορίες από μεμονωμένα δίκτυα, έως ότου επιτευχθεί σύγκλιση, με το τελικό συγχωνευμένο δίκτυο να ομαδοποιείται, χρησιμοποιώντας φασματική ομαδοποίηση (Mamatjan et al., 2017).

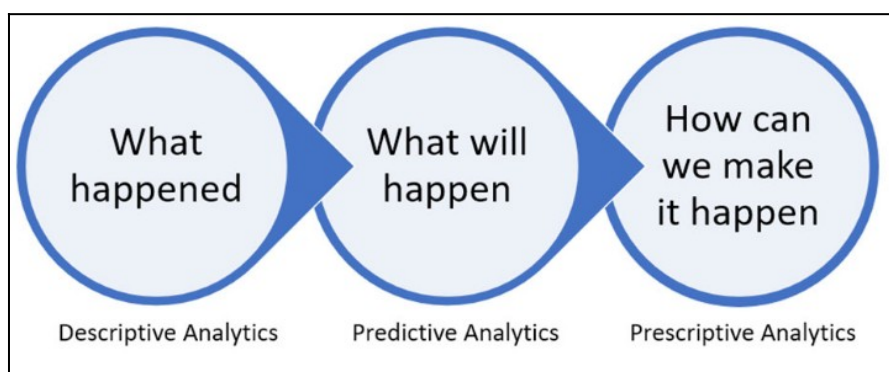
Η Προγνωστική Ανάλυση λειτουργεί καλύτερα με δομημένα (structured) δεδομένα, αλλά μπορεί να είναι επιτυχής και με αδόμητα (unstructured) δεδομένα. Για να επιτευχθούν ακριβείς προβλέψεις, απαιτούνται μεγάλες ποσότητες δεδομένων, όπως είναι τα ΜΔΥ. Στη στατιστική, ως γνωστόν, όσο περισσότερα δεδομένα συλλέγονται, τόσο πιο ακριβείς είναι οι συσχετίσεις που προκύπτουν. Για παράδειγμα όσο μεγαλύτερο είναι το μέγεθος του δείγματος, τόσο πιο ακριβή είναι τα αποτελέσματα των δημοσκοπήσεων, αφού καθώς αυξάνεται το μέγεθος του δείγματος, μειώνεται το περιθώριο σφάλματος (Hurley, 2019).

Στην Προγνωστική Ανάλυση και τη Μηχανική Μάθηση, δεν πρέπει να αγνοείται η έννοια της μετατόπισης (drift), που σημαίνει ότι η σχέση μεταξύ της μεταβλητής στόχου, την οποία το μοντέλο προσπαθεί να προβλέψει και των ανεξάρτητων μεταβλητών, αλλάζει με το χρόνο με απρόβλεπτους τρόπους. Λόγω αυτής της μετατόπισης, το μοντέλο συνεχίζει να γίνεται ασταθές και οι προβλέψεις συνεχίζουν να γίνονται εσφαλμένες, με την πάροδο του χρόνου. Ο καλύτερος τρόπος ανίχνευσης της μετατόπισης, είναι η συνεχής παρακολούθηση. Οι μετρήσεις που σχετίζονται με τη σταθερότητα του μοντέλου πρέπει να παρακολουθούνται σε συνεχή διαστήματα. Αυτό το διάστημα μπορεί να είναι μια εβδομάδα, ένας μήνας, ένα τέταρτο κ.λπ. ανάλογα με τον τομέα και την επιχείρηση. Ο τρόπος παρακολούθησης μπορεί να είναι είτε χειροκίνητος, είτε ένα αυτοματοποιημένο σενάριο που ενεργοποιεί συναγερμούς και ειδοποιήσεις, όταν παρατηρούνται ξαφνικές ανωμαλίες (Shendre, 2020).

Ενώ η πρόβλεψη μπορεί να χρησιμοποιηθεί, προς το παρόν, για να προτείνει την ανάγκη για προληπτική παρέμβαση με βάση τον κίνδυνο που διατρέχει ο ασθενής, καθώς εξελίσσονται οι μέθοδοι εξαγωγής συμπερασμάτων, μπορούν να χρησιμοποιηθούν στο μέλλον, για την επιλογή των παρεμβάσεων (Pearson et al., 2020). Έτσι, στην περίπτωση των στελεχών των ΕΔ, αν η Προγνωστική Ανάλυση οδηγήσει σε συμπέρασμα, βάσει του οποίου ένας ασθενής διατρέχει αυξημένο κίνδυνο εμφάνισης καρδιακού επεισοδίου, η παρέμβαση που θα επιλεγεί μπορεί να είναι έγκαιρη και ουσιαστική, σε επίπεδο φαρμακευτικής αγωγής ή δίαιτας ή ακόμη παράλληλα και στα δύο επίπεδα, ανάλογα με την απόφαση του θεράποντος ιατρού, κατόπιν της υπόδειξης της μεθόδου εξαγωγής συμπερασμάτων.

Η παρέμβαση μπορεί να αποφασιστεί για συγκεκριμένο χρονικό διάστημα εφαρμογής, ακόμη και με τη βοήθεια στατιστικών τιμών, όπως η τιμή NNT (Number Needed to Treat), που αντιπροσωπεύει μια εκτίμηση του αριθμού των ατόμων που πρέπει να υποβληθούν στη συγκεκριμένη θεραπεία, προκειμένου να αποφευχθεί η εμφάνιση ενός επιπλέον αρνητικού αποτελέσματος. Για παράδειγμα, σύμφωνα με σχετική έρευνα (Lehman et al., 2015), η λήψη στατίνης για διάστημα πέντε ετών, παρουσιάζει NNT 104, ενώ η πιστή τήρηση δίαιτας με μεσογειακή διατροφή για το ίδιο χρονικό διάστημα, παρουσιάζει NNT 61, για αποφυγή εγκεφαλικού, καρδιακής προσβολής ή θανάτου.

Ο κύκλος επεξεργασίας που απαιτείται μέσω της ανάλυσης των δεδομένων, για την έγκαιρη πρόγνωση του κινδύνου που διατρέχει ένας ασθενής, προϋποθέτει: 1) ψηφιακές πληροφορίες σε πραγματικό χρόνο, διαθέσιμες από διάφορες πηγές ΜΔΥ, σε κατάλληλη μορφή, 2) αποτελεσματική ανάλυση των πληροφοριών αυτών, 3) χρησιμοποίηση κατάλληλων εργαλείων για παροχή υποστήριξης αποφάσεων και καθοδήγησης, βάσει κατευθυντήριων οδηγιών που βασίζονται σε αποδείξεις (Evidence Based Guidelines - EBG), 4) παρουσίαση αυτών των πληροφοριών σε ιατρικό και νοσηλευτικό προσωπικό, σε ασθενείς και υπεύθυνους χάραξης πολιτικής, με τρόπους που θα επιτρέπουν ενσωμάτωσή τους σε νέα και ήδη υπάρχοντα μοντέλα, για βελτίωση της αποτελεσματικότητας της περίθαλψης. Όσο οι εισηγήσεις και οι ενέργειες, θα λαμβάνουν χώρα όλο και περισσότερο «σε πραγματικό χρόνο», θα απαιτηθούν νέα συστήματα για να παρακολουθούνται τα οφέλη και οι κίνδυνοι των αλγορίθμων (Pearson et al., 2020).



Σχ. 18 Τα είδη Ανάλυσης Δεδομένων: Περιγραφική, Προγνωστική, Προδιαγραφική (Πηγή: vonvia.com).

Αξίζει να σημειωθεί ότι στάδιο πριν την Προγνωστική Ανάλυση αποτελεί η Περιγραφική Ανάλυση (Descriptive Analytics). Πρόκειται για πεδίο στατιστικής, που συνήθως αποτελεί το πρώτο βήμα ανάλυσης, το οποίο εξετάζει τα ιστορικά δεδομένα, για να δώσει πληροφορίες για περαιτέρω ανάλυση. Με τη χρήση τεχνικών εξόρυξης δεδομένων, μπορούμε να κατανοήσουμε τους λόγους επιτυχίας ή αποτυχίας στην περιγραφική ανάλυση. Η Προγνωστική Ανάλυση είναι το βήμα κατά το οποίο τα ιστορικά δεδομένα συνδυάζονται με αλγόριθμους για την πρόβλεψη των αποτελεσμάτων (Kolhe, 2020). Η Προγνωστική Ανάλυση μπορεί να συμπληρωθεί από την Προδιαγραφική Ανάλυση, η οποία μελετάται αμέσως παρακάτω.

7.2 Προδιαγραφική Ανάλυση (Prescriptive Analytics)

Η Προδιαγραφική Ανάλυση συμπληρώνει την Προγνωστική, με την επίλυση προβλημάτων βελτιστοποίησης στο χώρο της υγείας, μέσω στατιστικής/μαθηματικών. Μπορεί να βελτιώσει τη λήψη αποφάσεων και την αποτελεσματικότητα των διαδικασιών, βοηθώντας τους αναλυτές να προσεγγίσουν τα αποτελέσματα, συνδυάζοντάς τα με συγκεκριμένες καταστάσεις.

Αποτελεί την επόμενη φάση προγνωστικών σεναρίων, δηλαδή είναι ένας τύπος ανάλυσης δεδομένων, για δράση σε διαφορετικά σεναρία, ιδιαίτερα για προβλήματα βελτιστοποίησης, όπου απαιτείται χρήση εξελιγμένων αλγορίθμων, για την εύρεση της βέλτιστης λύσης.

Αυτός ο τύπος προβλημάτων, σύμφωνα με πρόσφατη έρευνα (Lopes et al., 2020), μπορεί να επιλυθεί με τέσσερις ξεχωριστές προσεγγίσεις: 1) Τυφλή Αναζήτηση (Blind Search), 2) Τοπική Αναζήτηση (Local Search), Αναζήτηση βάσει Πληθυσμού (Search based on Population) και 4) Βελτιστοποίησης πολλαπλών Στόχων (Multi-objective Optimization).

Η Τυφλή Αναζήτηση αναλαμβάνει την εξάντληση όλων των εναλλακτικών λύσεων, διασφαλίζοντας ότι έχουν δοκιμαστεί όλες οι λύσεις στο χώρο των αποτελεσμάτων, καθώς και τη βεβαιότητα ότι πάντοτε θα βρεθεί η βέλτιστη λύση.

Σε αντίθεση με την Τυφλή Αναζήτηση, οι πιο σύγχρονες τεχνικές βελτιστοποίησης βασίζονται στην Τοπική Αναζήτηση, όπου νέες λύσεις παράγονται από ήδη υπάρχουσες λύσεις, συμπεριλαμβανομένων μεθόδων, οι οποίες εστιάζουν σε μια τοπική γειτονιά (neighborhood), μέσω μιας δεδομένης αρχικής λύσης.

Η Αναζήτηση βάσει Πληθυσμού εξερευνά διαφορετικότερες περιοχές του χώρου αναζήτησης, επιτυγχάνοντας μεγαλύτερη ποικιλομορφία ως προς τον καθορισμό νέων λύσεων, οι οποίες μπορούν να προκύψουν όχι μόνο αλλάζοντας ελαφρώς κάθε μεμονωμένο σημείο αναζήτησης (search point), αλλά και συνδυάζοντας χαρακτηριστικά (attributes), που σχετίζονται με δύο ή περισσότερα σημεία αναζήτησης.

Τέλος, η Βελτιστοποίηση πολλαπλών Στόχων προκύπτει από την ανάγκη συνδυασμού πολλαπλών στόχων σε ένα μόνο αλγόριθμο, πολύ δε περισσότερο, αν η επίτευξη ορισμένων από αυτούς, προϋποθέτει την απώλεια άλλων, προκειμένου να επιτευχθεί η μεγαλύτερη δυνατή ισορροπία (Alharthi, 2018).

Η Προδιαγραφική Ανάλυση ως συμπλήρωμα της Προγνωστικής αποτελεί την τελική λύση υποστήριξης, για ένα πλήρως Ευφυές Σύστημα (Intelligent System), το οποίο δύναται να αποφεύγει σενάρια, καθώς και να προτείνει βέλτιστες ενέργειες. Έτσι, με την εισαγωγή των μαθηματικών και της στατιστικής στην ανάπτυξη αλγορίθμων, καθίσταται δυνατή η πλήρης διαχείριση ενός νοσοκομείου. Η βελτιστοποίηση των υπάρχοντων πόρων, μπορεί για παράδειγμα να επιτρέψει στο νοσοκομείο να συντομεύσει τις υπάρχουσες γραμμές αναμονής, ενώ με την ενσωμάτωση ενός συστήματος ικανού να προσδιορίζει προβλήματα υγείας, που δεν μπορούν να επιλυθούν εκτός νοσοκομείου, αυξάνονται οι δυνατότητες παροχής θεραπείας, σε όσους πραγματικά την χρειάζονται (Lopes et al., 2020).

7.3 Ιατρική Ακριβείας (Precision Medicine)

Σε πρόσφατη έρευνα (Hulsen et al., 2019) αναζητήθηκαν προοπτικές για το πώς μπορεί να σημειωθεί πρόοδος με τη χρήση ΜΔΥ, ώστε να φτάσουμε σε στρατηγικές «Ιατρικής Ακριβείας» (Precision Medicine). Η Ιατρική Ακριβείας πρέπει να αποτελεί στόχο της σύγχρονης ιατρικής επιστήμης, καθώς ορίζεται ως η παροχή της κατάλληλης θεραπείας, την κατάλληλη στιγμή, στον συγκεκριμένο ασθενή. Στην έρευνα αυτή εντοπίστηκαν βασικές εννοιολογικές και δομικές προκλήσεις, στις στρατηγικές Ιατρικής Ακριβείας που πρέπει να υιοθετηθούν και διαπιστώθηκε η ανάγκη τυποποίησης του περιεχομένου δεδομένων, της μορφής τους και των κλινικών ορισμών, η ανάγκη συνεργατικών δικτύων με κοινή χρήση δεδομένων και εξειδίκευσης. Επίσης, καταδείχθηκε ότι η έρευνα «παραγωγής υποθέσεων» (hypothesis-generating research) αλληλοσυμπληρώνεται με την παραδοσιακή μέθοδο υποθέσεων, που χρησιμοποιούσε ως τώρα η ιατρική επιστήμη.

Μέχρι πρόσφατα η πλειοψηφία των ιατρικών εξετάσεων αφορούσε στη διάγνωση της νόσου και σύμφωνα με μελέτες, η συμβολή τους στην πρόληψη ήταν χαμηλή. Νέες εξετάσεις χρησιμοποιούνται πλέον, οι οποίες αντί να στοχεύουν στη διάγνωση της νόσου, εντοπίζουν την επίδραση του περιβάλλοντος, του τρόπου ζωής, της διατροφής και των ελλείψεων στην υγεία. Η σύγχρονη αυτή προσέγγιση της Ιατρικής Ακριβείας χρησιμοποιεί εργαλεία, που επιτρέπουν στον ιατρό να παρέχει στο κάθε άτομο εξατομικευμένες θεραπευτικές οδηγίες, για πρόληψη ή θεραπεία χρόνιων νοσημάτων, όπως ο διαβήτης, οι καρδιολογικές παθήσεις, ο καρκίνος και τα αυτοάνοσα νοσήματα.

Το Εθνικό Συμβούλιο Έρευνας (National Research Council) υιοθέτησε τον ορισμό της Ιατρικής Ακριβείας το 2008, ορίζοντάς την ως την: «Προσαρμογή της ιατρικής περίθαλψης στα ιδιαίτερα χαρακτηριστικά κάθε ασθενούς, προς ταξινόμηση των ασθενών σε υπο-πληθυσμούς, που διαφέρουν ως προς την ευαισθησία τους σε μια συγκεκριμένη ασθένεια ή την απόκρισή τους σε μια συγκεκριμένη θεραπεία. Οι προληπτικές ή θεραπευτικές παρεμβάσεις μπορούν στη συνέχεια να επικεντρωθούν σε εκείνους που θα ωφεληθούν, εξοικονομώντας έξοδα και παρενέργειες για εκείνους που δεν θα το κάνουν».

Μέχρι σήμερα οι θεραπείες ήταν ενιαίες και αφορούσαν ασθενείς που είχαν παρόμοια συμπτώματα ή έπασχαν από την ίδια πάθηση. Κι αυτό διότι στοιχεία για αποτελεσματικότερες θεραπείες προέρχονται από κλινικές μελέτες που προτείνουν αγωγές με τα καλύτερα αποτελέσματα για τον «μέσο ασθενή».

Έτσι στο παράδειγμα χρήσης του ενιαίου φαρμάκου για όλες τις περιπτώσεις, αποτελεσματική θεραπεία είναι η θεραπεία κατά τεκμήριο ωφελεί το μεγαλύτερο μέρος του πληθυσμού-στόχου, ο οποίος συνήθως εντοπίζεται, χρησιμοποιώντας την έννοια του απαιτούμενου για θεραπεία αριθμού. Ο «αριθμός που απαιτείται για θεραπεία» (Number Needed to Treat - NNT) είναι ένα μέτρο που υποδεικνύει το μέσο αριθμό ατόμων, που χρειάζονται θεραπεία, προς αποφυγή ενός επιπλέον κακού αποτελέσματος. Για παράδειγμα, η χρησιμοποιημένη θεραπεία για τη χοληστερόλη έχει NNT 20, που σημαίνει ότι μόνο 1 στους 20, από αυτούς που υποβάλλονται σε θεραπεία θα έχουν τελικά όφελος από την εν λόγω θεραπεία (Schork, 2015). Οι υπόλοιποι όχι μόνο δεν θα ωφεληθούν, αλλά μπορεί η θεραπεία να επιδεινώσει την κατάστασή τους.

Ο κάθε ασθενής όμως, αποτελεί μοναδική βιολογική οντότητα, που αλληλοεπιδρά με το περιβάλλον και χρειάζεται διαφορετική αντιμετώπιση, σε διαφορετικές στιγμές στη ζωή του. Αυτή η ανάγκη της εξατομικευσης των θεραπειών οδήγησε στην Ιατρική Ακριβείας.

Ένας σημαντικός περιορισμός της σημερινής ιατρικής είναι η κακή κατανόηση της βιολογίας των ασθενειών. Μόνο συγκεντρώνοντας τεράστιες ποσότητες ΜΔΥ, όλες οι σχετικές μεταβλητές πολλαπλών πόρων, όπως δεδομένα DNA, RNA, πρωτεΐνης και μεταβολισμού θα μπορούν να ενσωματωθούν σε πιο ρεαλιστικά μοντέλα για να προβλέψουν πώς θα συμπεριφερθούν, για παράδειγμα, οι όγκοι και ποιοι ασθενείς θα ωφεληθούν καλύτερα από συγκεκριμένες, στοχευμένες θεραπείες. Αυτά τα ενσωματωμένα δεδομένα (omics) θα μπορούν να παρέχουν πιο ολοκληρωμένη εικόνα της βιολογικής συμπεριφοράς καρκινικών όγκων, των μηχανισμών που διέπουν την ανάπτυξή τους, το μεταστατικό δυναμικό τους, καθώς και την ανταπόκρισή τους σε στοχευμένες θεραπείες (Willems et al., 2019).

Πλέον με τη συλλογή στοιχείων από το DNA και την ανάλυση της αλληλουχίας του γονιδιώματος (genome sequencing), τον τρόπο ζωής και τη διατροφή, το ιατρικό ιστορικό, το μικροβίωμα (το σύνολο των μικροβίων που υπάρχουν φυσιολογικά στον οργανισμό), τους μεταβολίτες (μικρά μόρια που συμμετέχουν στις χημικές αντιδράσεις που μας κρατούν στη ζωή), είναι δυνατή η παροχή εξατομικευμένης θεραπείας και οδηγιών, προσαρμοσμένες στον κάθε ασθενή ξεχωριστά.

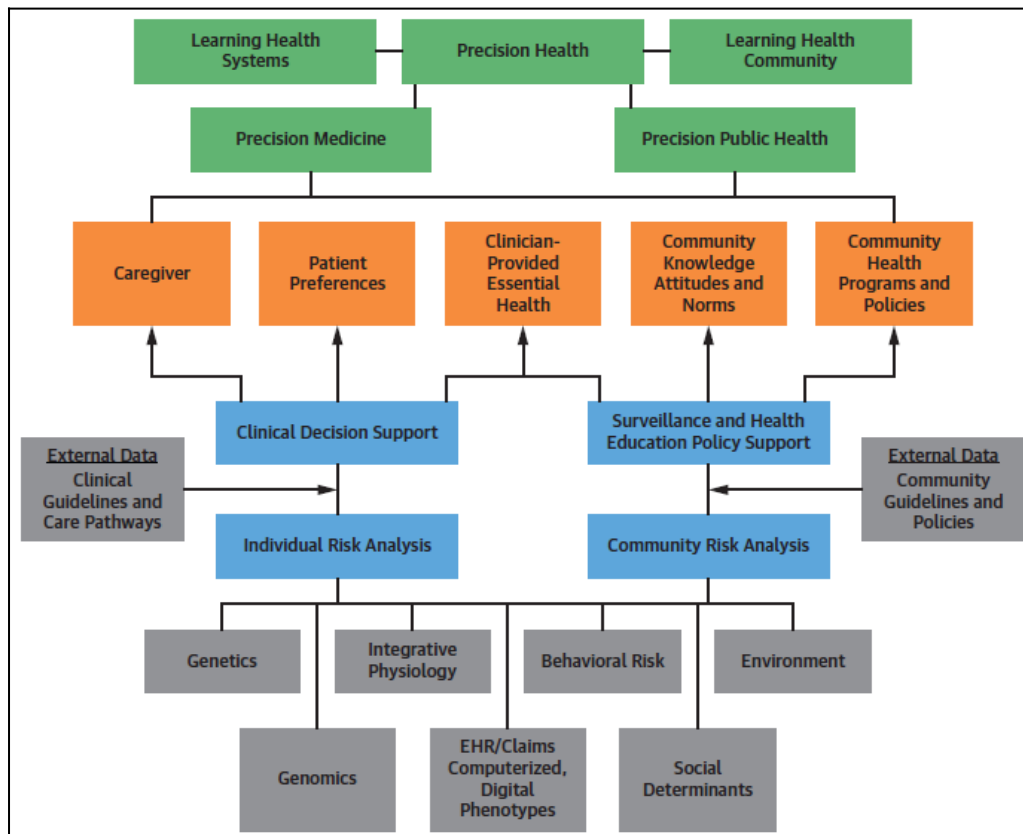
Έτσι, αναφορικά με τα στελέχη των ΕΔ και την πρόληψη ή θεραπεία κάποιων νόσων τους, διαφορετικά από το συνηθισμένο θα αντιμετωπιστεί η περίπτωση του στελέχους, που έλαβε μέρος σε ειρηνευτικές αποστολές και εκτέθηκε σε περιβάλλον ενδεχομένως μολυσμένο από ρίψεις χημικών, με ζημιογόνα για τον οργανισμό συστατικά ή διαβίωσε σε ανθυγιεινές τοπικές ατμοσφαιρικές συνθήκες ή ακολουθούσε διατροφή φτωχή σε θρεπτικά συστατικά ή δεν είχε ευκαιρίες εκγύμνασης, επαρκούς ξεκούρασης, ύπνου κ.λπ.

Η προσέγγιση αυτή επιτρέπει να διαχειριστεί κανείς τους πραγματικούς αιτιολογικούς παράγοντες που οδηγούν σε μια νόσο, αντί να διαχειρίζεται μόνο τα συμπτώματά της, αφότου αυτή εμφανιστεί. Η Ιατρική Ακριβείας ολοκληρώνει, ως επιπρόσθετο εργαλείο για τον ιατρό, τις υπάρχουσες διαγνωστικές και θεραπευτικές λύσεις. Η μέτρηση μικρών μορίων, παρέχει στοιχεία πολύ πριν την εμφάνιση της νόσου. Εξετάσεις μεταβολομικής ανιχνεύουν ανισορροπίες σε μοριακό επίπεδο, αρκετά χρόνια πριν την εκδήλωση της νόσου. Πρόσφατη έρευνα (Song et al., 2020) κατέδειξε τα πλεονεκτήματα της στοχευμένης με ακρίβεια πρόγνωσης, πρόληψης και θεραπείας συστηματικών νόσων και καρκίνων.

Έτσι μέσα από την έρευνα φάνηκε ότι η Ιατρική Ακριβείας, εκμεταλλεύομενη τις δυνατότητες της Ανάλυσης Δεδομένων και της Τεχνητής Νοημοσύνης, μπορεί να εφαρμοστεί πολύ αποτελεσματικά σε νόσους του

αναπνευστικού, του κυκλοφορικού, του ουροποιητικού, του πεπτικού, του ενδοκρινολογικού και άλλων συστημάτων του ανθρώπινου σώματος, όπως και σε περιπτώσεις ογκολογικών περιστατικών στον πνεύμονα, στον θυρεοειδή, στο συκώτι, στον προστάτη, στο μαστό, ενώ επισημάνθηκαν περιπτώσεις, όπου η προσωποποιημένη ιατρική μπορεί να εκμεταλλευτεί δυνατότητες τόσο της δυτικής, όσο και της παραδοσιακής Κινεζικής Ιατρικής.

Στο ανθρώπινο σώμα συμβαίνουν χιλιάδες χημικές αντιδράσεις στις οποίες συμμετέχουν πολύ μικρά μόρια, που δεν είναι εφικτό να μετρηθούν μέσα από τις κοινές εργαστηριακές εξετάσεις. Τα τελευταία 20 χρόνια οι τεχνολογικές εξελίξεις έχουν καταστήσει δυνατή τη μέτρηση αυτών των μορίων, που ονομάζονται «μεταβολίτες» και παρέχουν στοιχεία στον θεράποντα ιατρό, ώστε να μην χρειάζεται να περιμένει κανείς να αρρωστήσει και να προσπαθήσει στη συνέχεια να αλλάξει την πορεία της υγείας του. Ο ιατρός μπορεί πλέον να στοχεύσει με ακρίβεια και να προσαρμόσει οδηγίες και παρεμβάσεις εξατομικευμένες για κάθε άτομο, με στόχο τη βέλτιστη δυνατή υγεία του. Συγχωνευμένα, μεγάλα σύνολα δεδομένων από τις παραπάνω εξετάσεις (Big Data), επιτρέπουν πλέον πρόωρη πρόβλεψη κινδύνου και διαφορική διάγνωση, καθιστώντας τελικά δυνατή τη σύσταση των κατάλληλων θεραπειών, στην κατάλληλη δόση, στο συγκεκριμένο άτομο-ασθενή και στη σωστή στιγμή.



Σχ. 19 Η ολοκληρωμένη προσέγγιση της «Υγείας Ακρίβειας» (Πηγή: Pearson et al., 2020)

Για την εφαρμογή της Ιατρικής Ακρίβειας, από την πλευρά των δεδομένων υγείας απαιτούνται τέσσερα (4) δομικά στοιχεία: 1) τα δεδομένα υγείας του υποκειμένου, σε συνδυασμό με τα στοιχεία του από την κοινότητα στην οποία ζει/δραστηριοποιείται 2) οι μέθοδοι της επιστήμης των δεδομένων (data science) για την ενσωμάτωση στα παραπάνω και των επίκτητων δεδομένων, 3) η χρήση του συνόλου των δεδομένων με αποτελεσματικό τρόπο, προς αλλαγή συμπεριφορών και μείωση του κινδύνου ασθένειας και 4) η βελτίωση των δεξιοτήτων και των γνώσεων του ιατρικού-νοσηλευτικού προσωπικού για επίδραση της Ιατρικής Ακρίβειας στην κοινότητα (Pearson et al., 2020).

Στην έρευνά τους οι Pearson et al. (2020) κατέληξαν στην προσέγγιση της «Υγείας Ακριβείας», δηλαδή σε μια πιο ολοκληρωμένη προσέγγιση, η οποία προέρχεται από τη συγχώνευση της Ιατρικής Ακριβείας και της εφαρμογής της στη δημόσια υγεία. Η συγχώνευση υλοποιείται, όπως φαίνεται στο παρακάτω σχήμα, από την ενσωμάτωση τομέων δεδομένων (γκρίζα πλαίσια) με την οποία νέα δεδομένα των υποκειμένων και της κοινότητας συνδυάζονται με εξωτερικά δεδομένα, για να εκτιμηθεί πλήρως ο κίνδυνος που διατρέχουν τόσο τα υποκείμενα όσο και η κοινότητα. Οι παράλληλες αναλυτικές μέθοδοι ακριβείας (μπλε πλαίσια) δημιουργούν τα εργαλεία υποστήριξης αποφάσεων για ιατρούς και εργαζόμενους στον τομέα της δημόσιας υγείας. Αυτά τα εργαλεία είναι στη συνέχεια διαθέσιμα σε ένα ευρύτερο φάσμα ενδιαφερομένων (πορτοκαλί πλαίσια) που μπορούν να εφαρμόσουν και να αξιολογήσουν την αποτελεσματικότητα των προϊόντων ανάλυσης ακριβείας. Ένα κλινικό ή κοινοτικό ζήτημα υγείας, σε ένα αποτελεσματικό σύστημα υγείας ακριβείας, θα μπορούσε στη συνέχεια να αντιμετωπιστεί στο υποκείμενο, στην κοινότητα ή σε συνδυασμένα επίπεδα. Αν τα υποκείμενα της προτεινόμενης προσέγγισης είναι στελέχη των ΕΔ και η κοινότητα το στράτευμα, τότε η «Υγεία Ακριβείας» θα μπορούσε αναλογικά να εφαρμοστεί στις ΕΔ.

7.4 Μεταγραφική Ιατρική (Translational Medicine)

Η Μεταγραφική Ιατρική (translational medicine) είναι η διαδικασία κατά την οποία η βασική έρευνα μεταβαίνει από τις ανακαλύψεις στο εργαστήριο σε πραγματικές κλινικές εφαρμογές. Αποτελεί σημαντικό νέο επιστημονικό κλάδο, που απαιτεί τη συνεργασία μεταξύ πολλαπλών ειδικοτήτων, αλλά και την εκμετάλλευση μεγάλων συνόλων ιατρικών δεδομένων.

Αριθμός αναλύσεων υψηλής απόδοσης παράγουν δεδομένα από πολλά δείγματα ασθενών. Τα σύνολα δεδομένων στη συνέχεια διαρθρώνονται σε αναγνώσιμη μορφή από τον ηλεκτρονικό υπολογιστή και πιθανώς αναγνωρίζονται σημαντικές μεταβλητές, χρησιμοποιώντας έναν αλγόριθμο ML. Ο αλγόριθμος θα «μάθει» τις σχέσεις μεταξύ των μεταβλητών και θα μπορεί να εκτελέσει έξυπνες εργασίες, όπως η ομαδοποίηση ασθενών ή η πρόβλεψη των αποτελεσμάτων τους.

Πριν από την εφαρμογή της ML στη Μεταγραφική Ιατρική, υπάρχουν αρκετά μη τεχνικά ζητήματα που πρέπει να αντιμετωπιστούν. Συχνά, για την εκπαίδευση των αλγορίθμων ML απαιτείται η χρησιμοποίηση ευαίσθητων προσωπικών δεδομένων. Η πρόσβαση στα δεδομένα αυτά πρέπει να πραγματοποιείται προσεκτικά, ώστε να εξασφαλίζεται η προστασία της ιδιωτικής ζωής, χωρίς παράλληλα να εμποδίζεται η καινοτομία και η τεχνολογική πρόοδος, για τη βελτίωση των αποτελεσμάτων.

Στα σύνολα δεδομένων, που χρησιμοποιούνται για την εκπαίδευση των αλγορίθμων ML, πρέπει να αποφεύγεται η μεροληπτική (biased) χρήση τους, για μείωση του κινδύνου αποτυχίας των μεθόδων ML, ως προς την γενίκευση των αποτελεσμάτων. Η επανεξέταση της ευθύνης και της λογοδοσίας των ατόμων ή των οργανισμών, που επιλέγουν σύνολα δεδομένων για εκπαίδευση αλγορίθμων ML, είναι βασικές παράμετροι για να αντιμετωπιστεί αυτό. Έτσι, πρέπει να θεσπιστούν δεοντολογικά πλαίσια από επιστημονικές επιτροπές και ρυθμιστικούς φορείς, για να αναγνωρίσουν και να ελαχιστοποιήσουν την επίδραση των μεροληπτικών (biased) μοντέλων, καθοδηγώντας παράλληλα τις σχεδιαστικές επιλογές, για εισαγωγή συστημάτων που οικοδομούν εμπιστοσύνη, κατανόηση και διατήρηση της ιδιωτικής ζωής (Toh et al., 2019).

Η αναπαραγωγιμότητα είναι μια άλλη πτυχή που πρέπει να αντιμετωπιστεί για να διασφαλιστεί η ευρεία υιοθέτηση της ML στη Μεταγραφική Ιατρική. Πρέπει να δίδεται προσοχή κατά την εξαγωγή συμπερασμάτων αποκλειστικά από μεγάλες δεξαμενές κλινικών δεδομένων, καθώς συχνά είναι ετερογενή ως προς την ποιότητά τους. Η υπεύθυνη κοινή χρήση δεδομένων και κώδικα θα πρέπει να είναι απαιτήσεις για τους συγγραφείς, παράλληλα με τις δημοσιεύσεις τους, για να διασφαλιστεί η εμπιστοσύνη στα αποτελέσματα της έρευνάς τους.

Αποτελεί πρωταρχικής σημασίας η προσήλωση στο προς απάντηση βιολογικό ερώτημα που τίθεται κάθε φορά, καθώς και στα δεδομένα που απαιτούνται για την απάντησή του, αντί να εφαρμόζονται αδιάκριτα τεχνικές ML, σε οποιοδήποτε σύνολο δεδομένων.

Ωστόσο οι εξελίξεις στο χώρο της ML φέρνουν ολοένα πιο αντικειμενικές μεθόδους για την αντιμετώπιση της ασφάλειας που παρατηρείται στη Μεταγραφική Ιατρική, επιτρέποντας πιο σταθερές, βάσει δεδομένων λήψεις αποφάσεων, για την επόμενη γενιά διαγνωστικών εργαλείων και θεραπειών στους ασθενείς.

Τέλος, υπάρχει η ανάγκη εξήγησης και εύκολης ερμηνείας των προβλέψεων των συστημάτων ML για την εφαρμογή τους σε κλινικές ρυθμίσεις. Οι αλγόριθμοι «μαύρου κουτιού» (black-box) μπορεί να έχουν καλή ακρίβεια πρόβλεψης, οι προβλέψεις τους όμως είναι δύσκολο να ερμηνευθούν, δεν είναι εφικτές, περιορίζοντας την κλινική τους εφαρμογή. Ευτυχώς, έχουν επινοηθεί νέες μέθοδοι που επιτρέπουν στους ερευνητές να ερμηνεύουν αλγόριθμους «μαύρου κουτιού» για να διασφαλίσουν ότι οι προβλέψεις τους είναι λογικές (Toh et al., 2019).

8. Ασφάλεια Δεδομένων Υγείας - Διατάξεις και Περιορισμοί του ΓΚΠΔ

Παραδοσιακά, ο όρος ασφάλεια δεδομένων (data security) χρησιμοποιείται για να περιγράψει τη μεθοδολογία, καθώς και τις μεθόδους και τεχνικές που ακολουθούνται, προκειμένου να επιτευχθούν οι εξής στόχοι: 1) Εμπιστευτικότητα (confidentiality), δηλαδή ότι τα δεδομένα δεν πρέπει να αποκαλύπτονται σε μη εξουσιοδοτημένα άτομα, 2) Ακεραιότητα (integrity), δηλαδή ότι τα δεδομένα πρέπει να είναι ακριβή, ακέραια και γνήσια – όχι εσφαλμένα, αλλοιωμένα ή μη ενημερωμένα, 3) Διαθεσιμότητα (availability), δηλαδή ότι τα δεδομένα πρέπει να είναι στη διάθεση των χρηστών, όποτε απαιτείται η χρήση τους.

Στη σημερινή εποχή, κατά την οποία κρίσιμες λειτουργίες της κοινωνίας βασίζονται σε πληροφοριακά συστήματα, ένας φορέας οφείλει να εξασφαλίζει ότι τα συστήματα και οι εφαρμογές του θα συνεχίζουν να λειτουργούν υπό δυσμενείς συνθήκες, όπως μετά από ένα φυσικό ή τεχνικό περιστατικό, και ότι θα είναι σε θέση να τα επαναφέρει σε λειτουργία. Έτσι, στους τρεις ανωτέρω στόχους της ασφάλειας, ο ΓΚΠΔ προσθέτει και την αξιοπιστία (resilience) των συστημάτων (ΑΠΔΠΧ, 2021).

Παρά τις συμβατικές προβλέψεις και τη λήψη μέτρων ασφαλείας, ελλοχεύει πάντα ο κίνδυνος απώλειας δεδομένων, μη εξουσιοδοτημένης πρόσβασης τρίτων οντοτήτων σε προσωπικά δεδομένα και μεταδεδωμένα επικοινωνιών (cloud, τηλεφωνία, διευθύνσεις IP, online προτιμήσεις, προφίλ πελατών, δεδομένα ροής δρομολογητή κ.λπ.), αλλά και λοιπές κρίσιμες πληροφορίες (ιδιωτικά απόρρητα, εφευρέσεις, έρευνες) κάθε εταιρείας ή οργανισμού (Κανέλλος, 2020).

Μιλώντας για ΜΔΥ και μάλιστα στελεχών των Ενόπλων Δυνάμεων, πιθανή απώλεια των οποίων θα στοίχιζε τόσο σε προσωπικό όσο ενδεχομένως και σε εθνικό επίπεδο, καθίσταται σαφές ότι η ασφάλεια και η ιδιωτικότητα των χρησιμοποιούμενων δεδομένων είναι πρωταρχικής σημασίας. Οι συγκλίνουσες τεχνολογίες που εξετάζονται στην παρούσα εργασία, θα μπορούσαν να χρησιμοποιηθούν με μεγάλες όμως επιφυλάξεις και εγγυήσεις, ως προς τη χρήση της υπολογιστικής νέφους, όπου προς το παρόν, η ασφάλεια των δεδομένων δεν μπορεί να είναι εγγυημένη.

Στην κατεύθυνση αυτή, σχεδιάστηκε και προτάθηκε από τους Kaur et al., (2018) ένα εφευρετικό και ασφαλές μοντέλο υγειονομικής περίθαλψης τεσσάρων επιπέδων (layers). Το κατώτερο επίπεδο πηγής δεδομένων (data source) ασχολείται με ετερογενείς πηγές δεδομένων. Αυτό το επίπεδο έχει σκοπό να διαχειρίζεται ετερογενή δεδομένα και να τα μετασχηματίζει σε ομοιογενή, εάν απαιτείται.

Το δεύτερο επίπεδο αποθήκευσης δεδομένων (data storage) προορίζεται για διαχείριση της διαδικασίας βελτιστοποίησης της αποθήκευσης. Το τρίτο επίπεδο ασφάλειας δεδομένων και ιδιωτικότητας (data security) παρέχει αρκετές προηγμένες εγγυήσεις και χαρακτηριστικά προστασίας ιδιωτικού απορρήτου, όπως η απόκρυψη δεδομένων (data masking), η παρακολούθηση δραστηριότητας, η ομοιομορφική κρυπτογράφηση, ο ενσωματωμένος έλεγχος ταυτότητας (Pluggable Authentication Module) και ο λεπτομερής έλεγχος πρόσβασης (granular access control).

Τέλος, το τέταρτο επίπεδο εφαρμογής Μηχανικής Μάθησης είναι υπεύθυνο για διεργασίες όπως η διάγνωση της νόσου, η ανακάλυψη φαρμάκων, η Ανάλυση Δεδομένων και η υποστήριξη της οπτικοποίησης. Εδώ, τεχνικές Μηχανικής Μάθησης όπως η παραδοσιακή εξόρυξη δεδομένων (data mining), μπορούν να χρησιμοποιηθούν για την έγκαιρη διάγνωση της νόσου. Η ακρίβεια στη διάγνωση ασθενειών μπορεί να ενισχυθεί περαιτέρω, ενσωματώνοντας τις έννοιες της λογικής fuzzy και της θεωρίας των πληροφοριών (Kaur et al., 2018).

Παρόμοιας στόχευσης είναι και το σύστημα που προτάθηκε από τους Kumar et al., (2018). Το περιβάλλον του νέφους παρέχει επαρκή χώρο για την αποθήκευση του μεγάλου όγκου ιατρικών δεδομένων, που αποθηκεύονται σε Hadoop με δυνατότητα κλιμάκωσης, δηλαδή σε σύνολο προγραμμάτων και διαδικασιών ανοιχτού κώδικα, που μπορούν να χρησιμοποιηθούν από τον καθένα, ως backbone για λειτουργίες Big Data. Επειδή η ασφάλεια των δεδομένων σε περιβάλλον νέφους αποτελεί μεγάλη πρόκληση, το προτεινόμενο σύστημα εισηγείται έναν «ασφαλή» αλγόριθμο αποθήκευσης ιατρικών δεδομένων.

Έτσι, χρησιμοποιείται ένας μηχανισμός αποθήκευσης, ο οποίος συλλέγει τα δεδομένα από μια μονάδα συλλογής. Αυτά αποθηκεύονται σε πέντε διαφορετικά στάδια, που περιλαμβάνουν, αποθήκευση, ανάκτηση, συγκέντρωση, διαμερισματοποίηση και συγχώνευση δεδομένων, ενώ τα «ασφαλή» ιατρικά δεδομένα μπορούν να αποθηκευτούν ξανά στη βάση δεδομένων του Νέφους. Ο προτεινόμενος αλγόριθμος λειτουργεί σε δύο κυρίως φάσεις, στη φάση της ασφαλούς αποθήκευσης (secured storage) και στη φάση της ανάκτησης (retrieval). Στα συμπεράσματα της έρευνας επισημαίνεται ότι προηγμένοι αλγόριθμοι κρυπτογράφησης θα εγγυώνται μεγαλύτερη ασφάλεια των δεδομένων υγείας στο μέλλον.

Τα ΜΔΥ, επειδή συνήθως βρίσκονται αποθηκευμένα στο πληροφοριακό σύστημα κάποιου νοσηλευτικού ιδρύματος, φαίνεται να ανήκουν σε αυτό, το οποίο αποτελεί και τον υπεύθυνο επεξεργασίας κατά τον ΓΚΠΔ. Ωστόσο, το κάθε ίδρυμα ενεργεί απλώς ως ο «θεματοφύλακας» αυτών των δεδομένων, αφού τα δεδομένα είναι ιδιοκτησία του ασθενούς και η πρόσβαση και χρήση τους, εκτός του κλινικού χώρου, απαιτεί τη ρητή συγκατάθεση του ίδιου του ασθενούς.

Αυτό το νομικά κατοχυρωμένο δικαίωμα αποτελεί ταυτόχρονα τροχοπέδη στην ταχεία εκμετάλλευση του μεγάλου όγκου δεδομένων υγείας, που διατηρούνται ήδη σε κλινικά αρχεία. Ενώ οι έρευνες οι βασισμένες σε υπόθεση (hypothesis driven) μπορούν να πραγματοποιηθούν σε συγκεκριμένα, ανώνυμα δεδομένα, μόλις αυτές ολοκληρωθούν, τα δεδομένα θα πρέπει να καταστραφούν. Για τεχνικές ΜΔΥ που χρησιμοποιούν χιλιάδες έως εκατομμύρια σημεία δεδομένων και που μπορεί να απαιτούν σημαντική επεξεργασία, η προοπτική απώλειας τέτοιου είδους πολύτιμων δεδομένων, στο τέλος του έργου, είναι απογοητευτική για την πρόοδο της ιατρικής γνώσης (Hulsen et al., 2019).

Σύμφωνα με μελέτη των Tao et al. (2019), οι ιατρικές πληροφορίες αξίζουν, στη μαύρη αγορά, 10 φορές περισσότερο από τον αριθμό της πιστωτικής κάρτας ενός ατόμου. Το FBI προειδοποίησε τους παρόχους υγειονομικής περίθαλψης να προφυλάσσονται από επιθέσεις στον κυβερνοχώρο, αφότου η Community Health Systems Inc, μια από τις μεγαλύτερες εταιρίες διαχείρισης 206 νοσοκομείων σε 29 πολιτείες των ΗΠΑ, ανακοίνωσε το 2014, ότι χάκερς είχαν εισέλθει στο δίκτυο υπολογιστών της και έκλεψαν προσωπικές πληροφορίες 4,5 εκατομμυρίων ασθενών. Ειδικοί στον τομέα της κυβερνο-ασφάλειας εκτιμούν ότι οι κυβερνο-εγκληματίες στοχοποιούν προοδευτικά πάνω από 3 τρισεκατομμύρια δολάρια, στον κλάδο υγειονομικής περίθαλψης των ΗΠΑ, όπου πολλές εταιρείες και επιχειρήσεις εξακολουθούν να εξαρτώνται από παλιά συστήματα δικτύων υπολογιστών, χωρίς να είναι εξοπλισμένα με νέες δυνατότητες και εργαλεία ασφάλειας στον κυβερνοχώρο.

Υπάρχουν δύο τύποι κοινών απειλών στον κυβερνοχώρο, τους οποίους συνήθως αντιμετωπίζει η βιομηχανία της υγειονομικής περίθαλψης: οι παραβιάσεις δεδομένων και η απαίτηση ψηφιακών λύτρων (ransomware). Σε σύγκριση με οποιονδήποτε άλλο κλάδο, η βιομηχανία υγειονομικής περίθαλψης έχει παράγοντες πολύ υψηλού κινδύνου, όταν αντιμετωπίζει παραβιάσεις δεδομένων και επιθέσεις ransomware.

Τα αρχεία υγειονομικής περίθαλψης θεωρούνται πολύτιμα για τους κυβερνο-εγκληματίες, λόγω της αφθονίας προσωπικών, ιατρικών και οικονομικών πληροφοριών που περιέχονται σε κάθε ατομικό ηλεκτρονικό φάκελο υγείας (ΑΗΦΥ). Με την πρόσβαση σε αυτές τις πληροφορίες, διαπράττεται κλοπή ταυτότητας, ασφαλιστική και οικονομική απάτη, προς αποκόμιση παράνομου κέρδους.

Σύμφωνα με στατιστικές μελέτες, το 88% όλων των επιθέσεων ransomware το 2017, είχαν στόχο τη βιομηχανία της υγειονομικής περίθαλψης (Tao et al., 2019). Σε μια επίθεση ransomware, ο εισβολέας απαιτεί αμοιβή λύτρων σε αντάλλαγμα για την ασφαλή επιστροφή των δεδομένων που παραβιάστηκαν. Εάν το θύμα επιλέξει να μην πληρώσει τα λύτρα, απειλείται να χάσει μόνιμα την πρόσβαση στα δεδομένα αυτά, τα οποία μπορεί αργότερα να μεταπωληθούν στο «dark web». Η βιομηχανία υγειονομικής περίθαλψης αποτελεί συνεπώς τον μεγαλύτερο στόχο επιθέσεων ransomware, και οι υπεύθυνοι των φορέων υγείας είναι συνήθως αναγκασμένοι να συμμορφωθούν με τις απαιτήσεις του επιτιθέμενου και να πληρώσουν με αντάλλαγμα τα αρχεία που παραβιάστηκαν, καθώς η αδυναμία πρόσβασης σε αρχεία ασθενών, μπορεί να θέσει σε μεγάλο κίνδυνο, ακόμη και τη ζωή τους.

Καθώς οι επιτιθέμενοι στον κυβερνοχώρο καταφεύγουν σε περισσότερο καινοτόμες προσεγγίσεις για να κερδίσουν χρήματα, η βιομηχανία της υγειονομικής περίθαλψης γίνεται ολοένα και πιο ελκυστικός στόχος, λόγω της δυνατότητας πώλησης τεράστιων συνόλων ιδιωτικών και ευαίσθητων δεδομένων, προς αποκόμιση παράνομου κέρδους. Τα νοσοκομεία συνήθως έχουν χαμηλή ασφάλεια, επομένως είναι σχετικά εύκολο για τους χάκερς να λάβουν μεγάλο αριθμό προσωπικών δεδομένων, για ιατρική απάτη. Αυτό δεν σημαίνει ότι η ασφάλεια των Μαζικών Δεδομένων είναι ασήμαντη σε άλλους κλάδους, αποδεικνύει ωστόσο, ότι κανένας κλάδος δεν εξαιρείται από παραβιάσεις δεδομένων, εφόσον αυτά συλλέγονται και υφίστανται επεξεργασία σε μεγάλες ποσότητες και ότι κάθε κλάδος πρέπει να επενδύει στην ασφάλεια των δεδομένων του, κατά προτεραιότητα (Tao et al., 2019).

Ειδικότερα, τα δεδομένα υγείας των στελεχών των ΕΔ, εκτός από ευαίσθητα και προσωπικά, όπως οποιουδήποτε άλλου ασθενούς, είναι ιδιαιτέρως πολύτιμα για τον εχθρό, καθώς μπορούν κατόπιν κατάλληλης ανάλυσης και επεξεργασίας, να του παρέχουν χρήσιμες πληροφορίες, περί της μαχητικής ικανότητας του στρατεύματος, κατά ηλικία, γεωγραφική περιοχή, ειδικότητα, περιόδους έντασης και αντιπαλότητας κ.λπ.

8.1 Ασφάλεια Μαζικών Δεδομένων Υγείας στο Νέφος

Η τεχνολογία της Υπολογιστικής Νέφους (Cloud Computing), όπως αναλύθηκε στη σχετική ενότητα, συμβάλλει στην ευελιξία, την αποτελεσματικότητα, στην εξασφάλιση υπολογιστικής ισχύος, στην εξοικονόμηση πόρων και χρόνου από επιχειρήσεις και οργανισμούς. Ωστόσο, οι απειλές για την ασφάλεια και την ιδιωτικότητα με τη χρήση αυτής της τεχνολογίας είναι αρκετές, ενώ και οι προκλήσεις που τίθενται, ως προς την εφαρμογή των σχετικών διατάξεων του ΓΚΠΔ περί απορρήτου, ακεραιότητας και εμπιστευτικότητας είναι σοβαρές.

Οι κίνδυνοι για την ασφάλεια και ιδιωτικότητα των ΜΔΥ που αποθηκεύονται στο νέφος διακρίνονται σε δύο μεγάλες κατηγορίες: 1) στην έλλειψη του απαιτούμενου ελέγχου του υπευθύνου επεξεργασίας επί του εκτελούντος την επεξεργασία των δεδομένων (αρχή λογοδοσίας) και 2) στη μη παροχή επαρκών πληροφοριών σχετικά με την ίδια την επεξεργασία, το χώρο αποθήκευσης, τα τεχνικά και οργανωτικά μέτρα (τύπος κρυπτογράφησης, είδος κρυπτογραφικών κλειδιών, συχνότητα λήψης εφεδρικών αντιγράφων), όπως απαιτεί ο ΓΚΠΔ (αρχή διαφάνειας).

Ο κίνδυνος παραβίασης της ιδιωτικότητας στο Νέφος μπορεί να αντιμετωπιστεί με κρυπτογράφηση των ΜΔΥ, πριν ακόμη αποθηκευτούν σε αυτό. Έτσι μειώνεται ο κίνδυνος παραβίασής τους, είτε από το υποκείμενο, είτε από τον πάροχο των υπηρεσιών, είτε από τρίτο φορέα. Αν ωστόσο χαθεί το κλειδί κρυπτογράφησης, τότε καθίσταται αδύνατη η πρόσβαση στα δεδομένα.

Για το λόγο αυτό είναι αφενός καλή η πρακτική τήρησης εφεδρικών αντιγράφων τους, ώστε να μειωθούν οι επιπτώσεις μιας καταστροφικής απώλειάς τους, σταθμίζοντας αφετέρου το γεγονός πως η εν λόγω πρακτική αυξάνει την έκθεση των ΜΔΥ σε παραβιάσεις.

Το ζήτημα της φορητότητας των ΜΔΥ είναι ένα άλλο σημαντικό θέμα στο περιβάλλον του Νέφους. Τα υποκείμενα των ΜΔΥ κατά κανόνα δεν γνωρίζουν σε ποιες φυσικές μηχανές (servers) είναι αποθηκευμένα τα δεδομένα τους ή σε ποιο γεωγραφικό χώρο βρίσκονται τα data centers που τα τηρούν ή σε ποια απομακρυσμένα συστήματα τηρούνται τα αντίγραφα ασφαλείας τους. Επειδή εξάλλου το σύνολο των δεδομένων του φορέα υγείας είναι δυνατό να αλλάζει συχνά, τίθεται εν αμφιβόλω η ακεραιότητα των ΜΔΥ, ενώ ταυτόχρονα είναι δυσχερής η απόδειξη ανάκτησής τους, αφού δεν υπάρχει τρόπος να επιβεβαιωθεί ανά πάσα ώρα και στιγμή η πληρότητα και ακεραιότητά τους στο Νέφος.

Για τους παραπάνω λόγους, οι υπεύθυνοι και εκτελούντες επεξεργασία των ΜΔΥ οφείλουν να επιλέγουν παρόχους, οι οποίοι προσφέρουν επαρκείς εγγυήσεις εμπιστευτικότητας, ακεραιότητας και διαθεσιμότητας των δεδομένων αυτών, μέσω υπογραφής κατάλληλης Συμφωνίας Επιπέδου Υπηρεσιών (Service Level Agreement - SLA)(Κανέλλος, 2020).

8.2 Ασφάλεια Δεδομένων στο Διαδίκτυο των Πραγμάτων Υγείας

Όπως αναλύθηκε στη σχετική ενότητα, δεδομένα υγείας είναι δυνατόν να περιλαμβάνονται στο περιεχόμενο που διακινεί και επεξεργάζεται το επίπεδο εφαρμογής, ένα από τα πέντε επίπεδα της αρχιτεκτονικής του Διαδικτύου των Πραγμάτων Υγείας (τα υπόλοιπα τέσσερα είναι: αντίληψης, δικτύου, επεξεργασίας και διαχείρισης υπηρεσιών).

Επομένως κάθε οργανισμός που χρησιμοποιεί τη συγκεκριμένη τεχνολογία, οφείλει, κατά τη φάση του σχεδιασμού, να καταγράφει με ακρίβεια τα συστήματα που χρησιμοποιούνται στο επίπεδο αντίληψης και τις κατηγορίες των δεδομένων των σχετικών με την υγεία, που συλλέγονται άμεσα ή έμμεσα. Για παράδειγμα ένας φορέας υπηρεσιών υγείας, πρέπει να καταρτίσει αρχείο επεξεργασίας δεδομένων που συλλέγει για τα υποκείμενα - ασθενείς (μετρήσεις σχετικές με τα ζωτικά τους σημεία, φαρμακευτική αγωγή, έλεγχο δραστηριότητας ευπαθών ατόμων κ.λπ.). Εξάλλου, το επίπεδο εφαρμογής πρέπει να διαθέτει πλήρως σχεδιασμένες διαδικασίες για οποιαδήποτε συναφή επεξεργασία των δεδομένων υγείας, ώστε να προστατεύονται τόσο τα δικαιώματα των υποκειμένων, όσο και η ασφάλεια των δεδομένων τους.

Καθώς οι συσκευές ΙοΗΤ με τη χρήση κατάλληλων τεχνολογιών (GPS, RFID κ.λπ.) μπορούν να επικοινωνούν μεταξύ τους χωρίς τη μεσολάβηση του χρήστη, εγείρονται θέματα ιδιωτικότητας και ασφάλειας. Προκύπτουν συνεπώς ερωτήματα όπως ποιος συλλέγει κι ελέγχει τις πληροφορίες, σε ποια νομική βάση τις επεξεργάζεται, για πόσο χρόνο τις διατηρεί και πώς λαμβάνει τη συγκατάθεση του υποκειμένου, αφού οι εν λόγω συσκευές δεν διαθέτουν μηχανισμούς λήψης συναίνεσης. Είναι εξάλλου πρακτικά δύσκολο αν όχι αδύνατο να εφαρμοστεί η αρχή της διαφάνειας και η πρακτική δυνατότητα παροχής εξηγήσεων στο μέσο άνθρωπο, για τον τρόπο συνδυαστικής επεξεργασίας, μέσω διάχυτης υπολογιστικής ισχύος (ubiquitous computing), δεδομένων που προέρχονται από πολλές διαφορετικές πηγές, όπως αισθητήρες και έξυπνες συσκευές.

Πέρα από τα στάδια ελέγχου του υπευθύνου επεξεργασίας των δεδομένων υγείας, της εθελοντικής πιστοποίησης των συσκευών ΙοΗΤ και το δικαίωμα προσφυγής του υποκειμένου στην ΑΠΔΠΧ, το τελευταίο έχει ακόμη δικαίωμα αποζημίωσης από τον υπεύθυνο επεξεργασίας, εφόσον ζημιώθηκε από την επεξεργασία των δεδομένων του από συσκευή ΙοΗΤ, κατά παράβαση των όσων προβλέπει ο ΓΚΠΔ (Κανέλλος, 2020).

8.3 Ασφάλεια Μαζικών Δεδομένων Υγείας και Τεχνητή Νοημοσύνη

Σύμφωνα με το άρθρο 22 του ΓΚΠΔ, «το υποκείμενο των δεδομένων έχει το δικαίωμα να μην υπόκειται σε απόφαση που λαμβάνεται αποκλειστικά βάσει αυτοματοποιημένης επεξεργασίας, συμπεριλαμβανομένης της κατάρτισης προφίλ, η οποία παράγει έννομα αποτελέσματα που το αφορούν ή το επηρεάζει σημαντικά με παρόμοιο τρόπο».

Σύμφωνα με τον κ. Κανέλλο (2020), συγγραφέα του «The GDPR Handbook», η παραπάνω νομοθετική απαίτηση, έρχεται αντιμέτωπη με τις ραγδαίες τεχνολογικές εξελίξεις, αφού με την ήδη διαμορφωθείσα κατάσταση και με τη βοήθεια συστημάτων που χρησιμοποιούν Τεχνητή Νοημοσύνη, Μηχανική Μάθηση και Νευρωνικά Δίκτυα, υπεύθυνοι και εκτελούντες επεξεργασία βρίσκονται αντιμέτωποι με σημαντικές προκλήσεις συμμόρφωσης με τον ΓΚΠΔ.

Εξελιγμένες τεχνικές εξόρυξης δεδομένων (data mining) επιτρέπουν την δευτερογενή παραγωγή, εκτός αρχικού σκοπού και την επεξεργασία υπό συχνά αδιαφανείς συνθήκες, πλήθους προσωπικών δεδομένων και μεταδεδομένων, από αρχικά ανώνυμα αρχεία Μαζικών Δεδομένων.

Σταδιακά η δυνατότητα εξόρυξης προσωπικών δεδομένων από μικτά αρχεία (mixed datasets) προσωπικών και μη προσωπικών δεδομένων μπορεί να χρησιμοποιηθεί για την κατασκευή αναγνωρίσιμων προσωπικών προφίλ, αν όχι στην εξατομικευμένη παρακολούθηση των υποκειμένων. Κατά συνέπεια, είναι πρόδηλο ότι, με τη χρήση υπολογιστών υψηλής ισχύος, ο διαχωρισμός μεταξύ προσωπικών και μη προσωπικών δεδομένων γίνεται ασαφής, δημιουργώντας μια ευρεία γκρίζα περιοχή ανασφάλειας δικαίου.

Η χρήση προγνωστικής ανάλυσης μέσω μηχανικής μάθησης και διενέργειας προβλέψεων, εγείρει ερωτήματα ως προς τη διασφάλιση των όρων εφαρμογής του ΓΚΠΔ. Ένα τέτοιο σύνθετο ερώτημα είναι για παράδειγμα το εξής: ποια είναι η δυνατότητα ανθρώπινης παρέμβασης και επεξήγησης αυτοματοποιημένων αποφάσεων που λαμβάνονται βάσει σύνθετων μαθηματικών αλγορίθμων, ώστε να ικανοποιείται η αρχή της «εξηγησιμότητας» και της «λογοδοσίας» του υπευθύνου επεξεργασίας αναφορικά με τη χρήση έμπειρων συστημάτων; (Κανέλλος, 2020).

Σύμφωνα με τη «Λευκή Βίβλο» της Ευρωπαϊκής Επιτροπής για την AI, η σχετική με αυτή τεχνολογία πρέπει να λειτουργεί με εμπιστοσύνη για τους ανθρώπους σε μια δίκαιη και ανταγωνιστική οικονομία, με στόχο μια ανοικτή, δημοκρατική και βιώσιμη κοινωνία. Για την ενίσχυση εφαρμογής αλλά και για την εποπτεία και διαφάνεια των συστημάτων AI είναι αναγκαία η λειτουργία κατάλληλων νομικών κανόνων, ιδιαίτερος σε συστήματα υψηλού κινδύνου όπως στο χώρο της υγείας, ενώ οι ελεγκτικές αρχές θα πρέπει να είναι σε θέση να δοκιμάζουν και να πιστοποιούν τα δεδομένα που χρησιμοποιούνται από τους αλγορίθμους (EC, White Paper on AI, 2020).

8.4 Εθνική Άμυνα – Εθνική Ασφάλεια

Οι έννοια της Εθνικής Άμυνας, η έννοια των δεδομένων υγείας των στελεχών ως ειδικής κατηγορίας δεδομένων καθώς και η έννοια της επιστημονικής έρευνας, κρίνεται σκόπιμο να αποσαφηνιστούν, ως προς τον τρόπο που αντιμετωπίζονται από τον ΓΚΠΔ, ενόψει της μελέτης του ζητήματος της «Προγνωστικής Ανάλυσης των Μαζικών Δεδομένων Υγείας των στελεχών των ΕΔ», που προφανώς εντάσσεται στο ευρύτερο πλαίσιο της επιστημονικής ιατρικής έρευνας, προς εξαγωγή χρήσιμων συμπερασμάτων και λήψη αποφάσεων.

Τονίζεται, ότι σύμφωνα με την παρ. 2 του άρ. 4 της Συνθήκης για την Ευρωπαϊκή Ένωση (ΣΕΕ) «... (η Ένωση) σέβεται τις ουσιαστικές λειτουργίες του κράτους, ιδίως δε τις λειτουργίες που αποβλέπουν στη διασφάλιση της εδαφικής ακεραιότητας, τη διατήρηση της δημόσιας τάξης και την προστασία της εθνικής ασφάλειας. Ειδικότερα, η εθνική ασφάλεια παραμένει στην ευθύνη κάθε κράτους μέλους».

Σύμφωνα με το άρθρο 1 του ν. 2292/1995 «Οργάνωση και Λειτουργία Υπουργείου Εθνικής Άμυνας, διοίκηση και έλεγχος των ενόπλων δυνάμεων και άλλες διατάξεις» (ΦΕΚ Α΄ 35/15-02-1995), στο οποίο προσδιορίζεται η αποστολή των Ενόπλων Δυνάμεων, «Η Εθνική Άμυνα περιλαμβάνει το σύνολο των λειτουργιών και δραστηριοτήτων, που αναπτύσσονται από το Κράτος, με σκοπό την προστασία της εδαφικής ακεραιότητας, της εθνικής ανεξαρτησίας και κυριαρχίας και της ασφάλειας των πολιτών εναντίον οποιασδήποτε εξωτερικής επίθεσης ή απειλής, καθώς και την υποστήριξη των εθνικών συμφερόντων».

Σύμφωνα με την αιτιολογική σκέψη 16 του ΓΚΠΔ ο Κανονισμός αυτός δεν εφαρμόζεται σε ζητήματα προστασίας θεμελιωδών δικαιωμάτων και ελευθεριών ή στην ελεύθερη κυκλοφορία δεδομένων προσωπικού χαρακτήρα που σχετίζονται με δραστηριότητες μη υπαγόμενες στο πεδίο εφαρμογής του ενωσιακού δικαίου, όπως δραστηριότητες που αφορούν την εθνική ασφάλεια.

Μπορεί λοιπόν άραγε να θεωρήσει κανείς, ότι εφόσον τα στελέχη των Ελληνικών ΕΔ, αποτελούν το έμπυχο δυναμικό που προασπίζεται την εθνική μας ασφάλεια, τα δεδομένα τους και συγκεκριμένα τα δεδομένα υγείας τους δεν εμπίπτουν στις διατάξεις του ΓΚΠΔ;

Σύμφωνα με πρόσφατη απόφαση (20/2020) που αφορούσε καταγγελία υποκειμένου των δεδομένων κατά του 401 ΓΣΝΑ, για παράνομη επεξεργασία προσωπικών δεδομένων κατά την είσοδό του στην πύλη του Νοσοκομείου, η Αρχή Προστασίας Δεδομένων Προσωπικού Χαρακτήρα (ΑΠΔΠΧ), αναφέρεται ότι κατά τις διατάξεις του άρθρου 2, που αφορούν το ουσιαστικό πεδίο εφαρμογής του, ο ΓΚΠΔ δεν εφαρμόζεται «στο πλαίσιο δραστηριότητας η οποία δεν εμπίπτει στο πεδίο εφαρμογής του δικαίου της Ένωσης» (παρ. 2 στοιχ. α΄) και «από αρμόδιες αρχές για τους σκοπούς της πρόληψης, της διερεύνησης, της ανίχνευσης ή της δίωξης ποινικών αδικημάτων ή της εκτέλεσης ποινικών κυρώσεων, συμπεριλαμβανομένης της προστασίας και πρόληψης έναντι κινδύνων που απειλούν τη δημόσια ασφάλεια» (παρ. 2 στοιχ. δ΄).

Στο άρθρο 23 ΓΚΠΔ, προβλέπεται η δυνατότητα θέσπισης μέσω νομοθετικού μέτρου περιορισμών στα δικαιώματα στις υποχρεώσεις και τα δικαιώματα των άρθρων 12 – 22 και στα άρθρα 34 και 5, όταν οι περιορισμοί αυτοί σέβονται την ουσία των θεμελιωδών δικαιωμάτων και ελευθεριών και συνιστά αναγκαίο και αναλογικό μέτρο σε μια δημοκρατική κοινωνία για τη διασφάλιση, μεταξύ άλλων, της ασφάλειας του κράτους (στοιχ. α΄), της δημόσιας ασφάλειας (στοιχ. γ΄) και της πρόληψης, της διερεύνησης, της ανίχνευσης ή της δίωξης ποινικών αδικημάτων ή της εκτέλεσης ποινικών κυρώσεων, περιλαμβανομένης της προστασίας από απειλές κατά της δημόσιας ασφάλειας και της πρόληψης αυτών (στοιχ. δ΄).

Τέλος, στο άρθρο 45 παρ. 2 ΓΚΠΔ αναφορικά με τις διαβιβάσεις δεδομένων προσωπικού χαρακτήρα προς τρίτη χώρα ή διεθνή οργανισμό, εφόσον αρκετές φορές τα στελέχη των ΕΔ καλούνται να υπηρετήσουν στο εξωτερικό (ΝΑΤΟ, ειρηνευτικές αποστολές κ.λπ., άρα είναι αναγκαία η διαβίβαση ιατρικών τους δεδομένων), ορίζεται ότι κατά την εκτίμηση της επάρκειας του επιπέδου προστασίας, η Επιτροπή λαμβάνει υπόψη, ιδίως, τα ακόλουθα στοιχεία: «το κράτος δικαίου, τον σεβασμό των ανθρωπίνων δικαιωμάτων και των το κράτος δικαίου, τον σεβασμό των ανθρωπίνων δικαιωμάτων και των θεμελιωδών ελευθεριών, τη σχετική νομοθεσία, τόσο τη γενική όσο και την τομεακή, μεταξύ άλλων όσον αφορά τη δημόσια ασφάλεια, την άμυνα θεμελιωδών ελευθεριών, τη σχετική νομοθεσία, τόσο τη γενική όσο και την τομεακή, μεταξύ άλλων όσον αφορά τη δημόσια ασφάλεια, την άμυνα, την εθνική ασφάλεια και το ποινικό δίκαιο και την πρόσβαση των δημόσιων αρχών σε δεδομένα προσωπικού χαρακτήρα, καθώς και την εφαρμογή αυτής της νομοθεσίας...».

Από τις ανωτέρω διατάξεις προκύπτει ότι η έννοια της εθνικής ασφάλειας συναντάται στο άρθρο 45 του ΓΚΠΔ που αφορά στις διαβιβάσεις βάσει απόφασης επάρκειας (και στις αιτιολογικές σκέψεις 16 του ΓΚΠΔ για το πεδίο εφαρμογής και 104 για τις διαβιβάσεις), ενώ κυρίως γίνεται αναφορά στην έννοια της δημόσιας ασφάλειας (άρθρα 2, 23 και 45 και αιτ. σκ. 19, 50, 73, 104) και ασφάλειας του κράτους (άρθρο 23).

Εξάλλου, ο ν. 4624/2019, καθορίζοντας την αρμοδιότητα της Αρχής, προβλέπει στο άρθρο 10 παρ. 5: «Η Αρχή δεν είναι αρμόδια να ελέγχει πράξεις επεξεργασίας δεδομένων προσωπικού χαρακτήρα που διενεργούνται από τις δικαστικές και εισαγγελικές αρχές στο πλαίσιο της δικαστικής λειτουργίας και των δικαστικών τους καθηκόντων, καθώς και πράξεις επεξεργασίας διαβαθμισμένων δεδομένων προσωπικού

χαρακτήρα που διενεργούνται για τις δραστηριότητες που αφορούν την εθνική ασφάλεια». Από τη συστηματική δε ερμηνεία των διατάξεων των παραγράφων 4 και 5 του άρ. 10 του νόμου 4624/2019 προκύπτει ότι «ο νομοθέτης δεν εξαιρεί εν γένει τα ζητήματα εθνικής ασφάλειας από την αρμοδιότητα της Αρχής, καθώς όπου από διεθνείς και διακρατικές συμβάσεις προβλέπεται ρητά αρμοδιότητα της Αρχής, η τελευταία διατηρεί τη σχετική αρμοδιότητα».

Σύμφωνα με την ίδια απόφαση (20/2020), η ΑΠΔΠΧ αποφάσισε ότι δεν έχει αρμοδιότητα ελέγχου των πράξεων επεξεργασίας που διενεργούνται για δραστηριότητες που αφορούν την εθνική ασφάλεια εφόσον συντρέχουν σωρευτικά οι ακόλουθες δυο προϋποθέσεις:

α) πρόκειται για πράξεις επεξεργασίας διαβαθμισμένων δεδομένων προσωπικού χαρακτήρα, και ορθότερα, πληροφοριών καταχωρημένων σε συστήματα αρχειοθέτησης που έχουν αρμοδίως διαβαθμισθεί και οι οποίες συνιστούν δεδομένα προσωπικού χαρακτήρα και β) πρόκειται για πράξεις επεξεργασίας δεδομένων προσωπικού χαρακτήρα που διενεργούνται για δραστηριότητες που αφορούν την εθνική ασφάλεια.

Επίσης, ενώ η Αρχή δεν έχει οποιουδήποτε είδους αρμοδιότητα επί των μέτρων φυσικής ασφαλείας που θα εφαρμοστούν σε ένα στρατιωτικό νοσοκομείο (καθώς το τελευταίο τα επιλέγει και εφαρμόζει), στο μέτρο που μια επεξεργασία δεν διενεργείται για δραστηριότητα σχετιζόμενη με την εθνική ασφάλεια, η Αρχή αρμοδίως ασκεί την ελεγκτική της αρμοδιότητα για να κρίνει την νομιμότητα επεξεργασίας και τα περαιτέρω σχετικά ζητήματα π.χ. εάν ο υπεύθυνος επεξεργασίας έχει λάβει τα αναγκαία τεχνικά και οργανωτικά μέτρα ασφαλείας σύμφωνα με τα άρθρα 5 παρ. 1, 24 και 32 ΓΚΠΔ.

Η «εθνική ασφάλεια» συνιστά αόριστη νομική έννοια, η οποία πρέπει, με βάση τα δεδομένα που συντρέχουν σε κάθε περίπτωση, να ερμηνεύεται, προσδιορίζεται και τεκμηριώνεται ad hoc από τον υπεύθυνο επεξεργασίας. Συνεπώς, δεν αρκεί η αόριστη και γενική επίκληση της συνδρομής της εθνικής ασφάλειας για τον αποκλεισμό της ΑΠΔΠΧ από την άσκηση των αρμοδιοτήτων της, αλλά απαιτείται στη συγκεκριμένη περίπτωση τεκμηρίωση ότι η επεξεργασία διενεργείται για δραστηριότητα που αφορά την εθνική ασφάλεια της Ελλάδας, εκτός αν είναι πρόδηλη η συνδρομή της έννοιας αυτής (π.χ. αμιγώς στρατιωτικές εγκαταστάσεις με πολεμικό υλικό, μη προσβάσιμες στο κοινό).

Η Αρχή μάλιστα με την Γνωμοδότηση υπ' άρ. 01/2020 έκρινε αναφορικά με τη διάταξη του άρθρου 10 παρ. 5 του Ν 4624/2019 ότι: «Λόγω της γενικότητας της διάταξης, κατά την οποία εξαιρούνται του ελέγχου της Αρχής οι πράξεις επεξεργασίας διαβαθμισμένων δεδομένων προσωπικού χαρακτήρα που διενεργούνται για τις δραστηριότητες που αφορούν την εθνική ασφάλεια, ανακύπτουν δυσχέρειες στον χαρακτηρισμό των πράξεων αυτών. Ενόψει αυτού, είναι σκόπιμο να προστεθεί στην παρ. 5 η εξής διάταξη: «Οι αρχές που επεξεργάζονται διαβαθμισμένα δεδομένα προσωπικού χαρακτήρα στο πλαίσιο δραστηριοτήτων εθνικής ασφάλειας, ενημερώνουν και συνεργάζονται με την Αρχή για την τήρηση της νομοθεσίας περί προστασίας προσωπικών δεδομένων και ιδίως τη λήψη και την τήρηση των αναγκαίων γενικών μέτρων ασφαλείας».

Έχει προταθεί εξάλλου να συμπεριληφθεί σε μελλοντική τροποποίηση του νόμου, ρήτρα συνεργασίας αρχών που ασχολούνται με δραστηριότητες εθνικής ασφάλειας (όπως η Εθνική Υπηρεσία Πληροφοριών - ΕΥΠ), για τη διασφάλιση ενός κατάλληλου επιπέδου προστασίας τέτοιων δεδομένων. Σε αρκετές περιπτώσεις ενδέχεται να δημιουργηθούν προβλήματα και συγκρούσεις αρμοδιοτήτων κατά την εφαρμογή του νόμου, αφού τα όρια μεταξύ εθνικής ασφάλειας και δημόσιας τάξης γίνονται διεθνώς ολοένα και πιο δυσδιάκριτα. Για παράδειγμα υπό το μανδύα του πρόσφυγα ή του λαθρομετανάστη, ο έλεγχος και η σύλληψη των οποίων ανήκει στο Υπουργείο Δημόσιας Τάξης, καλύπτονται συχνά μυστικοί πράκτορες που ενεργούν κατασκοπεία σε βάρος της χώρας μας υπέρ άλλων χωρών, επομένως ο εντοπισμός και σύλληψή τους αφορά την ΕΥΠ, ως ζήτημα εθνικής ασφάλειας (Κανέλλος 2020),

8.5 Δεδομένα Υγείας και Επιστημονική Έρευνα

Ο ΓΚΠΔ στην παράγραφο 1β του άρθρου 23 δίδει ευρύ περιθώριο δράσης στον εθνικό νομοθέτη αναφορικά με τον περιορισμό των δικαιωμάτων του υποκειμένου των δεδομένων, ως αναγκαίο και αναλογικό μέτρο για τη διασφάλιση της εθνικής άμυνας της χώρας.

Πέρα από την ιδιότητά τους ως υπηρετούντα στις ΕΔ, τα στελέχη-ασθενείς δεν χάνουν την ιδιότητά τους ως υποκείμενα κατά την έννοια του ΓΚΠΔ και του Ν 4624/2019, στα οποία ανήκουν και τα οποία αφορούν, ως προσωπικά και ιδιαίτερης κατηγορίας (άρ. 9, παρ. 1 του ΓΚΠΔ), προσωπικά δεδομένα για την υγεία τους. Σύμφωνα λοιπόν με το άρ. 9, παρ.1 του ΓΚΠΔ, απαγορεύεται η επεξεργασία δεδομένων προσωπικού χαρακτήρα, που μεταξύ άλλων αφορούν την υγεία. Η παράγραφος αυτή όμως δεν εφαρμόζεται σε περιπτώσεις που το υποκείμενο των δεδομένων (εν προκειμένω το στέλεχος των ΕΔ) έχει παράσχει ρητή συγκατάθεση (άρ 9, παρ. 2) για την επεξεργασία αυτών των δεδομένων για έναν ή περισσότερους συγκεκριμένους σκοπούς.

Η συγκατάθεση θα πρέπει να παρέχεται με σαφή θετική ενέργεια η οποία να συνιστά ελεύθερη, συγκεκριμένη, ρητή και εν πλήρει επιγνώσει ένδειξη της συμφωνίας του υποκειμένου των δεδομένων υπέρ της επεξεργασίας των δεδομένων που το αφορούν, για παράδειγμα με γραπτή δήλωση, μεταξύ άλλων με ηλεκτρονικά μέσα, ή με προφορική δήλωση. Όταν η επεξεργασία βασίζεται στη συναίνεση του υποκειμένου των δεδομένων, ο υπεύθυνος επεξεργασίας θα πρέπει να είναι σε θέση να αποδείξει ότι το υποκείμενο των δεδομένων συγκατατέθηκε στη πράξη επεξεργασίας.

Όταν το υποκείμενο των δεδομένων παρέσχε τη συναίνεσή του ή η επεξεργασία βασίζεται στο δίκαιο της Ένωσης ή κράτους μέλους που συνιστά αναγκαίο και αναλογικό μέτρο σε μια δημοκρατική κοινωνία για τη διασφάλιση, ειδικότερα, σημαντικών σκοπών στο πλαίσιο γενικού δημόσιου συμφέροντος, θα πρέπει να επιτρέπεται στον υπεύθυνο επεξεργασίας να προβαίνει στην περαιτέρω επεξεργασία των δεδομένων προσωπικού χαρακτήρα, ανεξάρτητα από τη συμβατότητα των σκοπών. Σε κάθε περίπτωση, θα πρέπει να διασφαλίζεται η εφαρμογή των αρχών που καθορίζονται ΓΚΠΔ και, ιδίως, η ενημέρωση του υποκειμένου των δεδομένων σχετικά με τους άλλους αυτούς σκοπούς και σχετικά με τα δικαιώματά του, συμπεριλαμβανομένου του δικαιώματος προβολής αντιρρήσεων.

Γενικά, επιτρέπεται η παρέκκλιση από την απαγόρευση της παρ. 1 του άρ. 9 για υγειονομικούς σκοπούς, συμπεριλαμβανομένων της δημόσιας υγείας και της διαχείρισης υπηρεσιών υγειονομικής περίθαλψης, προκειμένου ειδικότερα να διασφαλίζονται η ποιότητα και η αποδοτικότητα ως προς το κόστος των διαδικασιών που χρησιμοποιούνται για τον διακανονισμό αξιώσεων για παροχές και υπηρεσίες στο σύστημα υγειονομικής ασφάλισης, ή για σκοπούς αρχειοθέτησης προς το δημόσιο συμφέρον, σκοπούς επιστημονικής ή ιστορικής έρευνας ή στατιστικούς σκοπούς.

Η ΑΠΔΠΧ κρίνει παγίως (σχετικές αποφάσεις 46/2004, 47/2004, 16/2005, 32/2006) ότι η διενέργεια επιστημονικής έρευνας με δεδομένα υγείας ασθενών ενός νοσηλευτικού ιδρύματος, συνιστά νόμιμο σκοπό επεξεργασίας, μεταξύ άλλων, και λόγω του ότι σύμφωνα με το άρθρο 16 παρ. 1 του Συντάγματος η ανάπτυξη και η προαγωγή της έρευνας, αποτελεί υποχρέωση του Κράτους. Βέβαια η πρόσβαση στα αρχεία ενός νοσηλευτικού ιδρύματος συνιστά επεξεργασία ευαίσθητων προσωπικών δεδομένων (ειδικών κατηγοριών, κατά το άρ. 9 του ΓΚΠΔ). Προκειμένου η επεξεργασία να είναι νόμιμη θα πρέπει το νοσηλευτικό ίδρυμα να ζητήσει την άδεια της Αρχής για να επιτραπεί η πρόσβαση του ερευνητή στα αρχεία του. Η ΑΠΔΠΧ με τη σχετική άδεια καθορίζει τους όρους και τις προϋποθέσεις για την προστασία των δικαιωμάτων των ασθενών, όπως ανωνυμοποίηση δεδομένων, καταγραφή των δεδομένων εντός του χώρου του νοσηλευτικού ιδρύματος, κ.λπ. (Γενικό Νοσοκομείο Καβάλας).

Εξάλλου, πρέπει να διευκρινιστεί ότι αν στο πλαίσιο επιστημονικής έρευνας λαμβάνει χώρα η δημοσίευση στοιχείων του ιατρικού φακέλου ασθενών, κατά τρόπο που να μην είναι αμέσως ή εμμέσως προσδιορίσιμη η ταυτότητά τους, αυτό δεν συνιστά επεξεργασία προσωπικών δεδομένων, που εμπίπτει στο ρυθμιστικό πεδίο του νόμου, για την προστασία του ατόμου από την επεξεργασία δεδομένων προσωπικού χαρακτήρα

Σύμφωνα με την αιτιολογική σκέψη 35 του ΓΚΠΔ, τα δεδομένα προσωπικού χαρακτήρα σχετικά με την υγεία θα πρέπει να περιλαμβάνουν όλα τα δεδομένα που αφορούν την κατάσταση της υγείας του υποκειμένου των δεδομένων και τα οποία αποκαλύπτουν πληροφορίες για την παρελθούσα, τρέχουσα ή μελλοντική κατάσταση της σωματικής ή ψυχικής υγείας του υποκειμένου των δεδομένων (π.χ. αριθμός, σύμβολο, χαρακτηριστικό ταυτότητας για πλήρη ταυτοποίηση του φυσικού προσώπου για σκοπούς υγείας, πληροφορίες από εξετάσεις ή αναλύσεις, μέρος ή ουσία του σώματος, μεταξύ άλλων από γενετικά δεδομένα και βιολογικά δείγματα και κάθε πληροφορία, παραδείγματος χάριν, σχετικά με ασθένεια, αναπηρία, κίνδυνο ασθένειας, ιατρικό ιστορικό, κλινική θεραπεία ή τη φυσιολογική ή βιοϊατρική κατάσταση του υποκειμένου των δεδομένων).

Σύμφωνα με την αιτιολογική σκέψη 53 του ΓΚΠΔ, η επεξεργασία ειδικών κατηγοριών δεδομένων προσωπικού χαρακτήρα, που χρήζουν υψηλότερης προστασίας, θα πρέπει να γίνεται μόνο για σκοπούς που σχετίζονται με την υγεία, εφόσον αυτό είναι απαραίτητο για την επίτευξη των εν λόγω στόχων, προς όφελος των φυσικών προσώπων και της κοινωνίας στο σύνολό της, ιδίως στο πλαίσιο της διαχείρισης των υπηρεσιών και συστημάτων υγειονομικής περίθαλψης και κοινωνικής μέριμνας.

Η παράγραφος 1 του άρ. 9 δεν εφαρμόζεται κι όταν η επεξεργασία είναι απαραίτητη για την εκτέλεση των υποχρεώσεων και την άσκηση συγκεκριμένων δικαιωμάτων του υπευθύνου επεξεργασίας (εν προκειμένω της Διεύθυνση Υγειονομικού του στελέχους) ή του υποκειμένου των δεδομένων στον τομέα του εργατικού δικαίου και του δικαίου κοινωνικής ασφάλισης και κοινωνικής προστασίας. Τα στελέχη των ΕΔ είναι εργαζόμενοι του δημόσιου τομέα, η κοινωνική τους δε ασφάλιση και κοινωνική τους προστασία, έχει αναληφθεί από την ημέρα της κατάταξής τους, στις αντίστοιχες Διευθύνσεις Υγειονομικού των Κλάδων τους.

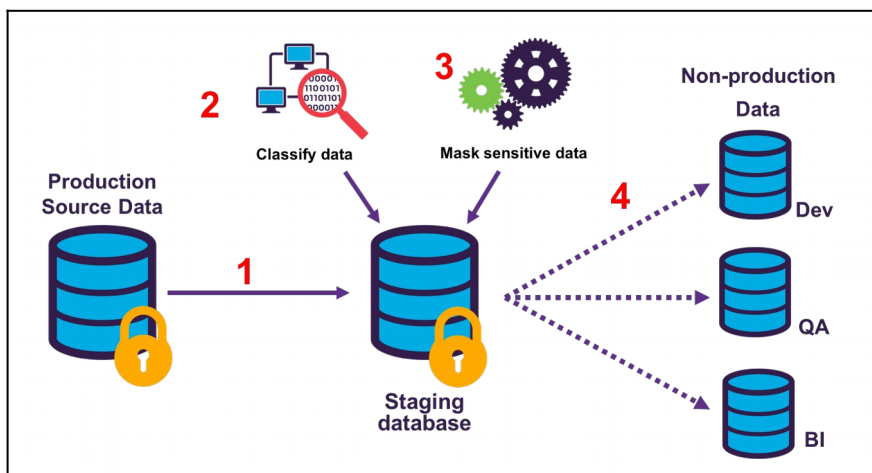
Εξάλλου η παράγραφος 1 δεν εφαρμόζεται όταν (άρ 9, παρ. 2ζ) η επεξεργασία είναι απαραίτητη για λόγους ουσιαστικού δημόσιου συμφέροντος, όταν (άρ 9, παρ. 2η) η επεξεργασία είναι απαραίτητη για σκοπούς προληπτικής ή επαγγελματικής ιατρικής, εκτίμησης της ικανότητας προς εργασία του εργαζομένου, ιατρικής διάγνωσης, παροχής υγειονομικής ή κοινωνικής περίθαλψης ή θεραπείας ή διαχείρισης υγειονομικών και κοινωνικών συστημάτων και υπηρεσιών, όταν (άρ 9, παρ 2ι) η επεξεργασία είναι απαραίτητη για σκοπούς επιστημονικής ή ιστορικής έρευνας ή για στατιστικούς σκοπούς σύμφωνα με το άρθρο 89 παράγραφος 1 (του ΓΚΠΔ). Στην τελευταία περίπτωση, προφανώς, εμπίπτει και η Προγνωστική Ανάλυση των δεδομένων υγείας των στελεχών, με σκοπό την έγκαιρη διάγνωση και θεραπεία ασθενειών τους.

Επίσης σύμφωνα με την παρ. 3 του άρ. 9, τα δεδομένα προσωπικού χαρακτήρα που αναφέρονται στην παράγραφο 1 του άρ. 9, μπορεί να τύχουν επεξεργασίας, όταν τα δεδομένα αυτά υποβάλλονται σε επεξεργασία από ή υπό την ευθύνη επαγγελματία που υπόκειται στην υποχρέωση τήρησης του επαγγελματικού απορρήτου ή από άλλο πρόσωπο το οποίο υπέχει επίσης υποχρέωση τήρησης του απορρήτου, όπως είναι το ιατρικό και νοσηλευτικό προσωπικό που περιθάλπει τα στελέχη των ΕΔ.

8.6 Ανωνυμοποίηση και Ψευδωνυμοποίηση Προσωπικών Δεδομένων

Τα προσωπικά δεδομένα που έχουν καταστεί ανώνυμα ή αλλιώς που έχουν υποστεί ανωνυμοποίηση, εξαιρούνται από όσα επιτάσσει η εφαρμογή του ΓΚΠΔ, καθώς δεν καθιστούν πλέον αναγνωρίσιμο το υποκείμενο που αφορούν, και ως εκ τούτου δεν υφίσταται κίνδυνος, για τα δικαιώματα και τις ελευθερίες του.

Τα ορθώς ανωνυμοποιημένα δεδομένα μπορούν επομένως θεωρητικά να υποστούν επεξεργασία και δημοσιοποίηση, χωρίς παραβίαση του ΓΚΠΔ. Για παράδειγμα, οι φορείς υγείας μπορεί να θέλουν να δημοσιεύσουν στατιστικά στοιχεία σχετικά με τα αποτελέσματα ερευνών με ανώνυμο τρόπο, χωρίς να απαιτείται να ζητήσουν την άδεια από τα υποκείμενα των στοιχείων αυτών.



Σχ. 20 Σχηματική απεικόνιση της λύσης απόκρυψης ευαίσθητων δεδομένων Imperva Camouflage (Πηγή: srcloudsolutions.co.uk)

Εξάλλου, τεχνικές απόκρυψης (camouflage) των ευαίσθητων δεδομένων είναι πλέον διαθέσιμες, με τη χρήση κατάλληλων τεχνικών κάλυψής τους. Για παράδειγμα η «Imperva Camouflage» είναι μια λύση κάλυψης δεδομένων που μειώνει το προφίλ κινδύνου του φορέα υγείας, αντικαθιστώντας ευαίσθητα δεδομένα με ρεαλιστικά φανταστικά (realistic fictional) δεδομένα. Τα «καμουφλαρισμένα» δεδομένα διατηρούν την ακεραιότητα των στοιχείων αναφοράς και είναι στατιστικά ακριβή, προστατεύοντας τα δεδομένα από κλοπή, διασφαλίζοντας παράλληλα τη συμμόρφωση με τους κανονισμούς και τις διεθνείς πολιτικές που υπαγορεύουν το απόρρητο και τη μεταφορά δεδομένων. Έτσι η εν λόγω λύση παρέχει προστασία στην απόκρυψη μεγάλου όγκου δεδομένων, διευκόλυνση της μεταφοράς δεδομένων εκτός του φορέα, αποτροπή απώλειας ευαίσθητων δεδομένων από μη παραγωγικά συστήματα και συμμόρφωση με τους διεθνείς κανονισμούς περί απορρήτου δεδομένων.

Ωστόσο στο πλαίσιο επεξεργασίας δεδομένων για τις ανάγκες διαχείρισης του ανθρώπινου δυναμικού ενός οργανισμού, η χρησιμότητα ανωνυμοποιημένων δεδομένων είναι μάλλον αμφίβολη. Εκτός της στατιστικής ανάλυσης, οι περισσότερες δραστηριότητες διαχείρισης ανθρώπινου δυναμικού απαιτούν την δυνατότητα αναγνώρισης υπαλλήλων/εργαζομένων μέσω των δεδομένων τους, και εφόσον αυτό δεν είναι πλέον απαραίτητο, τίθεται το ερώτημα κατά πόσο η διατήρηση των σχετικών πληροφοριών είναι πλέον απολύτως αναγκαία, οπότε και συνίσταται η καταστροφή τους.

Σύμφωνα με το άρθρο 4 παρ. 5 του ΓΚΠΔ, «ψευδωνυμοποίηση» ορίζεται ως η επεξεργασία δεδομένων προσωπικού χαρακτήρα κατά τρόπο ώστε τα δεδομένα να μην μπορούν πλέον να αποδοθούν σε συγκεκριμένο υποκείμενο των δεδομένων, χωρίς τη χρήση συμπληρωματικών πληροφοριών, εφόσον οι εν λόγω συμπληρωματικές πληροφορίες διατηρούνται χωριστά και υπόκεινται σε τεχνικά και οργανωτικά μέτρα, προκειμένου να διασφαλιστεί ότι δεν μπορούν να αποδοθούν σε ταυτοποιημένο ή ταυτοποιήσιμο φυσικό πρόσωπο.

Πρόκειται λοιπόν για ένα μέτρο ασφαλείας, που περιλαμβάνει μια διαδικασία επαναπροσδιορισμού, συνήθως μέσω της χρήσης κωδικών, όπου σε ένα σύνολο πληροφοριών ανατίθεται ένα μοναδικό (πρωτεύον) κλειδί ή αναγνωριστικό. Η χρήση ψευδωνυμοποιημένων δεδομένων μπορεί να βοηθήσει στη διαχείριση του κινδύνου προστασίας τους.

Στην αιτιολογική σκέψη 26 του ΓΚΠΔ αναφέρεται ρητά ότι ο Κανονισμός δεν αφορά την επεξεργασία ανωνύμων πληροφοριών, ούτε μεταξύ άλλων για στατιστικούς ή ερευνητικούς σκοπούς. Πιο συγκεκριμένα,, στην αιτιολογική σκέψη αναφέρεται ότι τα δεδομένα προσωπικού χαρακτήρα που έχουν υποστεί ψευδωνυμοποίηση, που θα μπορούσε να αποδοθεί σε φυσικό πρόσωπο με τη χρήση συμπληρωματικών πληροφοριών, θα πρέπει να θεωρούνται πληροφορίες σχετικά με ταυτοποιήσιμο φυσικό πρόσωπο. Για να κριθεί κατά πόσον ένα φυσικό πρόσωπο είναι ταυτοποιήσιμο, θα πρέπει να λαμβάνονται υπόψη όλα τα μέσα τα οποία είναι ευλόγως πιθανό ότι θα χρησιμοποιηθούν, δηλαδή και όλοι οι αντικειμενικοί παράγοντες, όπως απαιτούμενα έξοδα και χρόνος για την ταυτοποίηση, λαμβανομένων υπόψη της διαθέσιμης τεχνολογίας.

Οι αρχές της προστασίας δεδομένων δεν θα πρέπει συνεπώς να εφαρμόζονται σε ανώνυμες πληροφορίες, δηλαδή πληροφορίες που δεν σχετίζονται προς ταυτοποιημένο ή ταυτοποιήσιμο φυσικό πρόσωπο ή σε δεδομένα προσωπικού χαρακτήρα που έχουν καταστεί ανώνυμα, κατά τρόπο ώστε η ταυτότητα του υποκειμένου των δεδομένων να μην μπορεί ή να μην μπορεί πλέον να εξακριβωθεί.

Υπάρχουν διάφοροι τρόποι για ανωνυμοποίηση προσωπικών δεδομένων. Οι δύο πιο συχνά χρησιμοποιούμενοι είναι η τυχαιοποίηση και η γενίκευση. Η τυχαιοποίηση συνεπάγεται την αλλοίωση των δεδομένων, με σκοπό την κατάργηση της σύνδεσης μεταξύ του υποκειμένου και των δεδομένων του, χωρίς απώλεια της αξίας των τελευταίων. Η γενίκευση συνεπάγεται τροποποίηση του επιπέδου λεπτομέρειας στα δεδομένα. Έτσι για παράδειγμα, αν το υποκείμενο μας είναι ένας ασθενής βαθμού Ταγματάρχη, ηλικίας 40 ετών, με τόπο γέννησης τη Θεσσαλονίκη, η γενίκευση της πληροφορίας της ηλικίας του επιτυγχάνεται μέσω της χρήσης ηλικιακών τιμών διακύμανσης, π.χ. 35-45 (αντί 40), η γενίκευση της πληροφορίας του βαθμού του, επιτυγχάνεται με τη χρήση βαθμίδων ιεραρχίας π.χ. ανώτερος Αξιωματικός (όπου συμπεριλαμβάνονται οι βαθμοί του Ταγματάρχη, του Αντισυνταγματάρχη και του Συνταγματάρχη), ενώ η γενίκευση της πληροφορίας του τόπου γέννησής του, επιτυγχάνεται μέσω της χρήσης ευρύτερου γεωγραφικού διαμερίσματος π.χ. Μακεδονία (αντί Θεσσαλονίκης).

Quasi identifiers				Sensitive attribute
ID	Age	Country	Political views	
1	35	Greenland	Liberal	
2	35	Canada	Conservative	
3	38	Belize	Liberal	
4	40	Belize	Liberal	
5	37	Canada	Conservative	
6	37	Canada	Conservative	

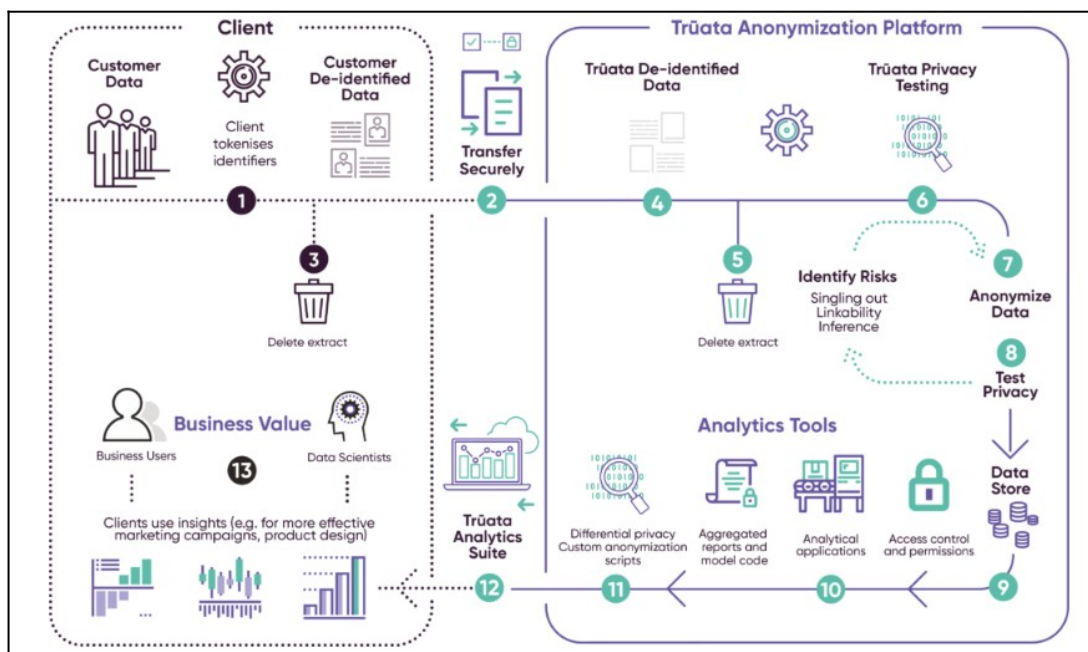
(a) Original data about the users

Quasi Identifiers			Sensitive Attribute
ID	Age	Country	Political views
1	35-40	America	Liberal
2	35-40	America	Conservative
3	35-40	America	Liberal
4	35-40	America	Liberal
5	35-40	America	Conservative
6	35-40	America	Conservative

(b) 2-anonymous data about the users

Σχ. 21 Παράδειγμα γενίκευσης 2 ψευδο-αναγνωριστικών (2-anonymity) της Ηλικίας και της Χώρας, για να μην είναι δυνατή η ταυτοποίηση των υποκειμένων και ο προσδιορισμός των πολιτικών τους πεποιθήσεων (Πηγή: Majeed and Lee, 2020)

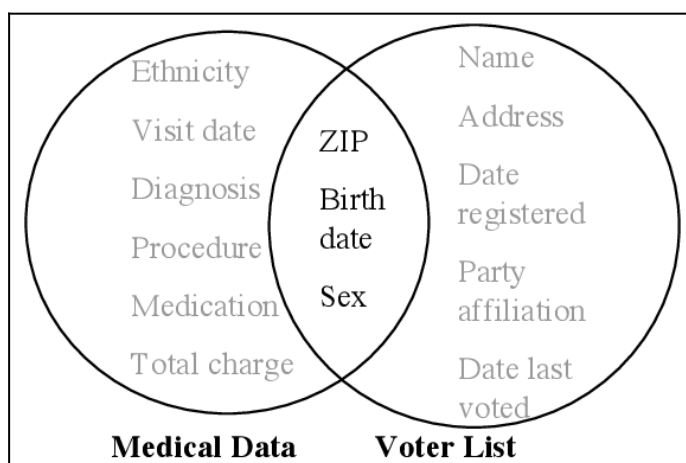
Η πράξη της ανωνυμοποίησης των προσωπικών δεδομένων αποτελεί από μόνη της μια πράξη επεξεργασίας κατά τον ΓΚΠΔ. Απαιτείται επομένως κατάλληλη νομική βάση για να διενεργηθεί και η επεξεργασία αυτή δεν πρέπει να γίνεται με τρόπο ασυμβίβαστο, με τον σκοπό για τον οποίο συλλέχθηκαν τα δεδομένα αρχικά.



Σχ. 22 Σχηματική απεικόνιση της πλατφόρμας ανωνυμοποίησης Truata (Πηγή: truata.com)

Υπάρχουν σύγχρονες πλατφόρμες ανωνυμοποίησης δεδομένων, για λογαριασμό πελατών που διαθέτουν δεδομένα, όπως η πλατφόρμα «Tūata». Η πλατφόρμα αυτή αναλύει τα ανώνυμα δεδομένα, χρησιμοποιώντας εργαλεία που επιλέγονται με βάση τις ανάγκες ανάλυσης δεδομένων, κάθε μεμονωμένου πελάτη. Οι πελάτες λαμβάνουν αναλυτικά στοιχεία, αλγόριθμους και συγκεντρωτικές αναφορές από την Tūata, για χρήση στα προϊόντα και τις λύσεις τους. Οι πελάτες εφαρμόζουν αυτά τα αναλυτικά στοιχεία, αλγόριθμους και αναφορές, για να ξεκλειδώσουν ισχυρές και πολύτιμες πληροφορίες από τα δεδομένα τους.

Ο κίνδυνος της ταυτοποίησης είναι κεντρικό ζήτημα για την εξασφάλιση της πραγματικής ανωνυμίας των προσωπικών δεδομένων. Ένα υποκείμενο δεν είναι απαραίτητο να κατονομαστεί για να ταυτοποιηθεί, αφού υπάρχουν πολλοί τρόποι με τους οποίους μπορεί να αναγνωριστεί και να ξεχωρίσει μέσα από μια ομάδα άλλων υποκειμένων. Επομένως, προκειμένου να διασφαλισθεί ότι η διαδικασία ταυτοποίησης είναι μη αναστρέψιμη, ο φορέας που επεξεργάζεται τα δεδομένα, πρέπει να εξετάσει τους διάφορους πιθανούς τρόπους με τους οποίους ένας κακόβουλος τρίτος (εισβολέας) χρησιμοποιεί την τεχνολογία και συνδυάζει αυτές τις πληροφορίες, από διαφορετικές ενδεχομένως πηγές, για να ταυτοποιήσει ένα άτομο. Για το λόγο αυτό χρησιμοποιούνται εξειδικευμένοι οργανισμοί που διασφαλίζουν ότι η διαδικασία ανωνυμοποίησης είναι πραγματικά μη αναστρέψιμη και με κανέναν τρόπο δεν οδηγεί σε ταυτοποίηση του υποκειμένου.



Σχ. 23 Παράδειγμα επίθεσης συσχετίσεων σε δύο πίνακες με κοινά πεδία ΤΚ, Ημερομηνία Γέννησης και Φύλο, για ταυτοποίηση των «ανωνυμοποιημένων» υποκειμένων (Πηγή: Sweeney, 2002)

Σε ένα σύνολο δεδομένων υπάρχουν αναγνωριστικά πεδία (identifiers), τα οποία μπορούν με μοναδικό τρόπο να ταυτοποιήσουν ένα υποκείμενο, όπως ο αριθμός ταυτότητάς του, ο αριθμός φορολογικού του μητρώου (ΑΦΜ), ο αριθμός διαβατηρίου του κ.λπ. Εκτός αυτών όμως των αναγνωριστικών, υπάρχουν κι άλλα πεδία σε ένα σύνολο δεδομένων, τα ψευδο-αναγνωριστικά (quasi-identifiers), που μπορούν να χρησιμοποιηθούν για να βοηθήσουν στην ταυτοποίηση, σε συσχέτιση με άλλες πληροφορίες. Τα ψευδο-αναγνωριστικά είναι σύνολα χαρακτηριστικών που, σε συνδυασμό, μπορούν να συνδεθούν με εξωτερικές πληροφορίες για να επαναπροσδιορίσουν (ή να μειώσουν την αβεβαιότητα) όλων ή ορισμένων από τα υποκείμενα, στα οποία αναφέρονται οι πληροφορίες (De Capitani di Vimercati and Foresti, 2011).

Μια μελέτη στις ΗΠΑ ανέφερε ότι με μόνο τρία δημόσια εκτεθειμένα πεδία πληροφοριών για ένα υποκείμενο (των ψευδο-αναγνωριστικών «ημερομηνία γέννησης», «ταχυδρομικός κώδικας» και «φύλο»), ερευνητές κατάφεραν να προσδιορίσουν με μοναδικό τρόπο, το 87% του πληθυσμού των ΗΠΑ (216 εκατομμύρια από 248 εκατομμύρια), όπου συνδυασμοί λίγων χαρακτηριστικών συσχετίστηκαν σε πληθυσμούς, ώστε να είναι εφικτή, με μοναδικό ή με σχεδόν μοναδικό τρόπο, η ταυτοποίηση μεμονωμένων ατόμων. Σαφώς, τα δεδομένα που κυκλοφόρησαν και περιείχαν τέτοιες πληροφορίες για αυτά τα άτομα, δεν μπορεί και δεν πρέπει να θεωρούνται ανώνυμα (Sweeney, 2002).

8.7 Συγκατάθεση και Εναντίωση των Υποκειμένων στην Προγνωστική Ανάλυση

Σύμφωνα με το άρθρο 6 του ΓΚΠΔ, η επεξεργασία είναι σύννομη εφόσον, μεταξύ άλλων, το υποκείμενο των δεδομένων έχει συναινέσει στην επεξεργασία των δεδομένων προσωπικού χαρακτήρα του για έναν ή περισσότερους συγκεκριμένους σκοπούς (παρ. 2α), η επεξεργασία είναι απαραίτητη για τη συμμόρφωση με έννομη υποχρέωση του υπευθύνου επεξεργασίας (παρ. 2γ), η επεξεργασία είναι απαραίτητη για την εκπλήρωση καθήκοντος που εκτελείται προς το δημόσιο συμφέρον ή κατά την άσκηση δημόσιας εξουσίας που έχει ανατεθεί στον υπεύθυνο επεξεργασίας (παρ. 2ε) κ.λπ.

Εν προκειμένω, το Υπουργείο Εθνικής Άμυνας με τους αντίστοιχους φορείς παροχής υγειονομικών υπηρεσιών (Διευθύνσεις Υγειονομικού των Κλάδων) προς τα στελέχη του, είναι ο υπεύθυνος επεξεργασίας κατά τον ΓΚΠΔ. Υποκείμενα των δεδομένων είναι τα στελέχη των ΕΔ και τα δεδομένα της υγείας τους, εκτός από προσωπικά, συγκαταλέγονται στα δεδομένα «ειδικών κατηγοριών» (άρ. 9).

Εκτός του ΓΚΠΔ, το Σύνταγμά μας ορίζει στο άρθρο 9Α ότι: «Καθένας έχει δικαίωμα προστασίας από τη συλλογή, επεξεργασία και χρήση, ιδίως με ηλεκτρονικά μέσα, των προσωπικών του δεδομένων, όπως νόμος ορίζει. Η προστασία των προσωπικών δεδομένων διασφαλίζεται από ανεξάρτητη αρχή, που συγκροτείται και λειτουργεί, όπως νόμος ορίζει».

Ειδικότερα, για τα δεδομένα ειδικών κατηγοριών (άρ. 9), ο ΓΚΠΔ ορίζει ότι «απαγορεύεται η επεξεργασία δεδομένων προσωπικού χαρακτήρα που αποκαλύπτουν τη φυλετική ή εθνοτική καταγωγή, τα πολιτικά φρονήματα, τις θρησκευτικές ή φιλοσοφικές πεποιθήσεις ή τη συμμετοχή σε συνδικαλιστική οργάνωση, καθώς και η επεξεργασία γενετικών δεδομένων, βιομετρικών δεδομένων με σκοπό την αδιαμφισβήτητη ταυτοποίηση προσώπου, δεδομένων που αφορούν την υγεία ή δεδομένων που αφορούν τη σεξουαλική ζωή φυσικού προσώπου ή τον γενετήσιο προανατολισμό».

Στην παρ. 2 του άρ 9, ορίζεται ότι η παρ. 1 του άρ. 9 δεν εφαρμόζεται σε ορισμένες, ειδικά αναφερόμενες, περιπτώσεις. Έτσι δεν εφαρμόζεται (άρ 9, παρ. 2α) όταν «το υποκείμενο των δεδομένων έχει παράσχει ρητή συγκατάθεση για την επεξεργασία αυτών των δεδομένων προσωπικού χαρακτήρα για έναν ή περισσότερους συγκεκριμένους σκοπούς, εκτός εάν το δίκαιο της Ένωσης ή κράτους μέλους προβλέπει ότι η απαγόρευση που αναφέρεται στην παράγραφο 1, δεν μπορεί να αρθεί από το υποκείμενο των δεδομένων».

Επίσης, η παρ. 1 του άρ. 9 δεν εφαρμόζεται (άρ 9, παρ. 2β) όταν «η επεξεργασία είναι απαραίτητη για την εκτέλεση των υποχρεώσεων και την άσκηση συγκεκριμένων δικαιωμάτων του υπευθύνου επεξεργασίας ή του υποκειμένου των δεδομένων στον τομέα του εργατικού δικαίου και του δικαίου κοινωνικής ασφάλισης και κοινωνικής προστασίας, εφόσον επιτρέπεται από το δίκαιο της Ένωσης ή κράτους μέλους ή από συλλογική συμφωνία σύμφωνα με το εθνικό δίκαιο παρέχοντας κατάλληλες εγγυήσεις για τα θεμελιώδη δικαιώματα και τα συμφέροντα του υποκειμένου των δεδομένων».

Ακόμη, η παρ. 1 του άρ. 9 δεν εφαρμόζεται (άρ 9, παρ. 2η) όταν «η επεξεργασία είναι απαραίτητη για σκοπούς προληπτικής ή επαγγελματικής ιατρικής, εκτίμησης της ικανότητας προς εργασία του εργαζομένου, ιατρικής διάγνωσης, παροχής υγειονομικής ή κοινωνικής περίθαλψης ή θεραπείας ή διαχείρισης υγειονομικών και κοινωνικών συστημάτων και υπηρεσιών βάσει του ενωσιακού δικαίου ή του δικαίου κράτους μέλους ή δυνάμει σύμβασης με επαγγελματία του τομέα της υγείας...».

Αφετέρου, σύμφωνα με το άρθρο 5 παρ. 1 του Συντάγματος, «Καθένας έχει δικαίωμα να αναπτύσσει ελεύθερα την προσωπικότητά του και να συμμετέχει στην κοινωνική, οικονομική και πολιτική ζωή της Χώρας, εφόσον δεν προσβάλλει τα δικαιώματα των άλλων και δεν παραβιάζει το Σύνταγμα ή τα χρηστά ήθη», ενώ σύμφωνα με το άρθρο 5Α παρ. 1 «Καθένας έχει δικαίωμα στην πληροφόρηση, όπως νόμος ορίζει. Περιορισμοί στο δικαίωμα αυτό είναι δυνατόν να επιβληθούν με νόμο μόνο εφόσον είναι απολύτως αναγκαίοι και δικαιολογούνται για λόγους εθνικής ασφάλειας, καταπολέμησης του εγκλήματος ή προστασίας δικαιωμάτων και συμφερόντων τρίτων». Η προσωπικότητα του ανθρώπου συντίθεται από τη σωματική, ψυχική, πνευματική, ηθική και κοινωνική υπόστασή του από την οποία απορρέουν επιμέρους δικαιώματα για την προστασία αγαθών, που συνυφαίνονται με την προσωπικότητά του, όπως είναι η ζωή, η σωματική ακεραιότητα, η υγεία και η τιμή (Παπαντωνίου, 1983).

Αν και εκτιμάται ότι η Προγνωστική Ανάλυση είναι προς όφελος του υποκειμένου των δεδομένων υγείας και των ΕΔ στις οποίες υπηρετεί, θα μπορούσε να ισχυριστεί κάποιος, ότι στο πλαίσιο της ελεύθερης ανάπτυξης της προσωπικότητας του υποκειμένου, το τελευταίο δεν επιθυμεί να υποβληθούν τα δεδομένα του σε Προγνωστική Ανάλυση, διότι π.χ. δεν θέλει να γνωρίζει, εκ των προτέρων, από τι θα νοσήσει στο μέλλον, έστω κι αν αυτό μπορεί τελικά να αποβεί σωτήριο για την υγεία του. Εκτός κι αν δεχτούμε ότι είναι αφενός δυνατή η Προγνωστική Ανάλυση, έτσι ώστε ο υπεύθυνος επεξεργασίας να γνωρίζει στοιχεία της πρόγνωσης, για να τροποποιήσει τρόπους εκπαίδευσης, διατροφής, άσκησης, φαρμακευτικής αγωγής κ.λπ. των υποκειμένων, χωρίς όμως να κοινοποιεί στα υποκείμενα τις προγνώσεις που έγιναν για αυτά, σεβόμενος την προτίμησή τους να μην γνωρίζουν.

Η παραπάνω αναφερόμενη επιλογή θα πρέπει να θεωρηθεί αρκετά πιθανή, αφού όπως έχει ήδη τονιστεί, η καλή κατάσταση της υγείας των υποκειμένων (στελεχών των ΕΔ) είναι προαπαιτούμενο της πολεμικής τους ετοιμότητας και αποτελεί αποφασιστικής σημασίας παράγοντα μαχητικής ισχύος του στρατεύματος. Η γνώση της κατάστασης της υγείας του προσωπικού, μπορεί να αποτελέσει ιδιαίτερα πολύτιμη πληροφορία για τον εκάστοτε επιχειρησιακό διοικητή, για την ανάληψη, συνέχιση ή μεταβολή των πολεμικών επιχειρήσεων ή των ανατιθέμενων, σε αυτόν, αποστολών. Η ηγεσία των ΕΔ εκτιμάται ότι θα ήθελε να έχει στη διάθεσή της αυτή την πολύτιμη γνώση όχι μόνο για το αξιόμαχο του προσωπικού της, αλλά και διότι υγιή στελέχη σημαίνει θετικό αντίκτυπο στην εθνική μας οικονομία, αφού από έρευνες έχει προκύψει ότι η καλή υγεία του πληθυσμού, έχει θετική και στατιστικά σημαντική επίδραση στην οικονομική ανάπτυξη μιας χώρας (Bloom et al., 2004).

Σύμφωνα εξάλλου με το άρθρο 21 παρ 1 του ΓΚΠΔ (δικαίωμα εναντίωσης), «το υποκείμενο των δεδομένων δικαιούται να αντιτάσσεται, ανά πάσα στιγμή και για λόγους που σχετίζονται με την ιδιαίτερη κατάστασή του, στην επεξεργασία δεδομένων προσωπικού χαρακτήρα που το αφορούν, η οποία βασίζεται στο άρθρο 6 παράγραφος 1 στοιχείο ε) ή στ), περιλαμβανομένης της κατάρτισης προφίλ, βάσει των εν λόγω διατάξεων. Ο υπεύθυνος επεξεργασίας δεν υποβάλλει πλέον τα δεδομένα προσωπικού χαρακτήρα σε επεξεργασία, εκτός εάν ο υπεύθυνος επεξεργασίας καταδείξει επιτακτικούς και νόμιμους λόγους για την επεξεργασία οι οποίοι υπερσχύουν των συμφερόντων, των δικαιωμάτων και των ελευθεριών του υποκειμένου των δεδομένων ή για τη θεμελίωση, άσκηση ή υποστήριξη νομικών αξιώσεων».

Σύμφωνα με την αιτιολογική σκέψη 69, «όταν δεδομένα προσωπικού χαρακτήρα μπορούν να υποβληθούν νόμιμα σε επεξεργασία επειδή η επεξεργασία είναι αναγκαία για την εκπλήρωση καθήκοντος που εκτελείται προς το δημόσιο συμφέρον ή κατά την άσκηση δημόσιας εξουσίας που έχει ανατεθεί στον υπεύθυνο επεξεργασίας ή για λόγους έννομων συμφερόντων του υπευθύνου επεξεργασίας ή τρίτου μέρους, κάθε υποκείμενο των δεδομένων θα πρέπει να δικαιούται παρ' όλα αυτά να αντιταχθεί στην επεξεργασία τυχόν δεδομένων προσωπικού χαρακτήρα που αφορούν την ιδιαίτερη κατάστασή του. Θα πρέπει να εναπόκειται στον υπεύθυνο επεξεργασίας να αποδείξει ότι τα επιτακτικά έννομα συμφέροντά του υπερσχύουν ενδεχομένως των συμφερόντων ή των θεμελιωδών δικαιωμάτων και ελευθεριών του υποκειμένου των δεδομένων».

Τα παραπάνω εμπόδια για την ενδεχόμενη εναντίωση του στελέχους των ΕΔ στην επεξεργασία των δεδομένων υγείας του μέσω της Προγνωστικής Ανάλυσης, μπορούν όπως είναι αυτονόητο να ξεπεραστούν, όταν το υποκείμενο δίνει τη συγκατάθεσή του, π.χ. λίγο πριν υποβληθεί σε υγειονομικές εξετάσεις, από τις οποίες θα προκύψουν τα δεδομένα υγείας του, για ενδεχόμενη, μελλοντικά, Προγνωστική Ανάλυση. Επίσης, εάν ο υπεύθυνος επεξεργασίας (Διευθύνσεις Υγειονομικού των Κλάδων των ΕΔ) κατορθώσει να καταδείξει επιτακτικούς και νόμιμους λόγους για την επεξεργασία οι οποίοι υπερσχύουν των συμφερόντων, των δικαιωμάτων και των ελευθεριών των υποκειμένων (στελεχών των ΕΔ), τα εν λόγω δεδομένα μπορούν να συνεχίσουν να υποβάλλονται σε επεξεργασία χωρίς πρόβλημα, μέσω της Προγνωστικής Ανάλυσης.

8.8 Αξία Προγνωστικής Ανάλυσης Μαζικών Δεδομένων Υγείας για τα στελέχη των ΕΔ

Όπως επισημάνθηκε στην βιβλιογραφική επισκόπηση της παρούσας έρευνας, εντοπίστηκε ερευνητικό κενό στην αξιοποίηση των δεδομένων υγείας, το οποίο κι επιχειρήθηκε να καλυφθεί μέσω αυτής. Το κενό αυτό έγκειται στην μέχρι στιγμής απουσία, των δυνατοτήτων που η Προγνωστική Ανάλυση των Μαζικών Δεδομένων Υγείας μπορεί να προσφέρει., από την φαρέτρα του ιατρικού και νοσηλευτικού προσωπικού των στρατιωτικών υγειονομικών υπηρεσιών των ΕΔ μας.

Έχει ήδη τονιστεί ότι οι δυνατότητες της Προγνωστικής Ανάλυσης, της Τεχνητής Νοημοσύνης, της Μηχανικής Μάθησης, της Ιατρικής Ακρίβειας και των λοιπών συγκλινουσών τεχνολογιών που αναλύθηκαν, δεν είναι πρωτόγνωρες. Αυτό όμως που λείπει από το στρατιωτικό σύστημα υγείας είναι η αξιοποίησή τους.

Τα Μαζικά Δεδομένα Υγείας των στελεχών των ΕΔ παράγονται διαρκώς, μπορούν να συγκεντρώνονται, να διατηρούνται και να υφίστανται επεξεργασία, με ασφάλεια. Ειδικά με τη νομοθέτηση του Ατομικού Ηλεκτρονικού Φακέλου Υγείας (ΑΗΦΥ), ως υποχρεωτικού, η ανάλυσή τους θα είναι εφικτή και χρησιμότητα.

Κατά τη συνήθη πρακτική των ετησίων υγειονομικών εξετάσεων, υποχρεωτικών για τα στελέχη των ΕΔ, από τη μέρα της κατάταξής τους μέχρι και την αποστρατεία τους, ο έλεγχος για τη γνωμάτευση της καλής κατάστασης της υγείας τους σε ετήσια βάση, κατά κανόνα βασίζεται στην εκτίμηση των τρεχόντων αποτελεσμάτων, της πιο πρόσφατης δηλαδή υγειονομικής τους εξέτασης. Για την εκτίμηση αυτή λαμβάνονται υπόψη πίνακες φυσιολογικών τιμών, παρόμοιων με αυτές του Σχ. 24, μικροβιολογικών εργαστηρίων, αιματολογικών, βιοχημικών, ανοσολογικών και λοιπών εξετάσεων. Σπανίως όμως γίνεται σύγκριση και συνεκτίμηση με τα αντίστοιχα αποτελέσματα της περσινής ή προπέρσινης εξέτασης, εκτός κι αν το ίδιο το στέλεχος φέρει μαζί του παλαιότερες εξετάσεις για σύγκριση ή ο γνωματεύων ιατρός τις ζητήσει, διότι κάτι τον ανησύχησε στις πιο πρόσφατες. Έτσι, αν ο εξεταζόμενος εμφανίζει στις τρέχουσες εξετάσεις του αποτελέσματα εντός των φυσιολογικών τιμών, με διαρκή όμως επιδείνωσή τους σε βάθος διαίτας, αυτό ενδεχομένως θα διαφύγει της δέουσας προσοχής.

Value	Normal Range	Unit	Value	Normal Range	Unit
COMPLETE BLOOD COUNT			URINALYSIS		
Red Blood Cell (RBC)	M: 4.5—5.5 F: 4.0—4.9	x10 ⁶ /ml	Volume	1,000—2,000	mL/day
White Blood Cell (WBC)	4,500—10,000	cells/mcL		30	mL/hour
Platelets	100,000—450,000	cells/mcL	Specific Gravity	1.010—1.020	
Hemoglobin (Hgb)	M: 13.5—16.5 F: 12.0—15.0 Pregnant: 10—15	g/dL	pH	4.5—8	
Hematocrit (Hct)	M: 41—50% F: 36—44%		Casts	1—2	per high power field
Mean Corpuscular Volume (MCV)	80—100	fL	Glucose	none	
			Protein	none	
			Uric Acid	F: 3.5—7.2 M: 2.6—6.0	mg/dL

Σχ. 24 Πίνακες φυσιολογικών εργαστηριακών τιμών εξετάσεων αίματος και ούρων (Πηγή: nurseslabs.com)

Αν λάβουμε υπόψη το γεγονός των συχνών μεταθέσεων, τόσο των στελεχών – ασθενών – εξεταζομένων, όσο και των ιατρών που γνωματεύουν τις ετήσιες εξετάσεις τους, συνήθως οι εξετάσεις πραγματοποιούνται σε διαφορετικά στρατιωτικά νοσοκομεία και γνωματεύονται από διαφορετικούς ιατρούς, αναλόγως της νέας έδρας του μετατιθέμενου (ασθενούς ή ιατρού).

Δεν υπάρχει η δυνατότητα παρακολούθησης του στελέχους από «προσωπικό» στρατιωτικό ιατρό, που να γνωρίζει με ακρίβεια το ιστορικό του. Βέβαια το εμπόδιο αυτό πλέον κάμπτεται, λόγω των ΑΗΦΥ και της εύκολης πρόσβασης στο ιστορικό του ασθενούς, ανεξάρτητα από τη γεωγραφική θέση που υπηρετεί.

Από τα παραπάνω, προκύπτει η αξία της Προγνωστικής Ανάλυσης των Μαζικών Δεδομένων Υγείας για τα στελέχη των ΕΔ, διότι όχι μόνο θα δίνει τη δυνατότητα συγκρίσεων με παλαιότερες εξετάσεις, αλλά θα επιτρέπει την έγκαιρη και αξιόπιστη πρόγνωση και πρόληψη πιθανής αναπτυσσόμενης ασθένειας ή αναχαίτιση της επιδείνωσης μια κατάστασης.

Τα μοντέλα πρόβλεψης, θα μπορούν να χρησιμοποιηθούν ως συστήματα υποστήριξης λήψης ιατρικών αποφάσεων, χωρίς ασφαλώς να καταργούν την επιστημονική γνώση και γνώμη των ιατρών, δρώντας όμως συμβουλευτικά και συμπληρωματικά. Επίσης οι δυνατότητες των τεχνολογιών που αναλύθηκαν και κυρίως της Ιατρικής Ακρίβειας θα επιτρέπουν τη στοχευμένη και όχι γενικευμένη αντιμετώπιση περιστατικών υγείας, προσαρμοσμένη στις ιδιαιτερότητες, ατομικές, κοινωνικές κ.λπ. του κάθε στελέχους.

Τα οφέλη από την αξιοποίηση της Προγνωστικής Ανάλυσης εκτιμάται ότι θα συμβάλλουν στην προαγωγή της σωματικής και ψυχικής υγείας των στελεχών, θα ενισχύσουν το επιστημονικό έργο ιατρών και νοσηλευτών και θα οδηγήσουν στην εξοικονόμηση πολύτιμων υγειονομικών πόρων, με άμεσο αντίκτυπο και στην εθνική μας οικονομία.

9. Ανάπτυξη και Δοκιμή Μοντέλων Προγνωστικής Ανάλυσης

Προς τεκμηρίωση της προτεινόμενης τεχνολογίας, όπως αυτή αναλύθηκε στις προηγούμενες ενότητες, θα επιχειρηθεί στη συνέχεια η δημιουργία μοντέλων Μηχανικής Μάθησης και ανάλυσης Μαζικών Δεδομένων Υγείας, από αρχεία μορφής .csv που θα αναζητηθούν στο Διαδίκτυο και θα επιχειρηθούν προγνωστικές αναλύσεις με μοντέλα «Ταξινόμησης» για προβλέψεις καταστάσεων, καθώς και με μοντέλα «Παλινδρόμησης» για προβλέψεις αριθμητικών τιμών.

9.1 Συντελεστές Μοντελοποίησης

9.1.1 Δεδομένα Υγείας Δοκιμών

Για τη δημιουργία των μοντέλων Μηχανικής Μάθησης και ανάλυσης Μαζικών Δεδομένων Υγείας, είναι απαραίτητη η χρήση αρχείων δεδομένων υγείας, σε μορφή .csv. Αυτά τα αρχεία θα πρέπει να έχουν ικανοποιητικό αριθμό χαρακτηριστικών (features) στις στήλες των αντίστοιχων πινάκων (π.χ. εργαστηριακές μετρήσεις από μικροβιολογικές εξετάσεις των υποκειμένων), τιμές αντιπροσωπευτικές και ρεαλιστικές, κατά το δυνατόν όχι κενές (null), ενώ το πλήθος των εγγραφών που αντιπροσωπεύουν το δείγμα προς ανάλυση (πλήθος υποκειμένων ασθενών), θα πρέπει να είναι αρκετά μεγάλο, ώστε να επιτρέπει αξιόπιστες δοκιμές και προβλέψεις, όχι όμως έτσι δομημένο που να οδηγεί το μοντέλο σε «over-fitting», κατάσταση κατά την οποία το μοντέλο θυμάται ένα τεράστιο αριθμό παραδειγμάτων, αντί να μαθαίνει να παρατηρεί τα χαρακτηριστικά.

Μετά από έρευνα στο Διαδίκτυο επιλέχθηκε ο ιστότοπος data.world, ο οποίος παρέχει περισσότερα από 4.400 datasets σχετικά με το λήμμα «health». Πολλά από αυτά τα σύνολα δεδομένων διαπιστώθηκε ότι χρησιμοποιούνται σε μεγάλο αριθμό projects από ερευνητές, γεγονός που εγγυάται σε αρκετά ασφαλή επίπεδα, την αξιοπιστία τους και τις προϋποθέσεις που αναφέρθηκαν στην προηγούμενη παράγραφο.

9.1.2 Πλατφόρμα Μηχανικής Μάθησης

Για την ανάπτυξη των κατάλληλων μοντέλων Μηχανικής Μάθησης και ανάλυσης Μαζικών Δεδομένων Υγείας, απαιτείται η χρήση του κατάλληλου λογισμικού ή εργαλείου, το οποίο παρέχει τη δυνατότητα μεταφόρτωσης σχετικών αρχείων μορφής .csv, όπως αυτά που αντλήθηκαν από τον ιστότοπο data.world και Προγνωστικής Ανάλυσής τους, με μεθόδους Μηχανικής Μάθησης.

Μετά από έρευνα στο Διαδίκτυο, επιλέχθηκε το εργαλείο που διατίθεται στον ιστότοπο Knowi.com, το οποίο παρέχει δυνατότητες Μηχανικής Μάθησης και ανάλυσης δεδομένων υγείας, που περιέχονται σε κατάλληλα δομημένα αρχεία .csv. Το συγκεκριμένο εργαλείο επιτρέπει, μεταξύ άλλων, την επιλογή μεθόδων Ταξινόμησης και Παλινδρόμησης. Η πρώτη θα χρησιμοποιηθεί για την πρόγνωση καταστάσεων (π.χ. το υποκείμενο είναι ή δεν είναι διαβητικός ασθενής), ενώ η δεύτερη θα χρησιμοποιηθεί για την πρόγνωση αριθμητικών τιμών (π.χ. ποια προβλέπεται να είναι η τιμή της γλυκοζυλιωμένης αιμοσφαιρίνης του υποκειμένου ασθενούς).

Μετά τη δημιουργία του κατάλληλου μοντέλου και την επιλογή της μεθόδου Μηχανικής Μάθησης (Ταξινόμησης ή Παλινδρόμησης), το εργαλείο δίνει τη δυνατότητα ανάλυσης και προετοιμασίας των δεδομένων σε έξι επιμέρους στάδια (επιβεβαίωσης των τύπων δεδομένων ή data types, επισήμανσης και απομάκρυνσης των outliers, ανίχνευσης πιθανώς ελλιπών τιμών, κλιμάκωσης των μεταβλητών, διακριτής ομαδοποίησης, ελέγχου των μεταβλητών dummy). Στη συνέχεια ακολουθεί εκπαίδευση του μοντέλου, επιλογή της μορφής αλγορίθμου για την εκπαίδευση και πρόβλεψη της επιθυμητής μεταβλητής (π.χ. με Δένδρο Απόφασης ή με Γραμμική Παλινδρόμηση Ελαχίστων Τετραγώνων), υπολογισμός του ποσοστού ακρίβειας της πρόβλεψης της εν λόγω μεταβλητής, εμφάνιση των αποτελεσμάτων, καθώς και δυνατότητα δημοσίευσης του μοντέλου που αναπτύχθηκε.

Για δεδομένα που δείχνουν μια γραμμική σχέση μεταξύ των μεταβλητών x και y , επιθυμούμε να βρούμε την καλύτερη γραμμή που ταιριάζει σε αυτήν τη σχέση και η γραμμή αυτή ονομάζεται γραμμή παλινδρόμησης με εξίσωση $y=a+bx$, που αποτελεί το απλούστερο γραμμικό μοντέλο, κι όπου x είναι η ανεξάρτητη μεταβλητή, y είναι η εξαρτημένη μεταβλητή, a είναι η σταθερά και b είναι η παράμετρος η οποία αφορά τη σχέση μεταξύ των μεταβλητών x και y . Αν υποθέσουμε ότι υπάρχουν n πλήθους παρατηρήσεις για τις μεταβλητές x και y , τότε μπορεί να τοποθετηθεί μια ευθεία γραμμή που να περνά όσο το δυνατόν πιο κοντά από τα σημεία που ορίζουν τα ζεύγη τιμών (x,y) και να εκτιμηθούν οι παράμετροι a και b . Η ευθεία θα τέμνει των άξονα Y στο σημείο a (σταθερά) και θα έχει κλίση b .

Μια μέθοδος εκτίμησης των παραμέτρων a και b είναι η μέθοδος παλινδρόμησης «Ελαχίστων Τετραγώνων» (Ordinary Least Squares – OLS) σύμφωνα με την οποία ελαχιστοποιείται το άθροισμα των τετραγώνων των διαφορών, μεταξύ των παρατηρημένων και των εκτιμημένων τιμών, της εξαρτημένης μεταβλητής y . Η γραμμή παλινδρόμησης ελαχίστων τετραγώνων (Least Squares Regression Line) είναι δηλαδή η γραμμή που κάνει την κατακόρυφη απόσταση από τα σημεία δεδομένων έως τη γραμμή παλινδρόμησης, όσο το δυνατόν μικρότερη. Ονομάζεται «Ελαχίστων Τετραγώνων» επειδή η καλύτερη γραμμή προσαρμογής είναι αυτή που ελαχιστοποιεί τη διακύμανση, δηλαδή το άθροισμα των τετραγώνων των σφαλμάτων (Glen, 2014).

9.2 Σενάρια Προγνωστικής Ανάλυσης - Αποτελέσματα

9.2.1 Πρόγνωση Κατάστασης Διαβητικός ή Όχι

Με τη μέθοδο της Ταξινόμησης, επιχειρήθηκε η πρόγνωση της κατάστασης (class) «Positive» ή «Negative», που αντιστοιχεί στην τιμή «είναι διαβητικός» ή «δεν είναι διαβητικός» αντίστοιχα, για τα υποκείμενα ενός συνόλου δεδομένων (dataset), που αντλήθηκε από τον ιστότοπο data.world.

Το εν λόγω σύνολο δεδομένων περιέχει 520 εγγραφές υποκειμένων, με 17 χαρακτηριστικά (features), όπως η ηλικία, το φύλο, η εμφάνιση ή όχι πολυουρίας και πολυδιψίας, η ξαφνική απώλεια βάρους, η παχυσαρκία κ.ά, ενώ στην τελευταία στήλη του πίνακα δεδομένων υπάρχει η διάγνωση της κατάστασης «Positive» ή «Negative», δηλαδή η γνώμη του ιατρού για το αν, με βάση τα υπόλοιπα 16 χαρακτηριστικά, το υποκείμενο «είναι διαβητικός» ή «δεν είναι διαβητικός». Η κατάσταση αυτή αποτελεί και τη μεταβλητή (attribute) την οποία το μοντέλο μας θέλουμε να προβλέψει, με τη μέθοδο της Ταξινόμησης και πιο συγκεκριμένα με την προσέγγιση του Δέντρου Αποφάσεων (Decision Tree).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
	Age	Gender	Polyuria	Polydipsia	sudden weight loss	weakness	Polyphagia	Genital thrush	visual blurring	itching	irritability	delayed healing	partial paresis	muscle stiffness	Alopecia	Obesity	class
1	40	Male	No	Yes	No	Yes	No	No	No	Yes	No	Yes	No	Yes	Yes	Yes	Positive
2	58	Male	No	No	No	Yes	No	No	Yes	No	No	No	Yes	No	Yes	No	Positive
3	41	Male	Yes	No	No	Yes	Yes	No	No	Yes	No	Yes	No	Yes	Yes	No	Positive
4	45	Male	No	No	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No	No	No	No	Positive
5	60	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Positive
6	55	Male	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Positive
7	57	Male	Yes	Yes	No	Yes	Yes	Yes	No	No	No	Yes	Yes	No	No	No	Positive
8	66	Male	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No	Positive
9	67	Male	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Positive
10	70	Male	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No	No	Yes	No	Positive
11	44	Male	Yes	Yes	No	Yes	No	Yes	No	No	Yes	Yes	No	Yes	Yes	No	Positive
12	38	Male	Yes	Yes	No	No	Yes	Yes	No	Yes	No	Yes	No	Yes	No	No	Positive
13	35	Male	Yes	No	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	No	Positive
14	61	Male	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes	Positive
15	60	Male	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	No	No	Positive
16	58	Male	Yes	Yes	No	Yes	Yes	No	No	No	No	Yes	Yes	Yes	No	No	Positive
17	54	Male	Yes	Yes	Yes	Yes	No	Yes	No	No	No	Yes	No	Yes	No	No	Positive
18	67	Male	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Positive
19	66	Male	Yes	Yes	No	Yes	Yes	No	Yes	No	No	No	Yes	Yes	No	No	Positive
20	43	Male	Yes	Yes	Yes	Yes	No	Yes	No	No	No	No	No	No	No	No	Positive
21	62	Male	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	No	Yes	Yes	No	No	Positive
22	54	Male	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	No	Positive
23	39	Male	Yes	No	Yes	No	No	Yes	No	Yes	Yes	No	No	No	Yes	No	Positive
24	48	Male	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No	No	No	No	Positive
25	58	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	Yes	Yes	Yes	No	Yes	Positive
26	32	Male	No	No	No	No	No	Yes	No	No	Yes	Yes	No	No	No	Yes	Positive
27	42	Male	No	No	No	Yes	Yes	No	No	No	Yes	No	No	Yes	No	No	Positive
28	52	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	No	No	Positive
29	38	Male	No	Yes	No	No	No	Yes	No	No	No	No	No	No	Yes	No	Positive
30	53	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Positive
31	57	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No	Yes	No	No	No	Positive

Μετά την ανάλυση και προετοιμασία των δεδομένων στα έξι στάδια που προσφέρει το εργαλείο της Knowi.com, το μοντέλο που δημιουργήθηκε με την προσέγγιση του Δέντρου Αποφάσεων, προβλέπει τη μεταβλητή «class» με ακρίβεια άνω του 97%. Όπως φαίνεται από τη σύγκριση των δύο τελευταίων στηλών (της 17ης και της 18ης που προστέθηκε ως πρόβλεψη), οι προγνώσεις είναι επιτυχείς.

Age	Gender	Polyuria	Polydipsia	sudden weight loss	weakness	Polyphagia	Genital thrush	visual blurring	itching	irritability	delayed healing	partial paresis	muscle stiffness	Alopecia	Obesity	class	Predicted class
60	Male	No	No	No	Yes	No	No	No	No	No	No	Yes	Yes	No	No	Negative	Negative
60	Male	No	No	No	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	No	Negative	Negative
60	Male	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	No	No	Positive	Positive
60	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Positive	Positive
61	Male	No	No	No	Yes	No	Yes	No	Yes	No	Yes	No	No	Yes	No	Negative	Negative
61	Male	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Positive	Positive
61	Male	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Positive	Positive
61	Male	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Positive	Positive
62	Female	Yes	Yes	Yes	Yes	No	No	Yes	No	No	No	Yes	No	No	Yes	Positive	Positive
62	Male	Yes	No	No	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Positive	Positive
62	Male	Yes	No	No	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Positive	Positive
62	Male	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	No	Yes	Yes	No	No	Positive	Positive
62	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No	Yes	Yes	No	No	Positive	Positive
63	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No	No	Yes	Yes	Yes	Positive	Positive
63	Male	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No	No	Yes	Yes	Yes	Positive	Positive
64	Male	No	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Negative	Negative
64	Male	No	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Negative	Negative
64	Male	No	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Negative	Negative
64	Male	No	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Negative	Negative
65	Female	No	No	No	Yes	No	No	No	Yes	No	Yes	No	No	Yes	No	Negative	Negative
65	Female	Yes	Yes	No	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No	No	No	Positive	Positive
65	Female	Yes	Yes	No	Yes	Yes	No	No	Yes	No	No	Yes	Yes	No	No	Positive	Positive

9.2.2 Πρόγνωση Τιμής Γλυκοζυλιωμένης Αιμοσφαιρίνης

Με τη μέθοδο της Παλινδρόμησης, επιχειρήθηκε η πρόγνωση της τιμής της «Γλυκοζυλιωμένης Αιμοσφαιρίνης» (HbA1c), για τα υποκείμενα ενός συνόλου δεδομένων (dataset), που αντλήθηκε από τον ιστότοπο data.world. Το εν λόγω σύνολο δεδομένων περιέχει 1009 εγγραφές υποκειμένων, με 7 χαρακτηριστικά (features), όπως η ηλικία, το φύλο, η γλυκόζη νηστείας στο πλάσμα, η γλυκόζη πλάσματος δύο ώρες μετά από τη φόρτιση με 75 g γλυκόζης από το στόμα, η τιμή της γλυκοζυλιωμένης αιμοσφαιρίνης (HbA1c), ενώ στην τελευταία στήλη υπάρχει η διάγνωση της κατάστασης (Class), όπου «0» σημαίνει ότι το υποκείμενο είναι φυσιολογικό, «1» σημαίνει ότι έχει διαταραγμένη γλυκόζη νηστείας, «2» ότι έχει διαταραγμένη ανοχή στη γλυκόζη και «3» ότι πάσχει από σακχαρώδη διαβήτη. Η τιμή της γλυκοζυλιωμένης αιμοσφαιρίνης (HbA1c) αποτελεί τη μεταβλητή (attribute) την οποία το μοντέλο μας θέλουμε να προβλέψει, με τη μέθοδο της παλινδρόμησης και πιο συγκεκριμένα με τη μέθοδο των “Ελαχίστων Τετραγώνων” (Ordinary Least Squares – OLS).

	A	B	C	D	E	F	G
1	Age	Gender	Plasma F	Plasma 2Hr	Glucose R	HbA1c	Class
2	69	F	168	232	144	6.7	3
3	66	F	180	238	158	6.7	3
4	62	M	128	202	165	6.7	3
5	68	M	180	248	144	7.5	3
6	63	F	123	151	165	6.3	2
7	60	F	122	153	172	6.3	2
8	49	M	125	155	158	6.3	2
9	62	F	125	157	144	6.3	2
10	56	M	116	136	161	6.4	1
11	55	M	120	137	162	6.4	1
12	57	F	126	144	164	6.4	2
13	59	M	122	146	171	6.4	2
14	59	F	122	150	170	6.4	2
15	61	M	124	152	167	6.4	2
16	33	F	95	134	142	5.3	0
17	33	F	95	134	142	5.3	0
18	53	M	105	131	141	5.3	0
19	33	F	95	134	142	5.3	0
20	33	F	95	134	142	5.3	0
21	33	F	95	134	142	5.3	0
22	33	F	95	134	142	5.3	0
23	67	M	178	237	165	7.2	3
24	70	F	181	243	167	7.2	3
25	64	F	134	205	165	7.2	3
26	61	F	128	207	151	7.2	3
27	63	M	130	213	165	7.2	3
28	66	F	142	219	167	7.2	3
29	68	F	154	225	170	7.2	3

Μετά την ανάλυση και προετοιμασία των δεδομένων στα έξι στάδια που προσφέρει το εργαλείο της Knowi.com, το μοντέλο που δημιουργήθηκε με τη μέθοδο παλινδρόμησης των «Ελαχίστων Τετραγώνων» (OLS), προβλέπει την τιμή της γλυκοζυλιωμένης αιμοσφαιρίνης (HbA1c), με τιμή τετραγωνικής ρίζας του μέσου τετραγωνικού σφάλματος (Root Mean Square Error – RMSE), στο 20,0079. Η τιμή αυτή δείχνει πόσο μακριά είναι οι τιμές των μετρήσεων του πίνακα, από την παλινδρόμηση ή τη μοντελοποιημένη γραμμή, ενώ όπως φαίνεται από τη σύγκριση των δύο τελευταίων στηλών (7ης και της 8ης που προστέθηκε ως πρόβλεψη), οι προγνώσεις είναι επιτυχείς.

Age	Gender	Plasma F	Plasma 2Hr	Glucose R	Class	HbA1c	Predicted HbA1c
69	F	168	232	144	3	6.7	6.756250626311893
66	F	180	238	158	3	6.7	6.900512935728118
62	M	128	202	165	3	6.7	6.723024740631765
68	M	180	248	144	3	7.5	6.82826670875349
63	F	123	151	165	2	6.3	6.450417340456157
60	F	122	153	172	2	6.3	6.487046593977389
49	M	125	155	158	2	6.3	6.346360853581528
62	F	125	157	144	2	6.3	6.292297503641427
56	M	116	136	161	1	6.4	6.163193662084493
55	M	120	137	162	1	6.4	6.181438590064694
57	F	126	144	164	2	6.4	6.413935071845341
59	M	122	146	171	2	6.4	6.484900850130577
59	F	122	150	170	2	6.4	6.462936218149279
61	M	124	152	167	2	6.4	6.477527230261066
33	F	95	134	142	0	5.3	5.5927217414739
33	F	95	134	142	0	5.3	5.5927217414739
53	M	105	131	141	0	5.3	5.748513266590297
33	F	95	134	142	0	5.3	5.5927217414739
33	F	95	134	142	0	5.3	5.5927217414739
33	F	95	134	142	0	5.3	5.5927217414739
67	M	178	237	165	3	7.2	6.970637072662188
70	F	181	243	167	3	7.2	7.002106551835211

9.2.3 Πρόγνωση Κατάστασης Διαβήτη (4 Κατηγορίες)

Με τη μέθοδο της Ταξινόμησης, επιχειρήθηκε η πρόγνωση της κατάστασης 0,1,2 ή 3 για τα υποκείμενα του συνόλου δεδομένων (dataset), του προηγούμενου σεναρίου (που προβλέπει την τιμή της γλυκοζυλιωμένης αιμοσφαιρίνης - HbA1c).

Το εν λόγω σύνολο δεδομένων περιέχει 1009 εγγραφές υποκειμένων, με 7 χαρακτηριστικά (features), όπως η ηλικία, το φύλο, η γλυκόζη νηστείας στο πλάσμα, η γλυκόζη πλάσματος δύο ώρες μετά από τη φόρτιση με 75 g γλυκόζης από το στόμα, η τιμή της γλυκοζυλιωμένης αιμοσφαιρίνης (HbA1c), ενώ στην τελευταία στήλη υπάρχει η διάγνωση της κατάστασης (Class), όπου «0» σημαίνει ότι το υποκείμενο είναι φυσιολογικό, «1» σημαίνει ότι έχει διαταραγμένη γλυκόζη νηστείας, «2» ότι έχει διαταραγμένη ανοχή στη γλυκόζη και «3» ότι πάσχει από σακχαρώδη διαβήτη. Η κατάσταση «Class» αποτελεί τη μεταβλητή (attribute) την οποία το μοντέλο μας θέλουμε να προβλέψει, με τη μέθοδο της Ταξινόμησης και πιο συγκεκριμένα με την προσέγγιση του Δέντρου Αποφάσεων (Decision Tree). Το μοντέλο αυτό είναι ένα διακριτό μοντέλο πολλαπλών παραλλαγών (multivariate discrete), με ταξινομητές για 4 καταστάσεις (states 0,1,2,3), δηλαδή ο αριθμός τους είναι πεπερασμένος (Serrano, 2020).

	A	B	C	D	E	F	G
1	Age	Gender	Plasma F	Plasma 2Hr	Glucose R	HbA1c	Class
2	69	F	168	232	144	6.7	3
3	66	F	180	238	158	6.7	3
4	62	M	128	202	165	6.7	3
5	68	M	180	248	144	7.5	3
6	63	F	123	151	165	6.3	2
7	60	F	122	153	172	6.3	2
8	49	M	125	155	158	6.3	2
9	62	F	125	157	144	6.3	2
10	56	M	116	136	161	6.4	1
11	55	M	120	137	162	6.4	1
12	57	F	126	144	164	6.4	2
13	59	M	122	146	171	6.4	2
14	59	F	122	150	170	6.4	2
15	61	M	124	152	167	6.4	2
16	33	F	95	134	142	5.3	0
17	33	F	95	134	142	5.3	0
18	53	M	105	131	141	5.3	0
19	33	F	95	134	142	5.3	0
20	33	F	95	134	142	5.3	0
21	33	F	95	134	142	5.3	0
22	33	F	95	134	142	5.3	0
23	67	M	178	237	165	7.2	3
24	70	F	181	243	167	7.2	3
25	64	F	134	205	165	7.2	3
26	61	F	128	207	151	7.2	3
27	63	M	130	213	165	7.2	3
28	66	F	142	219	167	7.2	3
29	68	F	154	225	170	7.2	3

Μετά την ανάλυση και προετοιμασία των δεδομένων στα έξι στάδια που προσφέρει το εργαλείο της Knowi.com, το μοντέλο που δημιουργήθηκε με την προσέγγιση του Δέντρου Αποφάσεων, προβλέπει τη μεταβλητή «class» με ακρίβεια 100%. Όπως φαίνεται από τη σύγκριση των δύο τελευταίων στηλών (της 7ης και της 8ης που προστέθηκε ως πρόβλεψη), οι προγνώσεις είναι επιτυχείς.

Age	Gender	Plasma F	Plasma 2Hr	Glucose R	HbA1c	Class	Predicted Class
69	F	168	232	144	6.7	3	3
66	F	180	238	158	6.7	3	3
62	M	128	202	165	6.7	3	3
68	M	180	248	144	7.5	3	3
63	F	123	151	165	6.3	2	2
60	F	122	153	172	6.3	2	2
49	M	125	155	158	6.3	2	2
62	F	125	157	144	6.3	2	2
56	M	116	136	161	6.4	1	1
55	M	120	137	162	6.4	1	1
57	F	126	144	164	6.4	2	2
59	M	122	146	171	6.4	2	2
59	F	122	150	170	6.4	2	2
61	M	124	152	167	6.4	2	2
33	F	95	134	142	5.3	0	0
33	F	95	134	142	5.3	0	0
53	M	105	131	141	5.3	0	0
33	F	95	134	142	5.3	0	0
33	F	95	134	142	5.3	0	0
33	F	95	134	142	5.3	0	0
67	M	178	237	165	7.2	3	3
70	F	181	243	167	7.2	3	3

9.2.4 Πρόγνωση Κατάστασης Μεταβολικό Σύνδρομο ή Όχι

Με τη μέθοδο της Ταξινόμησης, επιχειρήθηκε η πρόγνωση της κατάστασης (MetabolicSyndrome) «MetSyn» ή «No MetSyn», που αντιστοιχεί στην τιμή “πάσχει από μεταβολικό σύνδρομο” ή “δεν πάσχει από μεταβολικό σύνδρομο” αντίστοιχα, για τα υποκειμένα ενός συνόλου δεδομένων (dataset), που αντλήθηκε από τον ιστότοπο data.world. Το εν λόγω σύνολο δεδομένων περιέχει 2401 εγγραφές υποκειμένων, με 15 χαρακτηριστικά (features), όπως η ηλικία, το φύλο, η φυλή, η εμφάνιση ή όχι αλβουμινουρίας, ο δείκτης μάζας σώματος (BMI), η τιμή του ζαχάρου, των τριγλυκεριδίων κ.ά, ενώ στην τελευταία στήλη του πίνακα δεδομένων υπάρχει η διάγνωση της κατάστασης «MetSyn» ή «No MetSyn», δηλαδή η γνώμη του ιατρού για το αν, με βάση τα υπόλοιπα 14 χαρακτηριστικά, το υποκείμενο “πάσχει” ή “δεν πάσχει” από μεταβολικό σύνδρομο. Η κατάσταση αυτή αποτελεί και τη μεταβλητή (attribute) την οποία το μοντέλο μας θέλουμε να προβλέψει, με τη μέθοδο της Ταξινόμησης και πιο συγκεκριμένα με την προσέγγιση του Δέντρου Αποφάσεων (Decision Tree).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	Seqn	Age	Sex	Marital	Income	Race	WaistCirc	BMI	Albuminuria	UrAlbCr	UricAcid	BloodGlucose	HDL	Triglycerides	MetabolicSyndrome	
2	62161	22	Male	Single	8200	White	81	23.3	0	3.88	4.9	92	41	84	No MetSyn	
3	62164	44	Female	Married	4500	White	80.1	23.2	0	8.55	4.5	82	28	56	No MetSyn	
4	62169	21	Male	Single	800	Asian	69.6	20.1	0	5.07	5.4	107	43	78	No MetSyn	
5	62172	43	Female	Single	2000	Black	120.4	33.3	0	5.22	5	104	73	141	No MetSyn	
6	62177	51	Male	Married	9000	Asian	81.1	20.1	0	8.13	5	95	43	126	No MetSyn	
7	62178	80	Male	Widowed	300	White	112.5	28.5	0	9.79	4.8	105	47	100	No MetSyn	
8	62184	26	Male	Single	9000	Black	78.8	22.1	0	9.21	5.4	87	61	40	No MetSyn	
9	62189	30	Female	Married	6200	Asian	80.2	22.4	0	8.78	6.7	83	48	91	No MetSyn	
10	62191	70	Male	Divorced	1000	Black			1	45.67	5.4	96	35	75	No MetSyn	
11	62195	35	Male		2500	Black		99	28.2	0	2.21	6.7	94	46	86	No MetSyn
12	62199	57	Male		9000	White	107.8	28	0	4.16	6	100	35	98	MetSyn	
13	62202	36	Male	Married	8200	MexAmerican	97	24.7	1	62.14	6.7	94	58	182	No MetSyn	
14	62205	28	Male	Single	9000	White	106	28.9	0	7.24	8.8	101	40	129	MetSyn	
15	62208	38	Male	Married	3500	Hispanic	82.7	22.2	0	5.53	4.5	91	46	62	No MetSyn	
16	62209	62	Female	Separated	2500	MexAmerican	92.9	26	0	2.33	3.6	138	31	565	MetSyn	
17	62214	22	Female	Separated	3500	Black	89	23.6	0	21.43	4.8	90	54	44	No MetSyn	
18	62215	65	Female	Married	1500	White	94	26.7	0	8.38	3.9	161	52	79	MetSyn	
19	62217	77	Female	Single	1600	MexAmerican	118.6	30.6	1	187.41	4.2	178	46	107	MetSyn	
20	62218	38	Female	Single	8200	Black	120.3	45.4	0	16.63	5.6	102	36	162	MetSyn	
21	62224	29	Female	Married	5400	Hispanic	90.1	27.2	0	7.39	4.5	84	56	121	No MetSyn	
22	62226	80	Male	Married	5400	Hispanic	97.2	28.4	0	6.94	4.2	111	37	105	No MetSyn	
23	62228	50	Male	Married	8200	White	136.1	43.4	0	4.71	6.5	204	33	192	MetSyn	
24	62231	48	Female	Married	9000	MexAmerican	111	33.2	0	4.84	5.7	99	59	122	No MetSyn	
25	62236	61	Male	Married	9000	Asian	90.8	23.2	0	2.34	5.3	102	39	135	MetSyn	
26	62239	22	Female		300	Hispanic	78.9	23.3	0	5.33	4.1	90	42	68	No MetSyn	
27	62248	65	Male	Married	4500	Asian	104.2	26.6	1	31.46	6.3	106	30	326	MetSyn	
28	62250	34	Male	Married	9000	MexAmerican	95.1	25.9	0	5.89	5	101	50	98	No MetSyn	
29	62264	77	Male	Married	3500	White	112	31.1	0	25.97	6.4	130	55	138	MetSyn	
30	62265	52	Male	Married	3500	White	110	31.4	1	120.73	4.1	258	39	223	MetSyn	

Μετά την ανάλυση και προετοιμασία των δεδομένων στα έξι στάδια που προσφέρει το εργαλείο της Knowi.com, το μοντέλο που δημιουργήθηκε με την προσέγγιση του Δέντρου Αποφάσεων, προβλέπει τη μεταβλητή «MetabolicSyndrome» με ακρίβεια άνω του 82%. Όπως φαίνεται από τη σύγκριση των δύο τελευταίων στηλών, οι προγνώσεις είναι στην πλειονότητά τους επιτυχείς, ενώ υπάρχουν ελάχιστες λανθασμένες προβλέψεις, όπως στην 9η εγγραφή με Seqn 62205, όπου ενώ θα έπρεπε να υπάρχει πρόβλεψη ότι το υποκείμενο «πάσχει», το μοντέλο πρόβλεψε ότι «δεν πάσχει».

Seqn	Age	Sex	Race	BMI	Albuminuria	UrAlbCr	UricAcid	BloodGlucose	HDL	Triglycerides	MetabolicSyndrom	Predicted Metabd
62161	22	Male	White	23.3	0	3.88	4.9	92	41	84	No MetSyn	No MetSyn
62164	44	Female	White	23.2	0	8.55	4.5	82	28	56	No MetSyn	No MetSyn
62169	21	Male	Asian	20.1	0	5.07	5.4	107	43	78	No MetSyn	No MetSyn
62172	43	Female	Black	33.3	0	5.22	5.0	104	73	141	No MetSyn	No MetSyn
62178	80	Male	White	28.5	0	9.79	4.8	105	47	100	No MetSyn	No MetSyn
62184	26	Male	Black	22.1	0	9.21	5.4	87	61	40	No MetSyn	No MetSyn
62189	30	Female	Asian	22.4	0	8.78	6.7	83	48	91	No MetSyn	No MetSyn
62202	36	Male	MexAmerican	24.7	1	62.14	6.7	94	58	182	No MetSyn	No MetSyn
62205	28	Male	White	28.9	0	7.24	8.8	101	40	129	MetSyn	No MetSyn
62208	38	Male	Hispanic	22.2	0	5.53	4.5	91	46	62	No MetSyn	No MetSyn
62209	62	Female	MexAmerican	26.0	0	2.33	3.6	138	31	565	MetSyn	MetSyn
62214	22	Female	Black	23.8	0	21.43	4.8	90	54	44	No MetSyn	No MetSyn
62215	65	Female	White	26.7	0	8.38	3.9	161	52	79	MetSyn	MetSyn
62217	77	Female	MexAmerican	30.6	1	187.41	4.2	178	46	107	MetSyn	MetSyn
62218	38	Female	Black	45.4	0	16.63	5.6	102	36	162	MetSyn	MetSyn
62224	29	Female	Hispanic	27.2	0	7.39	4.5	84	56	121	No MetSyn	No MetSyn
62226	80	Male	Hispanic	28.4	0	6.94	4.2	111	37	105	No MetSyn	No MetSyn
62228	50	Male	White	43.4	0	4.71	6.5	204	33	192	MetSyn	MetSyn
62231	48	Female	MexAmerican	33.2	0	4.84	5.7	99	59	122	No MetSyn	No MetSyn
62236	61	Male	Asian	23.2	0	2.34	5.3	102	39	135	MetSyn	No MetSyn
62248	65	Male	Asian	26.6	1	31.46	6.3	106	30	326	MetSyn	MetSyn
62250	34	Male	MexAmerican	25.9	0	5.89	5.0	101	50	98	No MetSyn	No MetSyn
62264	77	Male	White	31.1	0	25.97	6.4	130	55	138	MetSyn	MetSyn

9.2.5 Πρόγνωση Κατάστασης Καρδιοπαθής ή Όχι

Με τη μέθοδο της Ταξινόμησης, επιχειρήθηκε η πρόγνωση της κατάστασης (Heart Disease) «Presence» ή «Absence», που αντιστοιχεί στην τιμή “είναι καρδιοπαθής” ή “δεν είναι καρδιοπαθής” αντίστοιχα, για τα υποκείμενα ενός συνόλου δεδομένων (dataset), που αντλήθηκε από τον ιστότοπο data.world. Το εν λόγω σύνολο δεδομένων περιέχει 270 εγγραφές υποκειμένων, με 14 χαρακτηριστικά (features), όπως η ηλικία, το φύλο, η τιμή της αρτηριακής πίεσης, ο μέγιστος αριθμός σφίξεων, η παρουσία στηθάγχης μετά από άσκηση κ.ά, ενώ στην τελευταία στήλη του πίνακα δεδομένων υπάρχει η διάγνωση της κατάστασης «Presence» ή «Absence», δηλαδή η γνώμη του ιατρού για το αν, με βάση τα υπόλοιπα 13 χαρακτηριστικά, το υποκείμενο «είναι καρδιοπαθής» ή «δεν είναι καρδιοπαθής». Η κατάσταση αυτή αποτελεί και τη μεταβλητή (attribute) την οποία το μοντέλο μας θέλουμε να προβλέψει, με τη μέθοδο της Ταξινόμησης και πιο συγκεκριμένα με την προσέγγιση του Δέντρου Αποφάσεων (Decision Tree).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Age	Sex	Chest pain type	BP	Cholesterol	FBS over 120	EKG results	Max HR	Exercise angina	ST depression	Slope of ST	No vessels fluro	Thallium	Heart Disease
2	70	1	4	130	322	0	2	109	0	2,4	2	3	3	Presence
3	67	0	3	115	564	0	2	160	0	1,6	2	0	7	Absence
4	57	1	2	124	261	0	0	141	0	0,3	1	0	7	Presence
5	64	1	4	128	263	0	0	105	1	0,2	2	1	7	Absence
6	74	0	2	120	269	0	2	121	1	0,2	1	1	3	Absence
7	65	1	4	120	177	0	0	140	0	0,4	1	0	7	Absence
8	56	1	3	130	256	1	2	142	1	0,6	2	1	6	Presence
9	59	1	4	110	239	0	2	142	1	1,2	2	1	7	Presence
10	60	1	4	140	293	0	2	170	0	1,2	2	2	7	Presence
11	63	0	4	150	407	0	2	154	0	4,0	2	3	7	Presence
12	59	1	4	135	234	0	0	161	0	0,5	2	0	7	Absence
13	53	1	4	142	226	0	2	111	1	0,0	1	0	7	Absence
14	44	1	3	140	235	0	2	180	0	0,0	1	0	3	Absence
15	61	1	1	134	234	0	0	145	0	2,6	2	2	3	Presence
16	57	0	4	128	303	0	2	159	0	0,0	1	1	3	Absence
17	71	0	4	112	149	0	0	125	0	1,6	2	0	3	Absence
18	46	1	4	140	311	0	0	120	1	1,8	2	2	7	Presence
19	53	1	4	140	203	1	2	155	1	3,1	3	0	7	Presence
20	64	1	1	110	211	0	2	144	1	1,8	2	0	3	Absence
21	40	1	1	140	199	0	0	178	1	1,4	1	0	7	Absence
22	67	1	4	120	229	0	2	129	1	2,6	2	2	7	Presence
23	48	1	2	130	245	0	2	180	0	0,2	2	0	3	Absence
24	43	1	4	115	303	0	0	181	0	1,2	2	0	3	Absence
25	47	1	4	112	204	0	0	143	0	0,1	1	0	3	Absence
26	54	0	2	132	288	1	2	159	1	0,0	1	1	3	Absence
27	48	0	3	130	275	0	0	139	0	0,2	1	0	3	Absence
28	46	0	4	138	243	0	2	152	1	0,3	2	0	3	Absence
29	51	0	3	120	295	0	2	157	0	0,6	1	0	3	Absence
30	58	1	3	112	230	0	2	165	0	2,5	2	1	7	Presence
31	71	0	3	110	265	1	2	130	0	0,0	1	1	3	Absence
32	57	1	3	128	229	0	2	150	0	0,4	2	1	7	Presence

Μετά την ανάλυση και προετοιμασία των δεδομένων στα έξι στάδια που προσφέρει το εργαλείο της Knowi.com, το μοντέλο που δημιουργήθηκε με την προσέγγιση του Δέντρου Αποφάσεων, προβλέπει τη μεταβλητή «Heart Disease» με ακρίβεια άνω του 75%. Όπως φαίνεται από τη σύγκριση των δύο τελευταίων στηλών (της 14ης και της 15ης που προστέθηκε ως πρόβλεψη), οι προγνώσεις είναι επιτυχείς. Αξίζει να σημειωθεί πως ο αριθμός του δείγματος είναι σχετικά μικρός, γεγονός που δικαιολογεί το ποσοστό ακρίβειας στο 75%. Όταν το μοντέλο εκπαιδεύτηκε με τον τετραπλάσιο αριθμό εγγραφών (αντιγραφή των ίδιων εγγραφών τέσσερις φορές), ο εγγραφές ξεπέρασαν τις 1000 και το ποσοστό ακρίβειας άγγιξε το 100% («over-fitting»).

Age	Sex	Chest pain type	BP	Cholesterol	FBS over 120	EKG results	Max HR	Exercise angina	ST depression	Slope of ST	Number of ves	Thallium	Heart Disease	Predicted Heart
70	1	4	130	322	0	2	109	0	2,4	2	3	3	Presence	Presence
67	0	3	115	564	0	2	160	0	1,6	2	0	7	Absence	Absence
57	1	2	124	261	0	0	141	0	0,3	1	0	7	Presence	Presence
64	1	4	128	263	0	0	105	1	0,2	2	1	7	Absence	Absence
74	0	2	120	269	0	2	121	1	0,2	1	1	3	Absence	Absence
65	1	4	120	177	0	0	140	0	0,4	1	0	7	Absence	Absence
56	1	3	130	256	1	2	142	1	0,6	2	1	6	Presence	Presence
59	1	4	110	239	0	2	142	1	1,2	2	1	7	Presence	Presence
60	1	4	140	293	0	2	170	0	1,2	2	2	7	Presence	Presence
63	0	4	150	407	0	2	154	0	4,0	2	3	7	Presence	Presence
59	1	4	135	234	0	0	161	0	0,5	2	0	7	Absence	Absence
53	1	4	142	226	0	2	111	1	0,0	1	0	7	Absence	Absence
44	1	3	140	235	0	2	180	0	0,0	1	0	3	Absence	Absence
61	1	1	134	234	0	0	145	0	2,6	2	2	3	Presence	Presence
57	0	4	128	303	0	2	159	0	0,0	1	1	3	Absence	Absence
71	0	4	112	149	0	0	125	0	1,6	2	0	3	Absence	Absence
46	1	4	140	311	0	0	120	1	1,8	2	2	7	Presence	Presence
53	1	4	140	203	1	2	155	1	3,1	3	0	7	Presence	Presence
64	1	1	110	211	0	2	144	1	1,8	2	0	3	Absence	Absence
40	1	1	140	199	0	0	178	1	1,4	1	0	7	Absence	Absence
67	1	4	120	229	0	2	129	1	2,6	2	2	7	Presence	Presence
48	1	2	130	245	0	2	180	0	0,2	2	0	3	Absence	Absence
43	1	4	115	303	0	0	181	0	1,2	2	0	3	Absence	Absence

9.2.6 Πρόγνωση Κατάστασης Καρκίνος Πνεύμονα ή Όχι

Με τη μέθοδο της Ταξινόμησης, επιχειρήθηκε η πρόγνωση της κατάστασης (Level) «Low» ή «Medium» ή «High», που αντιστοιχεί στην τιμή «χαμηλή» ή «μέτρια» ή «μέση» αντίστοιχα πιθανότητα, να νοσήσει από καρκίνο του πνεύμονα, καθένα από τα υποκείμενα ενός συνόλου δεδομένων (dataset), που αντλήθηκε από τον ιστότοπο data.world.

Το εν λόγω σύνολο δεδομένων περιέχει 1000 εγγραφές υποκειμένων, με 25 χαρακτηριστικά (features), όπως η ηλικία, το φύλο, ενώ σε κλίμακες 1 έως 7 καταχωρήθηκαν καταστάσεις υγείας όπως: αλλεργία στη σκόνη, χρήση αλκοόλ, ισορροπημένη διαίτα, ενεργητικό ή παθητικό κάπνισμα, παχυσαρκία, απώλεια βάρους, κ.λπ. Στην τελευταία στήλη του πίνακα δεδομένων υπάρχει η διάγνωση της κατάστασης «Low» ή «Medium» ή «High», δηλαδή η γνώμη του ιατρού για το αν, με βάση τα υπόλοιπα 24 χαρακτηριστικά, η πιθανότητα του υποκειμένου να νοσήσει από καρκίνο του πνεύμονα είναι «χαμηλή» ή «μέτρια» ή «μέση» αντίστοιχα. Η κατάσταση αυτή αποτελεί και τη μεταβλητή (attribute) την οποία το μοντέλο μας θέλουμε να προβλέψει, με τη μέθοδο της Ταξινόμησης και πιο συγκεκριμένα με την προσέγγιση του Δέντρου Αποφάσεων (Decision Tree).

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	
2	P1	33	1	2	4	5	4	3	2	2	4	3	2	2	4	3	4	2	2	3	1	2	3	4	Low	
3	P10	17	1	3	1	5	3	4	2	2	2	2	2	4	2	3	1	3	7	8	6	2	1	7	2	Medium
4	P100	35	1	4	5	6	5	5	4	6	7	2	3	4	8	8	7	9	2	1	4	6	7	2	High	
5	P1000	37	1	7	7	7	7	6	7	7	7	7	7	7	8	4	2	3	1	4	5	6	7	5	High	
6	P101	46	1	6	8	7	7	7	6	7	7	8	7	7	9	3	2	4	1	4	2	4	2	3	High	
7	P102	35	1	4	5	6	5	5	4	6	7	2	3	4	8	8	7	9	2	1	4	6	7	2	High	
8	P103	52	2	2	4	5	4	3	2	2	4	3	2	2	4	3	4	2	2	3	1	2	3	4	Low	
9	P104	28	2	3	1	4	3	2	3	4	3	1	4	3	1	3	2	2	4	2	2	3	4	3	Low	
10	P105	35	2	4	5	6	5	6	5	5	5	6	6	6	5	1	4	3	2	4	6	2	4	1	Medium	
11	P106	46	1	2	3	4	2	4	3	3	3	2	3	4	4	1	2	4	6	5	4	2	1	5	Medium	
12	P107	44	1	6	7	7	7	7	6	7	7	8	7	7	8	5	3	2	7	8	2	4	5	3	High	
13	P108	64	2	6	8	7	7	7	6	7	7	7	8	7	7	9	6	5	7	2	4	3	1	4	High	
14	P109	39	2	4	5	6	6	5	4	6	6	6	6	6	6	5	3	2	4	3	1	7	5	6	Medium	
15	P11	34	1	6	7	7	7	6	7	7	7	7	7	8	4	2	3	1	4	5	6	7	5	High		
16	P110	27	2	3	1	4	2	3	2	3	3	2	2	4	2	2	2	3	4	1	5	2	6	2	Low	
17	P111	73	1	5	6	6	5	6	5	6	5	8	5	5	5	4	3	6	2	1	2	1	6	2	Medium	
18	P112	17	1	3	1	5	3	4	2	2	2	2	4	2	3	1	3	7	8	6	2	1	7	2	Medium	
19	P113	34	1	6	7	7	7	6	7	7	7	7	7	8	4	2	3	1	4	5	6	7	5	High		
20	P114	36	1	6	7	7	7	7	7	6	7	7	7	7	8	5	7	6	7	8	7	6	2	High		
21	P115	14	1	2	4	5	6	5	5	4	6	5	4	6	5	3	2	1	4	7	2	1	6	Medium		
22	P116	24	1	6	8	7	7	6	7	7	3	8	7	9	6	5	2	5	2	3	2	1	7	6	High	

Μετά την ανάλυση και προετοιμασία των δεδομένων στα έξι στάδια που προσφέρει το εργαλείο της Knowi.com, το μοντέλο που δημιουργήθηκε με την προσέγγιση του Δέντρου Αποφάσεων, προβλέπει τη μεταβλητή «Level» με ακρίβεια άνω του 99,5%. Όπως φαίνεται από τη σύγκριση των δύο τελευταίων στηλών (της 25ης και της 26ης που προστέθηκε ως πρόβλεψη), οι προγνώσεις είναι επιτυχείς.

Patient	Age	Gender	Air Pollu	Alcohol	Dust Alle	Occupat	Genetic	chronic	Balanced	Obesity	Smoking	Passive	Chest Pa	Coughin	Fatigue	Weight L	Shortnes	Wheezin	Swallow	Clubbing	Frequen	Dry Cou	Snoring	Level	Predicts
P1	33	1	2	4	5	4	3	2	2	4	3	2	2	4	3	4	2	2	3	1	2	3	4	Low	Low
P10	17	1	3	1	5	3	4	2	2	2	2	4	2	3	1	3	7	8	6	2	1	7	2	Medium	Medium
P100	35	1	4	5	6	5	5	4	6	7	2	3	4	8	8	7	9	2	1	4	6	7	2	High	High
P1000	37	1	7	7	7	7	6	7	7	7	7	7	7	8	4	2	3	1	4	5	6	7	5	High	High
P101	46	1	6	8	7	7	7	6	7	7	8	7	7	9	3	2	4	1	4	2	4	2	3	High	High
P102	35	1	4	5	6	5	5	4	6	7	2	3	4	8	8	7	9	2	1	4	6	7	2	High	High
P103	52	2	2	4	5	4	3	2	2	4	3	2	2	4	3	4	2	2	3	1	2	3	4	Low	Low
P104	28	2	3	1	4	3	2	3	4	3	1	4	3	1	3	2	2	4	2	2	3	4	3	Low	Low
P105	35	2	4	5	6	5	6	5	5	5	6	6	6	5	1	4	3	2	4	6	2	4	1	Medium	Medium
P106	46	1	2	3	4	2	4	3	3	3	2	3	4	4	1	2	4	6	5	4	2	1	5	Medium	Medium
P107	44	1	6	7	7	7	7	6	7	7	7	8	7	7	5	3	2	7	8	2	4	5	3	High	High
P108	64	2	6	8	7	7	7	6	7	7	7	8	7	7	9	6	5	7	2	4	3	1	4	High	High
P109	39	2	4	5	6	6	5	4	6	6	6	6	6	6	5	3	2	4	3	1	7	5	6	Medium	Medium
P11	34	1	6	7	7	7	6	7	7	7	7	7	8	4	2	3	1	4	5	6	7	5	High	High	
P110	27	2	3	1	4	2	3	2	3	3	2	2	4	2	2	2	3	4	1	5	2	6	2	Low	Low
P111	73	1	5	6	6	5	6	5	6	5	8	5	5	5	4	3	6	2	1	2	1	6	2	Medium	Medium
P112	17	1	3	1	5	3	4	2	2	2	2	4	2	3	1	3	7	8	6	2	1	7	2	Medium	Medium
P113	34	1	6	7	7	7	6	7	7	7	7	7	8	4	2	3	1	4	5	6	7	5	High	High	
P114	36	1	6	7	7	7	7	6	7	7	7	7	7	7	8	5	7	6	7	8	7	6	2	High	High
P115	14	1	2	4	5	6	5	5	4	6	5	4	6	5	3	2	1	4	7	2	1	6	Medium	Medium	
P116	24	1	6	8	7	7	6	7	7	3	8	7	9	6	5	2	5	2	3	2	1	7	6	High	High
P117	53	2	4	5	6	5	5	4	6	7	2	3	4	8	8	7	9	2	1	4	6	7	2	High	High
P118	62	1	6	8	7	7	6	7	7	8	7	7	9	3	2	4	1	4	2	4	2	3	High	High	

10. Επίλογος

10.1 Συμπεράσματα

Στόχος της παρούσα εργασίας ήταν η εξαγωγή χρήσιμων συμπερασμάτων, ως προς τις δυνατότητες της Προγνωστικής Ανάλυσης των Μαζικών Δεδομένων Υγείας, για την πρόγνωση και την πρόληψη σοβαρών και χρόνιων ασθενειών των στελεχών των ΕΔ. Κατά την έρευνα ελήφθησαν υπόψη τα νομοθετικά όρια που θέτει ο ΓΚΠΔ, ενώ παράλληλα μελετήθηκε η χρήση και σύγκλιση σύγχρονων τεχνολογιών που ενδυναμώνουν την Προγνωστική Ανάλυση των δεδομένων αυτών, όπως της Τεχνητής Νοημοσύνης, της Μηχανικής Μάθησης, της Υπολογιστικής Νέφους και του Διαδικτύου των Πραγμάτων Υγείας (IoHT), με την υποστήριξη του δικτύου κινητής τηλεφωνίας 5ης γενιάς (5G).

Κατά την αναζήτηση της βιβλιογραφίας, εντοπίστηκε το ερευνητικό κενό, ως προς τη μελέτη της ιδέας ανάπτυξης και αξιοποίησης μιας έγκαιρης, αξιόπιστης και πολύτιμης πρόγνωσης, βασισμένης στην ανάλυση δεδομένων υγείας, προσωπικών και συνάμα ευαίσθητων, του προσωπικού των ΕΔ. Το γεγονός αυτό εκτιμάται ότι κατέστησε το προς διερεύνηση ζήτημα, σημαντικό και σχετικά πρωτότυπο.

Τα συμπεράσματα που προέκυψαν από την εργασία εκτιμάται ότι θα συνεισφέρουν στη μελλοντική εκμετάλλευση των συγκλινουσών τεχνολογιών για την Προγνωστική Ανάλυση των δεδομένων υγείας, προς όφελος των υποκειμένων που αφορούν, τόσο στελεχών των ΕΔ όσο και του λοιπού πληθυσμού γενικότερα, ενώ προσδοκείται και η παροχή νέων δυνατοτήτων στο ερευνητικό, ιατρικό και νοσηλευτικό προσωπικό, κατά την έγκαιρη διάγνωση και θεραπεία σοβαρών και χρόνιων ασθενειών. Τα κυριότερα από τα συμπεράσματα, που προέκυψαν από την έρευνα, συνοψίζονται παρακάτω:

α. Αξία Προγνωστικής Ανάλυσης των Μαζικών Δεδομένων Υγείας για τις ΕΔ και το προσωπικό τους.

Εκτιμήθηκε συναφώς ότι η έγκαιρη γνώση, μέσω της πρόγνωσης, της κατάστασης της υγείας, είναι μάλλον προς όφελος, πρωτίστως του υποκειμένου των δεδομένων υγείας και δευτερευόντως των ΕΔ, στις οποίες υπηρετεί. Κι αυτό γιατί η καλή κατάσταση της υγείας των υποκειμένων (στελεχών των ΕΔ) είναι προαπαιτούμενο της πολεμικής τους ετοιμότητας και αποτελεί αποφασιστικής σημασίας παράγοντα μαχητικής ισχύος του στρατεύματος. Η γνώση της κατάστασης της υγείας του προσωπικού, μπορεί να αποτελέσει ιδιαίτερα πολύτιμη πληροφορία για τον εκάστοτε επιχειρησιακό διοικητή, για την ανάληψη, συνέχιση ή μεταβολή των πολεμικών επιχειρήσεων ή των ανατιθέμενων, σε αυτόν, αποστολών.

Επιπλέον, θεωρείται σχεδόν βέβαιο ότι η ηγεσία των ΕΔ θα επιθυμούσε να έχει στη διάθεσή της αυτή την πολύτιμη γνώση, όχι μόνο για να μπορεί να αξιολογήσει ανά πάσα χρονική στιγμή το αξιόμαχο, από άποψη υγείας, του προσωπικού της, αλλά και διότι η δυνατότητα να διατηρούν οι ΕΔ μας υγιή στελέχη, συνεπάγεται θετικό αντίκτυπο στην εθνική μας οικονομία, αφού από έρευνες έχει προκύψει ότι η καλή υγεία του πληθυσμού, έχει θετική και στατιστικά σημαντική επίδραση, στην οικονομική ανάπτυξη μιας χώρας (Bloom et al., 2004). Η γνώση αυτή μπορεί να αποτελεί οδηγό για μέγιστη αξιοποίηση του προσωπικού στο χρόνο υπηρεσίας του (30-40 έτη), με προσαρμογή των καθηκόντων του, παροχή ανάλογης εκπαίδευσης και διαφοροποίηση των δεξιοτήτων του, αναλόγως της κατάστασης της υγείας του.

Αν λοιπόν ο υπεύθυνος επεξεργασίας των Μαζικών Δεδομένων Υγείας, εν προκειμένω το Υπουργείο Εθνικής Άμυνας με τις Διευθύνσεις Υγειονομικού των Κλάδων, γνωρίζει στοιχεία της πρόγνωσης, που προέκυψαν από αξιόπιστες αναλύσεις και μπορούν να καταδεικνύουν πού κυμαίνεται η κατάσταση της υγείας του προσωπικού του, ενδεχομένως θα επιδιώξει να το συμβουλευθεί εγκαίρως, να τροποποιήσει τρόπους εκπαίδευσης, διατροφής, άσκησης, φαρμακευτικής αγωγής των υποκειμένων, ώστε να τα βοηθήσει να διατηρήσουν ή να ανακτήσουν την καλή τους υγεία. Η πρόγνωση τιμών διαφόρων ιατρικών εξετάσεων, εφόσον μπορούν να προβλεφθούν με βάση τιμές άλλων εξετάσεων, μπορεί να επιτρέψει εξοικονόμηση πόρων και να αποτρέψει αχρείαστη ταλαιπωρία των υποκειμένων. Αφετέρου, με την εστιασμένη παροχή υγειονομικών υπηρεσιών στο κάθε στέλεχος και όχι γενικά στο στράτευμα (π.χ. επιπτώσεις του καπνίσματος σε συγκεκριμένο καπνιστή και όχι γενικά), θα μπορούν να εξετάζονται απογραφικά όλα τα στελέχη και να ομαδοποιούνται τα αποτελέσματά τους κατά Κλάδο, Όπλο ή Σώμα, κατά εργασία σε ακραίες συνθήκες περιβάλλοντος (χιονοδρόμοι, πιλότοι, πληρώματα υποβρυχίων) κ.λπ.

Για τα δεδομένα υγείας των κληρωτών, εκτιμάται ότι δεν είναι αρκετά ώστε να οδηγήσουν σε συγκέντρωση και επεξεργασία ΜΔΥ ή τουλάχιστον δεν ανέρχονται σε ποσότητες ικανές να υποβληθούν σε Προγνωστική Ανάλυση, εκτός κι αν ενσωματωθούν και τα δεδομένα από τον πολιτικό, προ της στρατεύσής τους βίο, ειδικά όσων παρουσιάζουν αυτοκτονικές τάσεις. Στην περίπτωση, αυτή κατάλληλες τεχνικές ML μπορούν να χρησιμοποιηθούν ως συμπληρωματικό βοήθημα, για ταυτοποίηση και διαστρωμάτωση ατόμων σε κίνδυνο, λόγω παρατηρούμενης αυτοκτονικής συμπεριφοράς (Van Mens et al., 2020).

Τα μοντέλα πρόβλεψης, θα μπορούν να χρησιμοποιηθούν ως συστήματα υποστήριξης λήψης ιατρικών αποφάσεων, χωρίς να καταργούν την επιστημονική γνώση και γνώμη των ιατρών, δρώντας συμβουλευτικά και συμπληρωματικά. Οι δυνατότητες των τεχνολογιών που αναλύθηκαν και κυρίως της Ιατρικής Ακρίβειας θα επιτρέπουν τη στοχευμένη και όχι γενικευμένη αντιμετώπιση περιστατικών υγείας, προσαρμοσμένη στις ιδιαιτερότητες, ατομικές, κοινωνικές κ.λπ. του κάθε στελέχους.

Τα οφέλη από την αξιοποίηση της Προγνωστικής Ανάλυσης εκτιμάται ότι θα συμβάλλουν στην προαγωγή της σωματικής και ψυχικής υγείας των στελεχών, στην καλλιέργεια αισθήματος εμπιστοσύνης στο εν λόγω σύστημα, θα ενισχύσουν το επιστημονικό έργο ιατρών και νοσηλευτών, θα οδηγήσουν σε επέκταση του υφιστάμενου συστήματος Ιατρικής παρακολούθησης των στελεχών, με αξιόπιστα στοιχεία πρόβλεψης και θα οδηγήσουν στην εξοικονόμηση πολύτιμων υγειονομικών πόρων.

β. Εφαρμογές της έρευνας στις ΕΔ

Όλα τα σενάρια Προγνωστικής Ανάλυσης που παρουσιάστηκαν στην παρούσα έρευνα και δοκιμάστηκαν στην πλατφόρμα Knowi.com με ανωνυμοποιημένα δεδομένα δοκιμών, μπορούν να έχουν εφαρμογή στα στελέχη των ΕΔ, καθώς επιλέχθηκαν προγνώσεις, που εκτιμάται ότι αντιμετωπίζονται συνήθως στην καθημερινότητα των περιστατικών υγείας, στα στρατιωτικά νοσοκομεία. Επίσης, η παρέμβαση, κατόπιν ανάλυσης και επεξεργασίας δεδομένων υγείας, στην εντερική μικροχλωρίδα, μέσω αποτελεσματικής διαίτας και λήψης συμπληρωμάτων διατροφής, που προσαρμόζονται σε συγκεκριμένες πληθυσμιακές ομάδες, όπως αυτή των στελεχών, εκτιμάται ότι μπορεί να επιδράσει αποφασιστικά στην υγεία τους (Eetemadi et al., 2020).

Στην περίπτωση του στελεχών των ΕΔ, μπορεί να εφαρμοστεί η απομακρυσμένη παρακολούθηση με συσκευές IoHT, κατά τη διάρκεια της εργασίας, της διαμονής τους στο σπίτι ή σε δραστηριότητες εκτός υπηρεσίας και κυρίως κατά τη συμμετοχή τους σε ασκήσεις ή σε πολεμικές επιχειρήσεις. Ειδικά στην τελευταία περίπτωση, ο εξ αποστάσεως έλεγχος των ζωτικών σημείων των μαχόμενων στελεχών (πίεση, σφίξεις, θερμοκρασία, κορεσμός) μπορεί να δώσει πολύ κρίσιμη πληροφόρηση στον επιχειρησιακό διοικητή, σχετικά με την κατάσταση της υγείας του προσωπικού του και να τον διευκολύνει στη λήψη απόφασης για τις ενδεδειγμένες ενέργειες (π.χ. αίτηση αναπλήρωσης, παραπομπή προς εξέταση, αλλαγή τρόπου ενεργείας, αυτοματοποιημένη αναφορά απωλειών υγείας κ.λπ.). Απαραίτητη προϋπόθεση αποτελεί η ασφαλής διακίνηση των δεδομένων υγείας, μέσω κρυπτασφαλισμένου ασύρματου ευρυζωνικού δικτύου υψηλών ταχυτήτων.

Η εφαρμογή Whoops που λειτουργεί σε περιβάλλον νέφους Azure, για ανίχνευση περιστατικών πτώσης (Mrozek et al., 2020), θα ήταν χρήσιμη για εντοπισμό πτώσης μαχητών στο πεδίο της μάχης και έγκαιρη ενημέρωση των προϊσταμένων τους ή του ιατρικού προσωπικού, για τις απαραίτητες ενέργειες (εντοπισμό, περισυλλογή, αντικατάσταση, πρώτες βοήθειες κ.λπ.).

γ. Ρόλος Πληροφορικής

Επιβεβαιώθηκε στην εργασία ότι ένας από τους σημαντικούς ρόλους της Πληροφορικής είναι να παρέχει τα κατάλληλα εργαλεία, ώστε να μετατρέπει δεδομένα σε πληροφορίες και πληροφορίες σε γνώση, βοηθώντας στην κατανόηση των βαθέων σχέσεων μέσα από τα δεδομένα, ανακτώντας και εξάγοντας υποκείμενα μοτίβα (patterns) μέσα από αυτά. Αποκτά έτσι ζωτική σημασία στην εκμετάλλευση της αποκτηθείσας γνώσης, για την υποστήριξη της περίθαλψης των ασθενών και τελικά την καθοδήγηση της ατομικής τους συμπεριφοράς. Οι σύγχρονοι υπολογιστές μπορούν να επιτύχουν ταχύτητα ομαδοποίηση μεγάλων datasets, με βάση ποικίλα χαρακτηριστικά τους γνωρίσματα, δυνατότητα χρησιμότητας στην περίπτωση της μη Επιτηρούμενης ML.

δ. Αξία Μαζικών Δεδομένων Υγείας

Από την έρευνα φάνηκε ότι τα δεδομένα των ιατρικών φακέλων των ασθενών και οι υπολογιστικές αναλύσεις των δεδομένων αυτών, μπορούν να αποτελέσουν πολύτιμες και αξιόπιστες πηγές συγκέντρωσης ΜΔΥ. Σύνθετοι αλγόριθμοι, που υποστηρίζονται από υψηλής απόδοσης υπολογιστικά συστήματα, επιτρέπουν την μετατροπή αυτών των μεγάλων συνόλων δεδομένων σε πολύτιμη γνώση. Αν συνδυαστούν με μεγάλα σύνολα δεδομένων γονιδιωματικής ή περιβαλλοντολογικής προέλευσης, μπορούν να οδηγήσουν στην πρόγνωση σοβαρών ασθενειών όπως ο καρκίνος, τόσο για μεμονωμένους ασθενείς όσο και για ομάδες ασθενών με συγκεκριμένα χαρακτηριστικά (Willems et al., 2019). Ο συνδυασμός εξάλλου της Ενισχυτικής ML και των ΜΔΥ δημιουργεί μοναδικές ευκαιρίες, για βελτιστοποίηση των θεραπειών στην υγειονομική περίθαλψη, ειδικά για τις επαναλαμβανόμενες (Coronato et al., 2020).

Καθώς όλο και περισσότεροι επιστήμονες της βιοϊατρικής, τόσο σε ακαδημαϊκό όσο και σε επίπεδο βιοτεχνολογίας, χρησιμοποιούν τεχνολογίες υψηλής απόδοσης, όπως απεικονίσεις μικροσκοπίου και σύνθετους ελέγχους, σημειώνεται ραγδαία αύξηση στον όγκο και την ποιότητα των εργαστηριακών δεδομένων που παράγονται. Σε συνδυασμό με διάφορες εξελισσόμενες τεχνικές ML, καθίσταται πλέον όλο και περισσότερο εφικτό να αντληθεί βιολογική εικόνα από αυτά τα παραγόμενα Μαζικά Δεδομένα Υγείας, για να γίνουν κατανοητοί οι μηχανισμοί της νόσου, να εντοπισθούν νέες θεραπευτικές στρατηγικές και να βελτιωθούν τα διαγνωστικά εργαλεία, για κλινική εφαρμογή (Toh et al., 2019).

ε. Ζητήματα ασφάλειας Μαζικών Δεδομένων Υγείας

Όμως, η πληθώρα των δεδομένων υγείας, μέσα από τις νέες δυνατότητες και συγκλίνουσες τεχνολογίες, δημιούργησε, όπως ήταν αναμενόμενο, σοβαρές ανησυχίες σχετικά με την ανάκτηση, την αποθήκευση, την ανάλυση και την ασφάλεια των πληροφοριών των επαγγελματιών υγείας. Ταυτόχρονα δημιουργούνται προκλήσεις, όπως η έλλειψη τυποποίησης δεδομένων, η ανησυχία για ζητήματα απορρήτου, ο περιορισμός της ταχύτητας και η δυνατότητα υποκλοπής κατά τη μεταφορά δεδομένων, η αξιοπιστία αποθήκευσης των δεδομένων κ.λπ.

Μολονότι υπάρχουν εξαιρέσεις στην προστασία των δεδομένων, όταν ιδίως προτάσσεται η Εθνική Ασφάλεια και το δημόσιο συμφέρον ή εφόσον είναι απαραίτητη η επεξεργασία, για σκοπούς προληπτικής ή επαγγελματικής ιατρικής, εκτίμησης της ικανότητας προς εργασία κ.λπ., τεχνικές όπως αυτές της ανωνυμοποίησης και της ψευδωνυμοποίησης των δεδομένων, προσφέρουν μεγαλύτερη προστασία στα υποκείμενα. Τεχνικές όπως η παρακολούθηση δραστηριότητας, η ομοιομορφική κρυπτογράφηση, ο ενσωματωμένος έλεγχος ταυτότητας και ο λεπτομερής έλεγχος πρόσβασης, εγγυώνται ασφάλεια των δεδομένων υγείας (Kaur et al., 2018) και πρέπει να λαμβάνονται υπόψη, κατά την ανάπτυξη συστημάτων Μηχανικής Μάθησης σε ΜΔΥ. Επίσης τεχνικές απόκρυψης ευαίσθητων δεδομένων, τα αντικαθιστούν με ρεαλιστικά φανταστικά, διατηρώντας την ακεραιότητα των στοιχείων αναφοράς, τη στατιστική τους ακρίβεια, προστατεύοντάς τα από κλοπή και διασφαλίζοντας παράλληλα τη συμμόρφωση με τους κανονισμούς και τις διεθνείς πολιτικές, που υπαγορεύουν το απόρρητο και τη μεταφορά δεδομένων.

Τα δεδομένα υγείας των στελεχών των ΕΔ, εκτός από ευαίσθητα και προσωπικά, όπως οποιουδήποτε άλλου ασθενούς, είναι ιδιαίτερος πολύτιμος για τον εχθρό, καθώς μπορούν κατόπιν κατάλληλης ανάλυσης και επεξεργασίας, να του παρέχουν χρήσιμες πληροφορίες, περί της μαχητικής ικανότητας του στρατεύματος, κατά ηλικία, γεωγραφική περιοχή, ειδικότητα, περιόδους έντασης και αντιπαλότητας κ.λπ. Αντιστρόφως, πιθανή διαρροή ή παραβίαση εχθρικών δεδομένων υγείας, μπορούν να οδηγήσουν, κατόπιν αναλύσεων, σε συμπεράσματα για την εχθρική μαχητική ισχύ, προς όφελος των ΕΔ μας.

Ενώ οι προσφερόμενες τεχνολογικές λύσεις για την επεξεργασία των ΜΔΥ προσφέρουν ευελιξία, αποτελεσματικότητα, εξασφάλιση υπολογιστικής ισχύος, εξοικονόμηση πόρων και χρόνου, εντούτοις οι απειλές και οι προκλήσεις που ελλοχεύουν, σε περίπτωση μη ασφαλούς εκμετάλλευσής τους, είναι σοβαρές. Έτσι στην περίπτωση της Υπολογιστικής Νέφους, υπάρχουν κίνδυνοι παραβίασης της εμπιστευτικότητας, της ακεραιότητας και της διαθεσιμότητάς τους, καθώς και της μη συμμόρφωσης των φορέων που επεξεργάζονται τα ΜΔΥ, με τις αρχές της λογοδοσίας και της διαφάνειας, που επιτάσσει ο ΓΚΠΔ.

Επομένως, η διασφάλιση των ΜΔΥ στο Νέφος αποτελεί μεγάλη πρόκληση ασφαλείας, όπου απαιτείται η χρησιμοποίηση ασφαλών αλγορίθμων αποθήκευσης ή προηγμένων αλγορίθμων κρυπτογράφησης. Ειδικότερα, στην τεχνολογία της Κινητής Υγείας (mHealth), η ιδιωτικότητα και η τήρηση αρχών ασφαλείας, είναι απαραίτητες προϋποθέσεις για την οικοδόμηση εμπιστοσύνης από πλευράς των χρηστών της. Η κρυπτογράφηση των δεδομένων της υγείας τους θεωρείται αρκετά ασφαλής πρακτική, ενώ η ανωνυμοποίηση των δεδομένων αυτών πρέπει να είναι η προτιμώμενη επιλογή, για την προστασία της ιδιωτικότητας των υποκειμένων, σε συνδυασμό με την ύπαρξη μέτρων λογοδοσίας και ελέγχου.

στ. Σημεία προσοχής στα δεδομένα εκπαίδευσης αλγορίθμων Μηχανικής Μάθησης (ML)

Στα σύνολα δεδομένων, που χρησιμοποιούνται για την εκπαίδευση των αλγορίθμων ML, πρέπει να αποφεύγεται η μεροληπτική χρήση τους, για μείωση του κινδύνου αποτυχίας των μεθόδων ML, ως προς την γενίκευση των αποτελεσμάτων, ενώ κατάρτιση μοντέλων ML από εξαιρετικά «θορυβώδεις» πηγές δεδομένων, σε κάποιες περιπτώσεις μπορεί να οδηγήσει ακόμη και σε ανεπιθύμητα αποτελέσματα (Waring et al., 2020).

Η εκπαίδευση των εργαλείων ML πρέπει να γίνεται με ικανό πλήθος πραγματικών (όχι «συνθετικών») δεδομένων ασθενών και με χρήση κατευθυντήριων οδηγιών (guidelines) έμπειρων ιατρών, διαφορετικά είναι πιθανό να οδηγηθούμε σε εξαγωγή ανακριβών και λανθασμένων συμπερασμάτων (Ross and Swetlitz, 2018). Η ανισορροπία που οφείλεται στο μικρό μέγεθος των μελετών σειράς των ασθενών (patient cohorts), μπορεί να προκαλέσει ανεπιθύμητη «υπερπροσαρμογή» (over-fitting) όχι μόνο στα κλασικά μοντέλα AI αλλά και στα σύγχρονα μοντέλα βαθιάς μάθησης (Peng, 2015).

Επειδή η σχέση μεταξύ της μεταβλητής στόχου, την οποία το μοντέλο προσπαθεί να προβλέψει, και των ανεξάρτητων μεταβλητών αλλάζει με το χρόνο με απρόβλεπτους τρόπους, δεν πρέπει να αγνοείται η έννοια της μετατόπισης (drift). Λόγω αυτής της μετατόπισης, το μοντέλο συνεχίζει να γίνεται ασταθές και οι προβλέψεις συνεχίζουν να γίνονται εσφαλμένες, με την πάροδο του χρόνου. Για το λόγο αυτό απαιτείται συνεχής παρακολούθηση, με μετρήσεις που σχετίζονται με τη σταθερότητα του μοντέλου, σε συνεχή διαστήματα (Shendre, 2020).

Αντί να εφαρμόζονται αδιάκριτα τεχνικές ML, σε οποιοδήποτε σύνολο δεδομένων, η προσήλωση στο προς απάντηση βιολογικό ερώτημα, που τίθεται κάθε φορά, καθώς και στα δεδομένα που απαιτούνται για την απάντησή του, είναι πρωταρχικής σημασίας (Toh et al., 2019). Οι ελεγκτικές αρχές θα πρέπει να είναι σε θέση να δοκιμάζουν και να πιστοποιούν τα δεδομένα που χρησιμοποιούνται από τους αλγορίθμους (EC, White Paper on AI, 2020).

ζ. Σημασία της Τεχνητής Νοημοσύνης και Ανάλυσης Δεδομένων για τα ΜΔΥ

Αυτό που κάνει την AI ιδιαίτερα σημαντική για την εκμετάλλευσή της στα δεδομένα υγείας, είναι η δυνατότητα της σταδιακής (incremental) μάθησης, στην οποία τα αποτελέσματα από ένα εκπαιδευμένο σύστημα AI ενσωματώνονται σε έναν κλειστό βρόχο ανατροφοδότησης δεδομένων και χρησιμοποιούνται για να τελειοποιήσουν την προγνωστική ακρίβεια του συστήματος, μέσω επαναληπτικής επανεκπαίδευσης του μοντέλου.

Η έρευνα κατέδειξε ότι η Προδιαγραφική Ανάλυση, ως συμπλήρωμα της Προγνωστικής, συμβάλει στην επίλυση προβλημάτων βελτιστοποίησης στο χώρο της υγείας, με τη χρήση δυνατοτήτων παρεχομένων από τις επιστήμες της στατιστικής και των μαθηματικών, αποτελώντας την τελική λύση υποστήριξης για ένα πλήρως Ευφυές Σύστημα (Intelligent System). Επίσης επιβεβαιώθηκε ότι η Τεχνητή Νοημοσύνη (AI), εκτός των άλλων, μπορεί να συνεισφέρει στη μείωση του διοικητικού φόρτου των φορέων υγειονομικής περίθαλψης, αφήνοντας έτσι περισσότερο πολύτιμο χρόνο στη διάθεση του ιατρικού και νοσηλευτικού προσωπικού, για την φροντίδα των ασθενών. Μαζί με άλλες ψηφιακές δυνατότητες και εργαλεία, τα σύγχρονα συστήματα υγείας μετασχηματίζονται και ενισχύονται δραστικά, με απώτερο σκοπό τη βελτίωση της υγείας και της ποιότητας των παρεχόμενων υγειονομικών υπηρεσιών στους ασθενείς.

Η βελτιστοποίηση των υπαρχόντων πόρων, μπορεί να επιτρέψει σε ένα νοσοκομείο να συντομεύσει τις υπάρχουσες γραμμές αναμονής, ενώ με την ενσωμάτωση ενός συστήματος ικανού να προσδιορίζει προβλήματα υγείας, που δεν μπορούν να επιλυθούν εκτός νοσοκομείου, αυξάνονται οι δυνατότητες παροχής θεραπείας, σε όσους πραγματικά τη χρειάζονται (Lopes et al., 2020). Ωστόσο η σχετική με την ΑΙ τεχνολογία πρέπει να λειτουργεί με εμπιστοσύνη για τους ανθρώπους σε μια δίκαιη και ανταγωνιστική οικονομία, με στόχο μια ανοικτή, δημοκρατική και βιώσιμη κοινωνία. Για την ενίσχυση εφαρμογής, αλλά και για την εποπτεία και διαφάνεια των συστημάτων ΑΙ, είναι αναγκαία η λειτουργία κατάλληλων νομικών κανόνων, ιδιαιτέρως σε συστήματα υψηλού κινδύνου όπως της υγείας (EC, White Paper on AI, 2020).

η. Αξιοποίηση των Ατομικών Ηλεκτρονικών Φακέλων Υγείας (ΑΗΦΥ)

Η επικράτηση των Μαζικών Δεδομένων και της ανάλυσής τους, προσφέρουν την ευκαιρία αξιοποίησης των Ατομικών Ηλεκτρονικών Φακέλων Υγείας ή ΑΗΦΥ (υποχρεωτικών πλέον στη χώρα μας με το Ν.4600/2019), παρέχοντας καλύτερα εργαλεία και περισσότερες πληροφορίες στον θεράποντα ιατρό, ως προς τη λήψη αποφάσεων και συνεπώς καλύτερη ποιότητα ιατρικών υπηρεσιών. Οι ΑΗΦΥ δύνανται να χρησιμοποιούν το «είδωλο υγείας», δηλαδή την εικονική αναπαράσταση με όλες τις σχετικές με την υγεία του υποκειμένου πληροφορίες και έξυπνους τρόπους διαχείρισης και πρόβλεψης της μελλοντικής του κατάστασης, το οποίο θα μπορεί να τροφοδοτεί με πληροφορίες, διάφορα ερευνητικά έργα, για τη δημιουργία νέων ακριβέστερων εργαλείων Ιατρικής Ακρίβειας (Prosperi et al., 2018).

θ. Σύγχρονα Συστήματα Υποστήριξης Κλινικών Αποφάσεων (CDSS)

Τα Συστήματα Υποστήριξης Κλινικών Αποφάσεων (CDSS) έχουν γίνει πλέον ισχυρότερα στους αλγόριθμους τους, πιο προσαρμόσιμα σε νέα δεδομένα, ενσωματωμένα πιο απρόσκοπτα στις υπάρχουσες τεχνολογίες και πιο αποτελεσματικά στην αντιμετώπιση ζητημάτων των ασθενών. Επίσης, μπορούν να χρησιμοποιούν μοντέλα πρόβλεψης για στοχευμένη διενέργεια θεραπειών ή αποφυγή της, με τα οποία αποφασίζεται αν η θεραπεία πρέπει να εφαρμοστεί σε άτομα με πολύ υψηλή πιθανότητα οφέλους, ή αν πρέπει να αποφευχθεί, για άτομα με μεγάλη πιθανότητα βλάβης. Τα CDSS δεν έρχονται να αντικαταστήσουν το ιατρικό και νοσηλευτικό προσωπικό, αντιθέτως έρχονται να υποστηρίξουν τη λήψη των αποφάσεών τους, ως προς την αντιμετώπιση ιατρικών περιστατικών, με ταχύτητα, ακρίβεια και αξιοπιστία, επιτρέποντας αναλύσεις και υπολογισμούς που ο ανθρώπινος νους, ακόμη και του πιο ικανού ιατρού ή νοσηλευτή, θα χρειαζόταν μήνες ή έτη για να επιτύχει.

ι. Προστασία δεδομένων υγείας: Δικαίωμα και Τροχοπέδη

Κάθε νοσηλευτικό ίδρυμα ενεργεί απλώς ως ο «θεματοφύλακας» των Μαζικών Δεδομένων Υγείας των ασθενών του, αφού τα δεδομένα αυτά ανήκουν στα υποκείμενα και η πρόσβαση και χρήση τους, εκτός του κλινικού χώρου, απαιτεί τη ρητή συγκατάθεση των ιδίων. Αυτό το νομικά κατοχυρωμένο δικαίωμα αποτελεί ταυτόχρονα τροχοπέδη στην ταχεία εκμετάλλευση του μεγάλου όγκου δεδομένων υγείας, που διατηρούνται ήδη σε κλινικά αρχεία. Για τεχνικές ΜΔΥ που χρησιμοποιούν χιλιάδες έως εκατομμύρια σημεία δεδομένων και που μπορεί να απαιτούν σημαντική επεξεργασία, η προοπτική απώλειας τέτοιου είδους πολύτιμων ανωνυμοποιημένων δεδομένων λόγω διαγραφής τους στο τέλος του έργου, είναι απογοητευτική για την πρόοδο της ιατρικής γνώσης (Hulsen et al., 2019).

ια. Επιστημονική έρευνα με δεδομένα υγείας

Η διενέργεια επιστημονικής έρευνας με δεδομένα υγείας ασθενών από τα αρχεία ενός νοσηλευτικού ιδρύματος, συνιστά νόμιμο σκοπό επεξεργασίας, επειδή όμως η πρόσβαση αυτή συνιστά επεξεργασία ειδικών κατηγοριών (ευαίσθητων) προσωπικών δεδομένων, θα πρέπει το νοσηλευτικό ίδρυμα να ζητήσει την άδεια της ΑΠΔΠΧ, για να επιτραπεί η πρόσβαση του ερευνητή στα εν λόγω αρχεία. Αν όμως, στο πλαίσιο επιστημονικής έρευνας λαμβάνει χώρα η δημοσίευση στοιχείων του ιατρικού φακέλου ασθενών, κατά τρόπο που να μην είναι αμέσως ή εμμέσως προσδιορίσιμη η ταυτότητά τους, αυτό δεν συνιστά επεξεργασία προσωπικών δεδομένων, που εμπίπτει στο ρυθμιστικό πεδίο του νόμου, για την προστασία τους.

10.2 Προτάσεις για Μελλοντική Έρευνα

α. Καλύτερη κατανόηση βιολογίας ασθενειών

Από την έρευνα προέκυψε ότι ένας σημαντικός περιορισμός της σημερινής ιατρικής είναι η κακή κατανόηση της βιολογίας των ασθενειών. Συγκεντρώνοντας τεράστιες ποσότητες Μαζικών Δεδομένων Υγείας (ΜΔΥ), όλες οι σχετικές μεταβλητές πολλαπλών πόρων, όπως δεδομένα DNA, RNA, πρωτεΐνης και μεταβολισμού θα μπορούν να ενσωματωθούν σε πιο ρεαλιστικά μοντέλα για να προβλέψουν πώς θα συμπεριφερθούν, για παράδειγμα, οι όγκοι και ποιοι ασθενείς θα ωφεληθούν καλύτερα από συγκεκριμένες, στοχευμένες θεραπείες (Willems et al., 2019).

β. Ολοκληρωμένη προσέγγιση της «Υγείας Ακριβείας»

Συγχώνευση της Ιατρικής Ακριβείας και της εφαρμογής της στη δημόσια υγεία μπορεί να οδηγήσει σε μια ολοκληρωμένη προσέγγιση της «Υγείας Ακριβείας», η οποία εκτιμάται ότι πρέπει να είναι βασική μελλοντική επιδίωξη, τόσο για το γενικό πληθυσμό όσο και για το στράτευμα ειδικότερα. Εργαλεία υποστήριξης αποφάσεων για ιατρούς και εργαζόμενους στον τομέα της δημόσιας υγείας, θα μπορούν να εφαρμόσουν και να αξιολογήσουν την αποτελεσματικότητα των προϊόντων ανάλυσης ακριβείας και να αντιμετωπίσουν δραστηκότερα τα ζητήματα υγείας, σε ατομικό ή κοινοτικό επίπεδο (Pearson et al., 2020).

γ. Ανακάλυψη νέων βιοδεικτών (biomarkers)

Η ανακάλυψη νέων βιοδεικτών στο μέλλον, μέσω την Ανάλυσης Μαζικών Δεδομένων Υγείας, αναμένεται να βελτιώσει την κατανόηση σχετικά με τους μηχανισμούς ασθένειας, να αυξήσει την αποτελεσματικότητα στην παροχή υγειονομικής περίθαλψης, να μειώσει το συνολικό κόστος για τον ασθενή, την οικογένεια, το νοσοκομείο και να διευκολύνει την υποστήριξη των κλινικών αποφάσεων (Zhang and Zhou, 2019), ενώ εκτιμάται ότι θα είναι δυνατή η χρησιμοποίηση νέων βιοδεικτών και ως εργαλείων ανάπτυξης φαρμάκων Ιατρικής Ακριβείας (Khan et al., 2019).

δ. Επιλογή παρεμβάσεων μέσω της πρόβλεψης

Σε πολλές περιπτώσεις, ένας προβλεπόμενος κίνδυνος υγείας δεν μπορεί να συνδυαστεί με μια παρέμβαση, που αποδεδειγμένα θα επιφέρει ευεργετικό αποτέλεσμα. Αυτό μπορεί να οδηγήσει σε μια κατάσταση στην οποία η απόπειρα παρέμβασης κάνει περισσότερο κακό παρά καλό στον ασθενή. Ενώ λοιπόν η πρόβλεψη μπορεί να χρησιμοποιηθεί, προς το παρόν, για να προτείνει την ανάγκη για προληπτική παρέμβαση με βάση το κίνδυνο που διατρέχει ο ασθενής, καθώς εξελίσσονται οι μέθοδοι εξαγωγής συμπερασμάτων, μπορούν να χρησιμοποιηθούν στο μέλλον, για την επιλογή των παρεμβάσεων (Pearson et al., 2020).

ε. Θέσπιση δεοντολογικών πλαισίων για αμερόληπτα μοντέλα

Για την αποφυγή χρήσης μεροληπτικών (biased) συνόλων δεδομένων, εκτιμάται ότι μελλοντικά θα πρέπει να θεσπιστούν δεοντολογικά πλαίσια από επιστημονικές επιτροπές και ρυθμιστικούς φορείς, για να αναγνωρίσουν και να ελαχιστοποιήσουν την επίδραση των μεροληπτικών μοντέλων, καθοδηγώντας παράλληλα τις σχεδιαστικές επιλογές, για εισαγωγή συστημάτων που οικοδομούν εμπιστοσύνη, κατανόηση και διατήρηση της ιδιωτικής ζωής (Toh et al., 2019).

στ. Ζητήματα ιδιωτικότητας και ασφάλειας δεδομένων στο ΙοΗΤ

Καθώς οι συσκευές ΙοΗΤ με τη χρήση κατάλληλων τεχνολογιών (GPS, RFID κ.λπ.) μπορούν να επικοινωνούν μεταξύ τους χωρίς τη μεσολάβηση του χρήστη, εγείρονται θέματα ιδιωτικότητας και ασφάλειας, ενώ προκύπτουν ερωτήματα που απαιτούν απαντήσεις, ενδεχομένως και τροποποιήσεις του νομοθετικού πλαισίου. Τέτοια ζητήματα αφορούν π.χ. ποιος συλλέγει κι ελέγχει τις πληροφορίες, σε ποια νομική βάση τις επεξεργάζεται, για πόσο χρόνο τις διατηρεί και πώς λαμβάνει τη συγκατάθεση του υποκειμένου, αφού οι εν λόγω συσκευές δεν διαθέτουν μηχανισμούς λήψης συναίνεσης (Κανέλλος, 2020).

11. Πίνακας Συγκριτικής Ανάλυσης Βασικών Βιβλιογραφικών Αναφορών

A/A	Συγγραφείς, Έτος	Τίτλος Άρθρου	Ερευνητικός Σκοπός	Μεθοδολογία	Κυριότερα Ευρήματα	Θέματα που Ερευνήθηκαν	Προτεινόμενες Λύσεις	Θέματα Ανοικτά για Μελλοντική Έρευνα
Ομάδα άρθρων: ANALYTICS								
1	Lopes et al., 2020	Predictive and Prescriptive Analytics in Healthcare: A Survey	Αναζήτηση της σημειωθείσας προόδου στον τομέα των Μαζικών Δεδομένων Υγείας, με τη χρήση Προγνωστικής και Προδιαγραφικής Ανάλυσης.	Έρευνα (survey) σχετική με τη χρήση Προγνωστικής και Προδιαγραφικής Ανάλυσης.	Η Προδιαγραφική Ανάλυση συμπληρώνει την Προγνωστική, προσφέροντας λύσεις, μέσω της επίλυσης προβλημάτων βελτιστοποίησης στο χώρο της υγείας, μέσω στατιστικής/μαθηματικών.	- Νοσοκομείο του μέλλοντος. - Καινοτόμες λύσεις στο χώρο της υγείας (σε διοικητικό και κλινικό επίπεδο). - Βελτιστοποίηση υπάρχοντων πόρων υγείας.	- Δημιουργία έξυπνου και δυναμικού περιβάλλοντος υγείας. - Υιοθέτηση Προγνωστικής Ανάλυσης - Συμπλήρωση της Προγνωστικής από την Προδιαγραφική Ανάλυση.	- Μοντελοποίηση αρχιτεκτονικής για την επίλυση θεμάτων στο χώρο της υγείας, σε διοικητικό και κλινικό επίπεδο.
2	Lepenioti et al., 2020	Prescriptive analytics: Literature review and research challenges	Έρευνα της σύγχρονης βιβλιογραφίας σχετικά με την Προδιαγραφική Ανάλυση και τις μεθόδους υλοποίησής της.	Ανάπτυξη τεκμηριωμένης γνώσης, μέσω συστηματικής αναθεώρησης.	Η Προδιαγραφική Ανάλυση μπορεί να βελτιώσει τη λήψη αποφάσεων και την αποτελεσματικότητα των διαδικασιών, βοηθώντας τους αναλυτές να προσεγγίσουν τα αποτελέσματα, συνδυάζοντάς τα με συγκεκριμένες καταστάσεις.	- Βελτίωση της Ανάλυσης Δεδομένων μέσω της Προγνωστικής Ανάλυσης. - Κάλυψη του κρίσιμου κενού στην Ανάλυση Δεδομένων των Επιχειρήσεων.	- Παροχή βελτιωμένων λύσεων στη διαδικασία λήψης απόφασης. - Χρήση κατανεμημένης επεξεργασίας και διαχείρισης δεδομένων.	- Συστήματα Ανάλυσης Δεδομένων της επόμενης γενιάς, ικανά να αναγνωρίζουν κινδύνους και πιθανά προβλήματα.
3	Zhang, 2020	Predictive analytics in the era of big data: opportunities and Challenges	Ανάδειξη της σημασίας της Προγνωστικής Ανάλυσης στα Μαζικά Δεδομένα Υγείας.	Αναφορά σε σύγχρονες επιστημονικές τάσεις.	Η Προγνωστική Ανάλυση αποτελεί τον ακρογωνιαίο λίθο της εξατομικευμένης φαρμακευτικής αγωγής και θεραπείας, ώστε ασθενείς με διαφορετικά κλινικά χαρακτηριστικά και γενετικά υπόβαθρα, να περιθάλπονται διαφορετικά.	- Διαμόρφωση κλινικού ερωτήματος για την Προγνωστική Ανάλυση Μαζικών Δεδομένων Υγείας. - Αντιμετώπιση εξειδικευμένων περιπτώσεων ασθενών (αγωγή, θεραπεία)	- Προγνωστική Ανάλυση τριών βημάτων (Διατύπωση Κλινικού Ερωτήματος, Σχεδίαση Μελέτης και Στατιστική Ανάλυση). - Χρήση Ιατρικής Ακρίβειας (Precision).	- Μελλοντικές δυνατότητες και ευκαιρίες χειρισμού των Μαζικών Δεδομένων, για μεγαλύτερη πρόοδο στον τομέα της υγείας.
4	Zhang and Zhou, 2019	Grand Challenges for Medtech Data Analytics	Καταγραφή των σημαντικότερων προκλήσεων στον χώρο της Ανάλυσης Δεδομένων Υγείας και πρόταση πιθανών λύσεων για την αντιμετώπισή τους.	Επισημάνση αναδυόμενων τεχνικών, που θα βοηθήσουν στην εξερεύνηση της Ανάλυσης σε εφαρμογές Μαζικών Δεδομένων Υγείας (MΔΥ), τόσο με τη χρήση παραδοσιακών τεχνικών επεξεργασίας σημάτων, όσο και μεθόδων τεχνητής νοημοσύνης.	- Καταγράφηκαν έξι βασικές προκλήσεις στον χώρο της Ανάλυσης Δεδομένων Υγείας και προτάθηκαν πιθανές λύσεις τους. - Η Ανάλυση Δεδομένων Υγείας, παρά τις προκλήσεις, συνεχίζει να εξελίσσεται ραγδαία καθημερινά.	- Ετερογένεια: τα δεδομένα υγείας είναι μείγμα δομημένων, ημι-δομημένων και μη δομημένων δεδομένων. - Χειρισμός των «Μεγάλων» Δεδομένων Υγείας. - Γενίκευση σαφούς «ορισμού» για την ΑΔΥ. - Μικρό μέγεθος μελετών σειράς των ασθενών - Κρίση αναπαραγωγιμότητας, έλλειψη αναλυτικών δεξιοτήτων που είναι απαραίτητες για την εξεύρεση σημασίας στα δεδομένα. - Απόρρητο και Ιδιοκτησία, των αρχείων των ασθενών.	- Συγχώνευση για ενσωμάτωση ετερογενών δεδομένων, προς δημιουργία αναλυτικών μοντέλων. - Διακομιστές υψηλών ταχυτήτων, με ενσωμάτωση κατανεμημένων υπολογιστικών συστημάτων, αλγορίθμων streaming ή υπολογιστικού νέφους. - Σαφής ορισμός της ΑΔΥ. - Χρήση πινάκων κόστους υπερ-δειγματοληψία, ανάλυση ευαισθησίας κ.ά - Αξιόπιστες τεχνικές επικύρωσης, υπερβαίνουν τις τρέχουσες τεχνικές στατιστικής επικύρωσης. - GDPR, αυστηρότερο νομικό πλαίσιο προστασίας, αξιόπιστες μέθοδοι κρυπτογράφησης.	Η Ανάλυση Δεδομένων Υγείας (ΑΔΥ) θα επιχειρήσει να ανακαλύψει νέους βιοδείκτες, να βελτιώσει την κατανόηση σχετικά με τους μηχανισμούς ασθένειας, να αυξήσει την αποτελεσματικότητα στην παροχή υγειονομικής περίθαλψης, να μειώσει το συνολικό κόστος για τον ασθενή, την οικογένεια, το νοσοκομείο και να διευκολύνει την υποστήριξη των κλινικών αποφάσεων.

5	Saheb and Izadi, 2019	Paradigm of IoT BDA (Big Data Analytics) in the healthcare industry: A review of scientific literature and mapping of research trends	Ανασκόπηση επιστημονικής βιβλιογραφίας και χαρτογράφηση ερευνητικών τάσεων, για την επίδραση της σχεδίασης, ανάπτυξης και υλοποίησης του IoT BDA στον τομέα της υγείας.	Συστηματική ποιοτική και ποσοτική βιβλιογραφική ανασκόπηση και βιομετρική ανάλυση.	Το πρότυπο IoT BDA συγκλίνει σε τρία κύρια επίπεδα: υπολογιστικής, αποθήκευσης (κρίσιμων και μη δεδομένων) και αφαίρεσης. Η σύγκλιση τους οδήγησε στην ανάπτυξη και πρόταση νέων συστημάτων και εφαρμογών IoT.	- Πώς συγκλίνουν οι τεχνολογίες IoT και BDA στη βιομηχανία της υγείας; - Ποιες εφαρμογές εμφανίζονται ως αποτέλεσμα του προτύπου IoT BDA;	- Το πρότυπο IoT BDA συγκλίνει σε τρία επίπεδα (υπολογισμού, συσώρευσης και αφαίρεσης). - Υπολογιστή Ομίχλης (Fog): κατάλληλη τεχνολογία για συστήματα υγείας αντιμετώπισης επειγόντων περιστατικών.	- Τα ευρήματα αυτής της μελέτης μπορούν να λειτουργήσουν ως βάση για μελλοντικές μελέτες, για τον προσδιορισμό νέων ερευνητικών ευκαιριών, στο IoT BDA.	
6	Stopa et al., 2019	Variance Reduction in Neurosurgical Practice: The Case for Analytics-Driven Decision Support in the Era of Big Data	Κατάδειξη της σημασίας της συμπερίληψης των Κατευθυντήριων Οδηγιών που Βασίζονται σε Αποδείξεις (Evidence Based Guidelines -EBG), στη διαδικασία λήψης κλινικής απόφασης.	Ανασκόπηση του υποβάθρου στις Κατευθυντήριες Οδηγίες που Βασίζονται σε Αποδείξεις (EBG) και στους Ηλεκτρονικούς Φακέλους Υγείας (ΗΦΥ), για την μείωση της διασποράς (variance reduction) στην νευροχειρουργική.	-Κρίθηκε πολύ χρήσιμη η συμπερίληψη των EBG, στους ΗΦΥ, για υποστήριξη της κλινικής απόφασης. -Οι εξελίξεις στα δεδομένα και στους αλγόριθμους AI, επιτρέπουν συνδυασμό ΜΔΥ και ταχύτατης επεξεργασίας τους, σε εργαλεία υποστήριξης κλινικών αποφάσεων.	- Μείωση διασποράς (variance reduction) μεταξύ παρόχων στη νευροχειρουργική, προς αποφυγή της αναποτελεσματικότητας και της κακής έκβασης της πορείας των ασθενών.	- Συμπερίληψη των EBG, στους ΗΦΥ, για μείωση λαθών, αύξηση της ασφάλειας των ασθενών και υποστήριξη της κλινικής απόφασης,	- Αξιοποίηση των δυνατοτήτων Ανάλυσης των Δεδομένων των Ηλεκτρονικών Φακέλων Υγείας, για παροχή στο ιατρικό προσωπικό, καλύτερων εργαλείων λήψης απόφασης.	
Ομάδα άρθρων: BIG DATA									
7	Pramanik et al., 2020	Healthcare informatics and analytics in big data	Κριτική ανάλυση των προηγμένων τεχνολογιών Πληροφορικής και Ανάλυσης Δεδομένων Υγείας (HCI&A) και παροχή οδικού χάρτη για προώθηση μελλοντικής έρευνας και μεταφοράς γνώσεων, της επόμενης γενιάς (HCI&A).	Βιβλιογραφική έρευνα στον τομέα των Μαζικών Δεδομένων, με έμφαση σε 4 τμήματα: υπάρχουσες τεχνολογίες, εφαρμογές συστημάτων, αξιολογήσεις συστημάτων και περιοχές ερευνών.	- HCI&A 1.0: αξιοποίηση τεχνολογιών σχεσιακών βάσεων δεδομένων. -HCI&A 2.0: υποστήριξη «συνεργατικής υγειονομικής περιθάλψης» συνδέοντας χρήστες, μέσω κοινωνικών δικτύων. -HCI&A 3.0: υιοθέτηση μεγάλης κλίμακας εργαλείων διαχείρισης κι ανάλυσης δεδομένων υγείας.	Παρουσίαση μιας συνολικής σύνοψης της Πληροφορικής και Ανάλυσης Δεδομένων Υγείας (HCI&A), υπογραμμίζοντας τις πολλές προκλήσεις και ευκαιρίες της.	- Παρουσίαση εξέλιξης των γενιών HCI&A, από την αρχή τους έως σήμερα. - Βασικές τεχνολογίες, για σύγχρονες μορφές HCI&A. - Καταγραφή ερευνητικών τάσεων σε ακαδημαϊκό και βιομηχανικό τομέα. - Κατάσταση σημερινών συστημάτων ακαδημαϊκής εκπαίδευσης, για κάλυψη αναγκών HCI&A.	Οδικός χάρτης για προώθηση μελλοντικής έρευνας και μεταφοράς γνώσεων της επόμενης γενιάς HCI & A, μέσω της μελέτης προηγμένων τεχνολογιών πληροφορικής και ανάλυσης Μαζικών Δεδομένων Υγείας.	
8	Aceto et al., 2020	Industry 4.0 and Health: Internet of Things, Big Data, and Cloud Computing for Healthcare 4.0 (HC4.0)	Χρήση της μελέτης αυτής ως αναφοράς για εξειδικευμένους ερευνητές και επαγγελματίες ΤΠΕ, για την αντιμετώπιση αναγκών στον τομέα της υγειονομικής περίθαλψης και των πληροφοριακών της συστημάτων, στο πλαίσιο των εξελίξεων του HC4.0.	Συστηματική ανασκόπηση σχετικής βιβλιογραφίας, χρησιμοποιώντας την προσέγγιση χιονοστιβάδας (snowballing), ως τρόπου αναζήτησης βιβλιογραφίας.	Η ηλεκτρονική υγεία πλέον υποστηρίζεται από τις ΤΠΕ, και θα μετασηματιστεί περαιτέρω στο νέο πλαίσιο του HC4.0 με νέες προσεγγίσεις, διαδικασίες και εφαρμογές, όπως τα ενισχυμένα περιβάλλοντα διαβίωσης, η αποκατάσταση στο σπίτι και η εξατομικευμένη υγειονομική περίθαλψη.	- Πώς διαμορφώνεται το όραμα της υγειονομικής περιθάλψης του μέλλοντος, μέσω των τεχνολογικών δυνατοτήτων και των καινοτομιών του HC4.0; - Ποιες καινοτόμες τεχνολογίες οδηγούν το όραμα αυτό; - Πώς θα χρησιμοποιηθούν τα ΜΔΥ για εξαγωγή πολύτιμης γνώσης, στο πλαίσιο του HC4.0;	- Διαμόρφωση υπηρεσίας υγειονομικής περιθάλψης πανταχού παρούσας, διαρκώς διαθέσιμης, προσφέροντας εξατομικευμένες ιατρικές υπηρεσίες. - Ειδικές φορητές συσκευές, αισθητήρες νανοκλίμακας, ασύρματα δίκτυα. - Χρήση εργαλείων - πλατφορμών ανάλυσης, και τεχνικών AI.	Συλλογή μεταδεδομένων μικρο-επιπέδου και μακρο-επιπέδου από έξυπνα φαρμακευτικά προϊόντα, με σκοπό να παρέχουν νέες πληροφορίες για τις ασθένειες, να βοηθήσουν στο σχεδιασμό υπηρεσιών, και να διευκολύνουν την εξατομικευμένη υγειονομική περίθαλψη.	
9	Mehtaa et al., 2019	Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study	Μελέτη της εξάπλωσης της εφαρμογής της Ανάλυσης Μαζικών Δεδομένων και της Τεχνητής Νοημοσύνης στον χώρο της υγείας.	Συστηματική μελέτη - αξιολόγηση 2.421 άρθρων σχετικών με το θέμα, που δημοσιεύτηκαν μεταξύ 2013 και 2019.	Η Προγνωστική Ανάλυση, η Μηχανική Μάθηση και η Τεχνητή Νοημοσύνη στα Μαζικά Δεδομένα Υγείας, καθιστούν δυνατή την ταυτοποίηση μοτίβων (patterns) και συσχετισμών (correlations), με πρακτικές ιδέες για τη βελτίωση της παρεχόμενης υγειονομικής περίθαλψης.	Τι είδους ερευνητικές μελέτες έχουν διεξαχθεί στα ΜΔΥ και στην AI, ποιες είναι οι συνεισφορές των μελετών αυτών, σε ποιους τομείς και διαστάσεις εστιάζουν, πώς εξελίχθηκε η έρευνα στην Ανάλυση Μαζικών Δεδομένων και στην AI στο χώρο της υγειονομικής περίθαλψης;	Κυρίως χρησιμοποιήθηκε έρευνα αξιολόγησης, οι περισσότερες πρωτογενείς μελέτες παρουσίασαν μοντέλα ML, ενώ η φροντίδα των ασθενών συγκεντρώνει το μεγαλύτερο ερευνητικό ενδιαφέρον. Κυρίως δίνεται έμφαση στη διαμόρφωση της κλινικής ειδικότητας, ενώ διαφαίνεται εντυπωσιακή αύξηση του ερευνητικού ενδιαφέροντος.	Ανάγκη για πιο ουσιαστικές μελέτες περιπτώσεων και εργασίες εμπειρίας, σχετικά με τη χρήση της Ανάλυσης Μαζικών Δεδομένων και της Τεχνητής Νοημοσύνης στο χώρο της υγειονομικής περιθάλψης στον πραγματικό κόσμο.	

10	Willems et al., 2019	The potential use of big data in oncology	Κατάδειξη των δυνατοτήτων των BD στην ογκολογία, καθώς και της ανάγκης συνδυασμού Μαζικών Δεδομένων κλινικών, παθολογικών και ποιότητας ζωής.	Ερμηνεία-σχολιασμός των δυνατοτήτων των BD στην ογκολογία.	-Σημαντική πηγή BD υγείας αποτελούν τα δεδομένα PROM και PREM. -Συνδυάζοντας μεγάλα σύνολα δεδομένων γονιδιώματος και περιβαλλοντικά, είναι δυνατή η πρόγνωση σοβαρών ασθενειών.	- Η κακή κατανόηση της βιολογίας των ασθενειών, ως περιορισμός της σημερινής ιατρικής. - Συνδυασμός ΜΔΥ με μεγάλα σύνολα δεδομένων γονιδιωματικής ή περιβαλλοντολογικής προέλευσης,	- Ενσωματωμένα δεδομένα (omics) θα μπορούν να παρέχουν ολοκληρωμένη εικόνα της βιολογικής συμπεριφοράς καρκινικών όγκων, των μηχανισμών που διέπουν την ανάπτυξή τους, το μεταστατικό δυναμικό τους και την ανταπόκρισή τους σε στοχευμένες - Πρόγνωση σοβαρών ασθενειών όπως ο καρκίνος, τόσο για μεμονωμένους ασθενείς όσο και για ομάδες.	- Δημιουργία οργανικών, αποκεντρωμένων, εικονικών δικτύων, αντί κεντρικών ΒΔ. - Γνώσεις που προέρχονται από ΜΔΥ, θα χρειαστεί να μεταφραστούν σε «Μικρά Δεδομένα» των μεμονωμένων περιθαλπόμενων ασθενών.	
11	Tao et al., 2019	Economic perspective analysis of protecting big data security and privacy	Παροχή οικονομικής αιτιολόγησης των τεχνικών αποφάσεων, που λαμβάνονται για την προστασία των Μαζικών Δεδομένων και του κόστους που δαπανούν οι οργανισμοί για αυτά.	Διερεύνηση των οικονομικών προοπτικών της ασφάλειας Μαζικών Δεδομένων και της ιδιωτικότητας, για την προστασία τους με ασφαλή και αποτελεσματικότερο τρόπο.	Η βιομηχανία υγειονομικής περίθαλψης αποτελεί τον μεγαλύτερο στόχο επιθέσεων ransomware, και οι υπεύθυνοι των φορέων υγείας είναι συνήθως αναγκασμένοι να συμμορφωθούν με τις απαιτήσεις του επιτιθέμενου.	Μελέτη των οικονομικών προοπτικών για την προστασία των Μαζικών Δεδομένων.	- Ασφάλειας και απορρήτου. - Επενδυτικών αποφάσεων για την ασφάλεια και ιδιωτικότητα. - Καταπολέμησης των εγκλημάτων κυβερνοχώρου. - Κυβερνητικών κανονισμών περί προστασίας	Η διερεύνηση των παρουσιαζόμενων τομέων, απαιτεί περαιτέρω λεπτομερή αναλυτικά αποτελέσματα και εργαλεία, στο μέλλον.	
12	Kaur et al., 2018	Big Data (BD) and Machine Learning (ML) Based Secure Healthcare Framework	Έλεγχος της επίδρασης της ML και των BDA στη βιομηχανία της υγείας, με βαρύτητα στην ασφάλεια των δεδομένων.	- Εμπειρική μελέτη για την ανάλυση του ρόλου των BD στη βιομηχανία της υγείας. -Σχεδίαση εφευρετικού και ασφαλούς υβριδικού μοντέλου υγειονομικής περίθαλψης 4 επιπέδων.	- Η ακρίβεια στη διάγνωση ασθενειών μπορεί να ενισχυθεί περαιτέρω, ενσωματώνοντας τις έννοιες της λογικής fuzzy και της θεωρίας των πληροφοριών. - Διαπιστώθηκε ότι κατά τη χρήση ML και BDA στη διάγνωση ασθενειών, δεν δινόταν η δέουσα σημασία στην ιδιωτικότητα και προστασία των δεδομένων.	- Τεχνικές που εγγυώνται ασφάλεια των δεδομένων υγείας. - Οι 4 πυλώνες στους οποίους βασίζονται τα Μαζικά Δεδομένα Υγείας.	- Απόκρυψη δεδομένων, παρακολούθηση δραστηριότητας, ομοιομορφική κρυπτογράφηση, ενσωματωμένος έλεγχος ταυτότητας και λεπτομερής έλεγχος πρόσβασης. - Περίθαλψη με επίκεντρο τον ασθενή, προγνωστική ανάλυση ασθενειών, παρακολούθηση ασθενών σε πραγματικό χρόνο, βελτίωση των μεθόδων θεραπείας.	Βελτίωση του προτεινόμενου μοντέλου, ώστε να μπορεί να λειτουργεί σε διάφορες φορητές συσκευές, με τεχνικές εμπνευσμένες από τη φύση και με χρήση υβριδισμού, για μεγαλύτερη ακρίβεια του διαγνωστικού του πλαισίου.	
Ομάδα άρθρων: PRECISION MEDICINE									
13	Song et al., 2020	Big data-driven precision medicine: Starting the custom-made era of iatrolology	Αναζήτηση και επισήμανση των πλεονεκτημάτων της Ιατρικής Ακρίβειας στη διάγνωση, τη θεραπεία και την πρόληψη συστηματικών και καρκινικών νόσων.	Έρευνα για την επισήμανση των πλεονεκτημάτων της Ιατρικής Ακρίβειας, σε οκτώ συστηματικές ασθένειες του ανθρώπινου σώματος και διάφορες μορφές καρκίνων.	-Η εφαρμογή της Ιατρικής Ακρίβειας (IA) βελτιώνει την αποτελεσματικότητα της διάγνωσης και της θεραπείας της νόσου και κάνει δυνατή την εξατομικευμένη θεραπεία για τους ασθενείς. -Αποτελεσματική αλληλεπίδραση δυτικής και παραδοσιακής κινεζικής ιατρικής, στο πλαίσιο της IA.	Πώς μπορεί να γίνει ακριβής και στοχευμένη διάγνωση, θεραπεία και πρόληψη συστηματικών ασθενειών και καρκίνων, προσωποποιημένη για κάθε ασθενή.	Εισαγωγή εφαρμογής της Ιατρικής Ακρίβειας (IA) σε οκτώ συστηματικές ασθένειες του ανθρώπινου σώματος και καρκίνους, επισημαίνοντας τα πλεονεκτήματα της IA στη διάγνωση, τη θεραπεία και την πρόληψή τους.	Διεύρυνση του πλήθους των δικαιούχων Ιατρικής Ακρίβειας (IA), επίλυση προβλημάτων στο πεδίο των omics, τυποποίηση και κανονικοποίηση. Ραγδαία ανάπτυξη της IA στο μέλλον, με μείωση κόστους ανάλυσης αλληλουχίας γονιδιώματος (genome sequencing).	
14	Pearson et al., 2020	Precision Health Analytics with Predictive Analytics and Implementation Research	Εξερεύνηση ευκαιριών και προκλήσεων από τη χρήση Προγνωστικής Ανάλυσης στα άτομα και στον πληθυσμό, καθώς και στα αντίστοιχα επίπεδα/τομείς της Ιατρικής Ακρίβειας.	Ανασκόπηση με χρήση τεχνολογίας αιχμής (state-of-the-art) – Έρευνα υλοποίησης.	-Η Προγνωστική Ανάλυση επεκτείνει την ποιότητα και ποσότητα δεδομένων που απαιτούνται για την ενημέρωση της περίθαλψης φροντίδα ασθενούς. -Νέα εργαλεία λήψης αποφάσεων απαιτούν	Προϋποθέσεις του κύκλου επεξεργασίας, μέσω της ανάλυσης των δεδομένων, που απαιτείται για την έγκαιρη πρόγνωση του κινδύνου που διατρέχει ένας ασθενής.	- Ψηφιακές πληροφορίες σε πραγματικό χρόνο, από διάφορες πηγές ΜΔΥ. - Αποτελεσματική ανάλυση των πληροφοριών αυτών. - Χρησιμοποίηση κατάλληλων εργαλείων για υποστήριξη αποφάσεων και	Ενώ η πρόβλεψη μπορεί να χρησιμοποιηθεί, προς το παρόν, για να προτείνει την ανάγκη για προληπτική παρέμβαση με βάση το κίνδυνο που διατρέχει ο ασθενής, καθώς εξελίσσονται οι μέθοδοι εξαγωγής συμπερασμάτων, μπορούν να	

					έρευνα, για καθορισμό χρήσης των πλεονεκτημάτων και της βιωσιμότητάς τους.		καθοδήγηση. - Παρουσίαση πληροφοριών για βελτίωση της περίθαλψης.	χρησιμοποιηθούν στο μέλλον, για την επιλογή των παρεμβάσεων.
15	Capobianco, 2020	Imprecise Data and Their Impact on Translational Research in Medicine	Η εξέταση εννοιών όπως η Μεταγραφική Ιατρική και τα ανακριβή δεδομένα υγείας.	Αξιοποίηση των διαστάσεων των δεδομένων υγείας, σε σχέση με υπάρχουσες ανεκπλήρωτες ανάγκες στην κλινική πρακτική.	Η Ιατρική προσδοκά από τα Μαζικά Δεδομένα Υγείας: την ικανότητα δημιουργίας προγνωστικών μοντέλων και τη δυνατότητα εφαρμογής τους, ώστε να ληφθούν ακριβή προφίλ κινδύνου ασθενών ή/και να ακολουθηθούν σωστές θεραπευτικές οδοί.	Προκλήσεις σε σχέση με τα δεδομένα υγείας, που πρέπει να αντιμετωπιστούν, τόσο στον γενικό πληθυσμό, όσο και σε εξατομικευμένες περιπτώσεις ασθενών, όπως καταδεικνύεται από το παράδειγμα της Ιατρικής Ακριβείας.	- Ανάγκη συγκέντρωσης, ολοκλήρωσης, μοντελοποίησης και ερμηνείας των δεδομένων από πηγές πληροφοριών και κλίμακες, που επηρεάζουν δυναμικά τις ασθένειες. - Τα προβλήματα στα δεδομένα συχνά απαιτούν υπολογιστική επεξεργασία μέσω λύσεων, χωρίς τις οποίες θα παρέμεναν δυσεπίλυτα.	Η διαλειτουργικότητα μοντέλων, ώστε παρόμοια δεδομένα να μπορούν να χρησιμοποιηθούν για μάθηση σε εργασίες γενίκευσης, με εντοπισμό στοιχείων που έχουν τα μοντέλα από κοινού και ο συνδυασμός τους σε ένα «μοντέλο - πυρήνα».
16	Khan et al., 2019	Unbiased data analytic strategies to improve biomarker discovery in Precision Medicine	Σχεδιασμός μελέτης, εκτιμήσεις ανάλυσης και προτεινόμενα πρότυπα λήψης αμερόληπτων αποφάσεων στην ανακάλυψη βιοδεικτών, για χρήση τους στην Ιατρική Ακριβείας.	Χρήση της αμερόληπτης (unbiased) προσέγγισης, όπου υιοθετείται το κατάλληλο πρωτόκολλο τυχαιοποίησης, έτσι ώστε η επιλεγμένη μεταβλητή να έχει πολύ καλή πιθανότητα να αποτελεί ακριβή εκπρόσωπο του στοχευμένου πληθυσμού.	Για την ανακάλυψη αναπαραγωγικών βιοδεικτών απαιτούνται: σωστός σχεδιασμός μελέτης, αμερόληπτη επεξεργασία δεδομένων, αναλύσεις ποιοτικού ελέγχου και μελετημένη εφαρμογή στατιστικής και αλγορίθμων Μηχανικής Μάθησης.	-Προ-επεξεργασία Μαζικών Δεδομένων Υγείας. -Στρατηγικές ανακάλυψης βιοδεικτών. -Εργαλεία Ανάλυσης Δεδομένων, για ανακάλυψη βιοδεικτών.	-Καθαρισμός δεδομένων (data cleaning), φιλτράρισμα και κανονικοποίηση. -Ποσοτικές και Ποιοτικές στρατηγικές. -Χρήση Στατιστικής και Μηχανικής Μάθησης.	Χρησιμοποίηση βιοδεικτών ως εργαλείων ανάπτυξης φαρμάκων Ιατρικής Ακριβείας.
17	Hulsen et al., 2019	From Big Data to Precision Medicine	Αναζήτηση προοπτικής για το πώς μπορεί να σημειωθεί πρόοδος στη χρήση Μαζικών Δεδομένων, ώστε να φτάσουμε σε επιτυχημένες στρατηγικές Ιατρικής Ακριβείας.	Εντοπισμός βασικών εννοιολογικών και δομικών προκλήσεων στις στρατηγικές Ιατρικής Ακριβείας, με συγκεκριμένα παραδείγματα.	-Διαπιστώθηκε η ανάγκη τυποποίησης του περιεχομένου δεδομένων, της μορφής τους και των κλινικών ορισμών, η ανάγκη συνεργατικών δικτύων με κοινή χρήση δεδομένων και εξειδίκευση. -Η έρευνα «κατασκευασμένης υποθέσης» συμπληρώνει την παραδοσιακή επιστήμη, που βασίζεται στην υπόθεση.	Συμβολή της Τεχνητής Νοημοσύνης και της Μηχανικής Μάθησης στις κλινικές απεικονιστικές εξετάσεις.	- Η ML εντοπίζει μη αναγνωρισμένα, μοναδικά χαρακτηριστικά εικόνες ή συνδυασμούς τους, που σχετίζονται με την εξέλιξη της νόσου. - Μετά εκπαίδευση, οι τεχνικές ML μπορούν να παρέχουν pre-screening εικόνων, προς αναζήτηση περιστατικών με μεγάλη πιθανότητα ασθένειας για ιεράρχησή τους.	Υπό τους περιορισμούς του ΓΚΠΔ, προσπάθεια από την επιστήμη των δεδομένων για εξισορρόπηση δύο ευθυνών: 1) της χρήσης των δεδομένων των ασθενών με κατάλληλη συγκατάθεσή τους και 2) της μεγιστοποίησης της ικανότητας βελτίωσης της υγείας τους.
18	Prosperi et al., 2018	Big data hurdles in precision medicine and precision public health	Διερεύνηση της παρούσας κατάστασης σχετικά με την Ιατρική Ακριβείας και τη Δημόσια Υγεία, για τη βελτίωση της υγείας, των αποτελεσμάτων της και για τη μείωση των ανισοτήτων στο χώρο της.	Ανάλυση τεχνικών και κοινωνικών εμποδίων που σχετίζονται με την ανάπτυξη μοντέλων πρόβλεψης κινδύνων για την υγεία, διαγνώσεων και αποτελεσμάτων από ολοκληρωμένες βιοϊατρικές βάσεις δεδομένων.	Η επιστήμη δεδομένων για την Ιατρική Ακριβείας και τη Δημόσια Υγεία, εγγυάται μια τυποποίηση βασισμένη στην Πληροφορική, σχεδιασμού μελέτης και διαλειτουργικότητας σε όλα τα επίπεδα της διαδικασίας συμπερασμάτων γνώσης, από την ερευνητική σημασιολογία, έως την ανάπτυξη μοντέλων και τελικά στην εφαρμογή.	Πρόταση χρήσης μοντέλου σύγχρονου «ειδώλου υγείας» (“health avatar), στην εποχή των ΜΔΥ των ασθενών.	-Το «είδωλο υγείας» πρέπει να συνδέει όλους τους σχετικούς με την υγεία τύπους δεδομένων, ακόμη και προερχόμενους εκτός υγειονομικής περίθαλψης. - Εργαλεία ΑΙ θα μπορούν να επεξεργαστούν εικόνες από αναρτήσεις του υποκειμένου στα κοινωνικά δίκτυα, για διατροφικές του συνήθειες. - Χρήση πληροφοριών γεωεντοπισμού, για έκθεσή του ασθενούς σε αλλεργιογόνα και ρύπους.	- Τα μοντέλα πρόβλεψης μελλοντικών καταστάσεων υγείας δεν είναι ακόμη ακριβή και η δυνατότητα δράσης τους δεν είναι ακόμη αξιόπιστη. - Θα απαιτηθεί ενσωμάτωση στην Ιατρική Ακριβείας συναφών θεμάτων ηθικής και χάραξης πολιτικής.

Ομάδα άρθρων: IoT

19	Alam και Rahmani, 2020	Intelligent context-based healthcare metadata aggregator in internet of medical things (IoMT) platform	Πρόταση ενσωμάτωσης του έξυπνου αθροιστή μεταδεδομένων υγείας από πλατφόρμες IoMT, με συστήματα λήψης απόφασης, για έγκαιρη διάγνωση ασθενειών.	Μελέτη περίπτωσης χρήσης του έξυπνου αθροιστή μεταδεδομένων υγείας, στην έγκαιρη και αυτόματη διάγνωση της σήψης.	Η κατάλληλη χρήση του νέου έξυπνου αθροιστή μεταδεδομένων, υγείας σε αποκεντρωμένες και καταναμημένες πλατφόρμες IoMT, μπορεί να συμβάλει σε διάφορους ιατρικούς τομείς και να συνδυαστεί με συστήματα λήψης ιατρικής απόφασης.	Μελέτη των διαφόρων τύπων μεταδεδομένων στο χώρο της υγείας, για κατανόηση της ποικιλομορφίας και της σημασίας τους, για τον έξυπνο αθροιστή.	Μεταδεδομένα υγείας σχετιζόμενα με: -Ασθενείς, παρόχους υγείας, φαρμακευτικές αγωγές, πληρωτές υπηρεσιών υγείας, υπηρεσίες διακυβέρνησης, παρόχους υπηρεσιών δεδομένων υγείας, ιατρική έρευνα και κατασκευαστές ιατρικών συσκευών.	Υλοποίηση του προτεινόμενου μοντέλου και κατάδειξη της πρακτικής του χρησιμότητας, κατά την ενσωμάτωσή του σε πλατφόρμες IoMT.
20	Mrozek et al., 2020	Fall detection in older adults with mobile IoT devices and machine learning in the cloud and on the edge	Διερεύνηση της δυνατότητας έγκαιρης ανίχνευσης περιστατικών πτώσης ηλικιωμένων, μέσω συσκευών IoT, για την παροχή υποστήριξής τους.	Παρουσίαση επεκτάσιμης αρχιτεκτονικής παρακολούθησης και ανίχνευσης περιστατικών πτώσης και επικύρωση μοντέλων Μηχανικής Μάθησης για την διαδικασία εντοπισμού τους.	Η απομακρυσμένη επιτήρηση ευπαθών ομάδων μεγάλης κλίμακας, απαιτεί όχι μόνο οι τοπικές συσκευές IoT να μπορούν να εντοπίζουν επικίνδυνες καταστάσεις, αλλά και τα κέντρα δεδομένων να παρέχουν σημαντικό χώρο αποθήκευσης και αυξημένες υπολογιστικές δυνατότητες.	-Δυνατότητες φορητών συσκευών IoT, να εκτελούν επεξεργασία και ανάλυση δεδομένων από αισθητήρα, σε τεχνολογία Edge. -Δυνατότητα μετριασμού εμποδίων της τεχνολογίας των Μαζικών Δεδομένων.	-Η αρχιτεκτονική συστήματος Whoops λειτουργεί σε περιβάλλον νέφους Azure, ανιχνεύοντας πτώσεις από κινητή συσκευή, κοντά στο άτομο που παρακολουθείται. -Μείωση της ποσότητας δεδομένων και της κίνησης του δικτύου.	Εξελίξεις στον τομέα της τηλεϊατρικής σε ανεπτυγμένες και νεο-βιομηχανοποιημένες χώρες, από τη χρήση τεχνολογιών όπως της Υπολογιστικής Νέφους, του IoT, των Big Data και της τεχνολογίας Edge.
21	Mani et al., 2020	An IoT Guided Healthcare Monitoring System for Managing Real-Time Notifications by Fog Computing Services	Κατάδειξη του ρόλου, που το IoT διαδραματίζει στην προσφορά υπηρεσιών υγείας.	Έρευνα-Σύγκριση τεχνολογιών IoT στον τομέα της Υγείας.	Η Υπολογιστική Ομίχλης (Fog) λειτουργεί μεταξύ συσκευών Νέφους (Cloud), με συνέπειες: διαθεσιμότητα επιπλέον μονάδας επεξεργασίας, μείωση καθυστέρησης, αξιοπιστία, ενεργειακή απόδοση και ιδιωτικότητα του συστήματος.	Κίνητρα, συμβολή και σημασία τεχνολογιών IoT, στον τομέα της Υγείας.	Παροχή συνεχούς επιτήρησης, οικοδόμηση εμπιστοσύνης, απομακρυσμένη παρακολούθηση ασθενών, μείωση κόστους, ειδοποιήσεις έκτακτης ανάγκης, γνώσης από τα δεδομένα, απομακρυσμένη ιατρική βοήθεια, παρακολούθηση προσωπικού, απογραφής και ασθενών, διαχείριση φαρμακευτικής αγωγής.	Χρήση φορητών έξυπνων ρολογιών που εκτελούν εργασίες φορητών υπολογιστών και κινητών τηλεφώνων, διευκολύνοντας τη ζωή των ανθρώπων, ως εναλλακτικές λύσεις.
22	Lee, 2019	IoT for enterprises: An ecosystem, architecture and IoT service business model	Η κάλυψη του ερευνητικού κενού που διαπιστώνεται στην έρευνα σχετικά με το επιχειρηματικό IoT και η παρακίνηση ενδιαφέροντος, για περαιτέρω αναζήτησή.	Παρουσίαση ενός οικοσυστήματος IoT, μιας αρχιτεκτονικής IoT και ενός επιχειρηματικού μοντέλου IoT, για την ορθή επιλογή και ανάπτυξη υπηρεσιών IoT (σε επιχειρηματικούς τομείς, μεταξύ των οποίων και αυτός της υγείας).	Ένα επιχειρηματικό IoT περιλαμβάνει: -5 επίπεδα: αντίληψης, δικτύου, επεξεργασίας, εφαρμογής, διαχείρισης υπηρεσιών. - 5 εμπλεκόμενους φορείς: δημιουργούς πλατφόρμας λογισμικού, κατασκευαστές πλατφόρμας εξοπλισμού, τεχνολόγους δικτυακών υποδομών, προγραμματιστές εφαρμογών και πελάτες.	Βασικά δομικά στοιχεία που πρέπει να περιλαμβάνει ένα επιχειρηματικό μοντέλο IoT	- Προτεινόμενη αξία, όταν νέες υπηρεσίες γίνονται αντιληπτές από χρήστες. -Δραστηριότητες δικτύου, η βελτίωση των οποίων πρέπει να επιδιώκεται. - Πόροι, π.χ. υπαλλήλους, τεχνολογίες, προϊόντα, υπηρεσίες, εγκαταστάσεις. - Βιωσιμότητα, σε μια προσπάθεια επίτευξης μακροπρόθεσμης εταιρικής ανάπτυξης και κερδοφορίας	Νέες διαστάσεις και ευκαιρίες, για μελλοντικές γενιές διαφορετικών υπηρεσιών, από τη συγχώνευση Μαζικών Δεδομένων και IoT, μέσω πλατφορμών δεδομένων IoT, σε τομείς όπως η υγειονομική περίθαλψη.
23	Kumar et al., 2018	Cloud and IoT based disease prediction and diagnosis system for healthcare using Fuzzy neural classifier	Αναζήτηση τρόπων προσφοράς καλύτερων υπηρεσιών παρακολούθησης και πρόγνωσης για τους ασθενείς, με τη χρήση των τεχνολογιών IoT και Cloud Computing.	Πρόταση χρήσης συστήματος 3 φάσεων: συλλογής δεδομένων (μέσω IoT συσκευών, αποθετηρίου UCI και ιατρικών φακέλων), ασφαλούς αποθήκευσης στο νέφος και πρόγνωσης-διάγνωσης σοβαρής ασθένειας (με χρήση κανόνων fuzzy).	Ο αναλυτής σοβαρότητας (severity analyser), μπορεί να χρησιμοποιηθεί για τον εντοπισμό της σοβαρότητας της ασθένειας, με χρήση κανόνων fuzzy, σύμφωνα με τις πληροφορίες που παρέχονται από τα ιατρικά αρχεία.	Διασφάλιση των Μαζικών Δεδομένων Υγείας στο Νέφος, ως πρόκλησης ασφαλείας, με χρήση ασφαλών αλγορίθμων αποθήκευσης δεδομένων ή προηγμένων αλγορίθμων κρυπτογράφησης τους.	Μηχανισμός αποθήκευσης, που συλλέγει δεδομένα από μονάδα συλλογής, τα οποία αποθηκεύονται σε πέντε στάδια: δλδ αποθήκευση, ανάκτηση, συγκέντρωση, διαμερισματοποίηση και συγχώνευση, ενώ τα «ασφαλή» ιατρικά δεδομένα μπορούν να αποθηκευτούν ξανά στη ΒΔ του Νέφους.	Ανάπτυξη προηγμένων αλγορίθμων κρυπτογράφησης, οι οποίοι θα εγγυώνται μεγαλύτερη ασφάλεια των Μαζικών Δεδομένων Υγείας στο μέλλον.

24	Thibaud et al., 2018	Internet of Things (IoT) in high-risk Environment, Health and Safety (EHS) industries: a comprehensive review	Αναζήτηση κυριότερων προκλήσεων και περιοχών μελλοντικής έρευνας και δράσης για τη χρήση του IoT στους τομείς υψηλού κινδύνου (συμπεριλαμβανομένης της υγείας).	Μελέτη του δημοσιευμένου συγγραφικού έργου σχετικού με το IoT, στους τομείς υψηλού κινδύνου (συμπεριλαμβανομένης της υγείας).	Οι σημαντικότερες προκλήσεις της χρήσης IoT στον τομέα της υγείας συνοψίζονται στην ασφάλεια – ιδιωτικότητα, στην τυποποίηση και στην επεκτασιμότητα.	Κοινωνικές και οικονομικές προκλήσεις της τεχνολογίας IoT, στον χώρο της υγείας.	- Ανάγκη για ανάπτυξη πρωτόκολλων, ειδικά για φορητές συσκευές, που λαμβάνουν υπόψη τη διαλειτουργικότητα.	- Διασύνδεση ετερογενών δικτύων που υποστηρίζουν τόσο ηλεκτρικούς όσο και βιοχημικούς τομείς. - Σύγκλιση συστημάτων υγειονομικής περίθαλψης και των έξυπνων περιβαλλοντολογικών συστημάτων	
Ομάδα άρθρων: MACHINE LEARNING									
25	Coronato et al., 2020	Reinforcement learning for intelligent healthcare applications: A survey	Επισκόπηση εφαρμογών της Ενισχυτικής ML στην υγειονομική περίθαλψη, τονίζοντας τις δυνατότητες να υποστηρίξει την ανάπτυξη εξατομικευμένων θεραπειών, στο πλαίσιο του γενικότερου οράματος της Ιατρικής Ακρίβειας.	Διερεύνηση προηγούμενων εργασιών (περίπου 150), επισημαίνοντας τυχόν περιορισμούς, καθώς και πιθανή μελλοντική συνεισφορά της Ενισχυτικής ML, στην υγειονομική περίθαλψη.	- Αν η Ενισχυτική ML συνδυαστεί με Μαζικά Δεδομένα Υγείας (MDY), υπάρχουν μοναδικές ευκαιρίες βελτιστοποίησης θεραπειών, ειδικά για τις επαναλαμβανόμενες. - Με τις δυνατότητες της ML, οι εφαρμογές υγειονομικής περίθαλψης είναι απεριόριστες.	Εφαρμογές Ενισχυτικής ML στον τομέα της υγειονομικής περίθαλψης, μέσω έρευνας σε περίπου 100 εργασίες από σχετικά περιοδικά και πρακτικά συνεδρίων και ομαδοποίησή τους σε 8 κατηγορίες.	- Ιατρική ακρίβειας. - Καθεστώς Δυναμικής Θεραπείας. - Εξατομικευμένη αποκατάσταση. - Ιατρική απεικόνιση. - Διαγνωστικά συστήματα. - Συστήματα ελέγχου. - Συστήματα διαλόγου. - Συστήματα Διαχείρισης Υγείας.	Αντιμετώπιση του θέματος της Ανεπάρκειας Δεδομένων (Data Deficiency) για χρήση τους στην Ενισχυτική ML, στην υγειονομική περίθαλψη.	
26	Van Mens et al., 2020	Applying machine learning on health record data from general practitioners to predict suicidality	Χρησιμοποίηση αλγορίθμων Μηχανικής Μάθησης για τον εντοπισμό αυτοκτονικών τάσεων, από τα ιατρικά δεδομένα, που συλλέγονται από Γενικούς Ιατρούς.	Έρευνα με χρήση ιατρικών δεδομένων, προερχομένων από βάση δεδομένων πρωτοβάθμιας περίθαλψης με 1,5 εκατομ. ασθενείς. Επιλέχθηκαν N=574 περιπτώσεις που έκαναν απόπειρα και N=207.308 ασθενείς με ψυχολογική ευπάθεια.	Οι τεχνικές Μηχανικής Μάθησης μπορούν να χρησιμοποιηθούν ως συμπληρωματικό βοήθημα στην ταυτοποίηση και τη διαστρωμάτωση ασθενών, οι οποίοι διατρέχουν κίνδυνο, λόγω παρατηρούμενης αυτοκτονικής συμπεριφοράς.	Υπάρχει χρονικό περιθώριο αντίδρασης, κι αν ναι πόσο πριν από την απόπειρα αυτοκτονίας, ώστε να αντιληφθεί ο Γενικός Ιατρός, με τη χρήση αλγορίθμων Μηχανικής Μάθησης, τον κίνδυνο που διατρέχει ο μελλοντικός αυτόχειρας;	Δημιουργία αυτόματης ειδοποίησης του Ιατρού, καθώς διαπιστώθηκε ότι: - Το 93% ζήτησε τη γνώμη του Ιατρού 1 χρόνο πριν από την απόπειρα. - Πάνω από 50% είχαν εγγραφεί 1 μήνα πριν από την απόπειρα. - Το 33% είχε εγγραφεί 1 εβδομάδα πριν την απόπειρα.	Η βελτίωση της τεχνικής που χρησιμοποιήθηκε στην έρευνα, μπορεί να βελτιωθεί περαιτέρω, με χρήση δεδομένων από την εκπαίδευση και την απασχόληση.	
27	Eetemadi et al., 2020	The Computational Diet: A Review of Computational Methods Across Diet, Microbiome and Health	Παροχή οδηγιών στη χρήση Μηχανικής Μάθησης και Ανάλυσης Δεδομένων για τη διατροφική υγεία και το μικροβίωμα του εντέρου, για την αναγνώριση κενών στις παρούσες μεθόδους και την περιγραφή νέων σεναρίων, για μελλοντική χρήση.	Έρευνα στην τρέχουσα τεχνολογία σχετικά με την διατροφική υγεία και το μικροβίωμα του εντέρου, από σχετικές πηγές δεδομένων, σε εργαλεία βιοπληροφορικής και σε δυνατότητες Μηχανικής Μάθησης.	Η δυνατότητα προσαρμογής στη διατροφή και σε άλλους περιβαλλοντικούς παράγοντες, καθιστούν το μικροβίωμα του εντέρου έναν εξαιρετικό στόχο για παρεμβάσεις που σχετίζονται με τη διατροφή για τη βελτίωση της υγείας.	Αναζήτηση κατάλληλων εργαλείων, που συντελούν στην κατάρτιση αποτελεσματικής δίαιτας και λήψης συμπληρωμάτων διατροφής, που προσαρμόζονται σε συγκεκριμένες πληθυσμιακές ομάδες ή άτομα	Πλατφόρμα 2 συστημάτων: - Προσδιορισμός του βέλτιστου μικροβιώματος, δεδομένου του επιθυμητού στόχου, ατόμου και περιβάλλοντος. - Παροχή συστάσεων για τη δημιουργία του μικροβιώματος-στόχου.	Περαιτέρω ώθηση στις μεθόδους επεξεργασίας και Ανάλυσης Δεδομένων, προσαρμοσμένες σε κάθε έργο, με βάση παράγοντες όπως, η τεχνολογία sequencing, η διαθεσιμότητα, η διάσταση και μεταβλητότητα των δεδομένων.	
28	Waring et al., 2020	Automated Machine Learning (AutoML): Review of the state-of-the-art and opportunities for healthcare	Μελέτη του συγγραφικού έργου στον τομέα της AutoML, ώστε να βοηθηθούν επαγγελματίες υγείας με περιορισμένες γνώσεις επιστήμης των δεδομένων, να χρησιμοποιούν καλύτερα τα μοντέλα ML.	Μελέτη του συγγραφικού έργου ακαδημαϊκών βάσεων δεδομένων, στον τομέα της Αυτοματοποιημένης Μηχανικής Μάθησης (AutoML).	- Η AutoML είναι ραγδαία εξελισσόμενος τομέας, που θα βοηθήσει τους επαγγελματίες υγείας να χρησιμοποιούν μοντέλα ML. - Υπάρχουν ακόμη περιορισμοί στην χρήσης της, ειδικά όταν πρόκειται για δεδομένα βιοϊατρικής, ως προς το μέγεθος και την ποικιλομορφία τους.	Τυπικά στοιχεία ενός προβλήματος ML, σε γραμμή παραγωγής (pipeline) για AutoML.	- Προετοιμασία δεδομένων (φόρτωση, καθαρισμός, μετασχηματισμοί και κανονικοποιήσεις). - Επιλογή χαρακτηριστικών (features) για δημιουργία του μοντέλου ML. - Επαναληπτική διαδικασία, όπου κάποιος ειδικός δημιουργεί, εκπαιδεύει, βελτιστοποιεί, επικυρώνει και επιλέγει έναν δεδομένο αλγόριθμο ML.	Ανάγκη περαιτέρω έρευνας για ευρεία υιοθέτηση της AutoML στην υγειονομική περίθαλψη, καθώς έχουν διαπιστωθεί δυσκολίες στον χειρισμό του όγκου και της ποικιλομορφίας των Μαζικών Δεδομένων.	

29	Toh et al., 2019	Looking beyond the hype: Applied AI and Machine Learning in translational medicine	Αναζήτηση τρόπων υπέρβασης των περιορισμών των μεθόδων Μηχανικής Μάθησης, ώστε να συμβάλλουν στη Μεταγραφική Ιατρική και στη λήψη αποφάσεων, με βάση τα δεδομένα.	Περιγραφή δημοφιλών μεθόδων Μηχανικής Μάθησης ή λιγότερο γνωστών, που όμως μπορούν να μάθουν από μεγάλα σύνολα δεδομένων, που παράγονται στο εργαστήριο.	Οι τεχνολογίες Μηχανικής Μάθησης έχουν περιορισμούς και αδυναμίες (όπως της διακυβέρνησης, της αναπαραγωγιμότητας και της ερμηνείας των δεδομένων).	Μη τεχνικά ζητήματα που πρέπει να αντιμετωπιστούν, πριν από την εφαρμογή της ML στη Μεταγραφική Ιατρική.	<ul style="list-style-type: none"> - Απαιτείται προσεκτική πρόσβαση στα δεδομένα. - Αποφυγή χρήσης μεροληπτικών (biased) συνόλων δεδομένων. - Προσήλωση στο προς απάντηση ερώτημα. - Απαιτείται εύκολη ερμηνεία των προβλέψεων των συστημάτων ML. 	Απαιτούνται περαιτέρω μελέτες για τον προσδιορισμό της ακρίβειας της ανάπτυξης του ταξινομητή (classifier), ως εργαλείου για την κλινική διάγνωση του καρκίνου, στο πλαίσιο της ανακάλυψης βιοδεικτών.
30	Johnston et al., 2019	Using Machine Learning (ML) Applied to Real-World Healthcare Data for Predictive Analytics: An Applied Example in Bariatric Surgery.	Συνεισφορά της Μηχανικής Μάθησης στην επιλογή ασθενών, με συγκεκριμένα υγειονομικά κριτήρια, από βάσεις δεδομένων υγείας.	Χρήση Πλατφόρμας Προγνωστικής Ανάλυσης (Ανοικτού Κώδικα).	Οι τεχνικές Μηχανικής Μάθησης σε πλατφόρμα Προγνωστικής Ανάλυσης Δεδομένων υγείας, μπορεί να συνεισφέρει στην επιτυχή επιλογή ασθενών για υποβολή τους σε συγκεκριμένη θεραπεία.	Δυνατότητες και περιορισμοί μοντέλων πρόβλεψης επιπέδου ασθενούς (Patient-Level Prediction - PLP).	<ul style="list-style-type: none"> - Πρόβλεψη πιθανότητας να βιώσει ο ασθενής ένα συγκεκριμένο αποτέλεσμα (ευεργετικό ή επιβλαβές), με βάση σύνολο προβλέψεων, όπως δημογραφικά, συννοσηρότητες κ.λπ. - Ένα PLP μοντέλο τείνει να εστιάζει σε μια μεμονωμένη πτυχή (κίνδυνο/πιθανότητα δυαδικού αποτελέσματος). - Μεμονωμένο PLP δεν μπορεί να συνεισφέρει αποτελεσματικά ως εργαλείο CDSS, χρειάζονται πληροφορίες από πολλά. 	Η εκπαίδευση των μοντέλων πρόβλεψης PLP με ολοκληρωμένα σύνολα δεδομένων, που περιλαμβάνουν δεδομένα των ασθενών από διοικητικές πηγές, από τον ΑΗΦΥ τους ή δεδομένα του τρόπου ζωής τους, θα κάνει αυτά τα μοντέλα πρόβλεψης ακόμη αποτελεσματικότερα στο μέλλον.

12. Βιβλιογραφία

12.1 Άρθρα – Συγγράμματα

- Εφημερίδα της Κυβερνήσεως της Ελληνικής Δημοκρατίας (2019) Νόμος 4600/2019 – ΦΕΚ Τεύχος Α' 43/09.03.2019, άρθρο 84, σ. 1135. Available from: <https://www.e-nomothesia.gr/kat-ygeia/nomos-4600-2019-phek-43a-9-3-2019.html> [Accessed 03 Jan 2020].
- Κανέλλος, Λ. (2020) The GDPR Handbook Για DPOs, Επιχειρήσεις & Οργανισμούς, Νομική Βιβλιοθήκη, Αθήνα, 2020, σ. 238-242, 351-370.
- Παπαντωνίου, Ν. (1983) Γενικές Αρχές του Αστικού Δικαίου, 3η εκδ., Αθήνα, 1983. Σπυριδάκη, Α. et al., (2019) Ανασκόπηση, Εφαρμογές της «κινητής υγείας» (mobile health) στα χρόνια νοσήματα και διερεύνηση της αποτελεσματικότητάς τους: Αρχεία Ελληνικής Ιατρικής 2019, 36(1), σ. 73-80.
- Aceto, G. (2020) Industry 4.0 and Health: Internet of Things, Big Data, and Cloud Computing for Healthcare 4.0. Available from: <https://doi.org/10.1016/j.jii.2020.100129> [Accessed 21 Nov 2020].
- Alam, M. and Rahmani, R. (2020) Intelligent context-based healthcare metadata aggregator in internet of medical things platform. Available from: <https://doi.org/10.1016/j.procs.2020.07.058> [Accessed 11 Oct 2020].
- Alharthi, H. (2018) Healthcare predictive analytics: An overview with a focus on Saudi Arabia. Journal of Infection and Public Health. Available from: <https://doi.org/10.1016/j.jiph.2018.02.005> [Accessed 11 Sep 2020].
- Andrews, J.G. et al., (2014) What will 5G be? IEEE J. Sel. Areas Commun. 32 (6) (2014) 1065–1082.
- Archenaa, J. et al., (2015) A Survey of Big Data Analytics in Healthcare and Government. Procedia Computer Science. 50: 408 – 413.
- Beck, A. and Kurz, M. (2020) A Perspective on Machine Learning Methods in Turbulence Modeling. Available from: DOI: [10.13140/RG.2.2.17469.69608](https://doi.org/10.13140/RG.2.2.17469.69608) [Accessed 27 Jan 2021].
- Bloom, D. et al., (2004) The Effect of Health on Economic Growth: A Production Function Approach. Available from: <https://doi.org/10.1016/j.worlddev.2003.07.002> [Accessed 17 Nov 2020].
- Boccardi, F. et al., (2014) Five disruptive technology directions for 5G, IEEE Commun. Mag. 52 (2) (2014) 74–80.
- Buchan, I. et al., (2009) The Fourth Paradigm: Data-Intensive Scientific Discovery. A unified modeling approach to data-intensive healthcare. Redmond: Microsoft Research; 2009. p. 91–8. Available from: <https://www.immagic.com/eLibrary/ARCHIVES/EBOOKS/M091000H.pdf> [Accessed 14 Nov 2020].
- Capobianco, E. (2020) Imprecise Data and Their Impact on Translational Research in Medicine. Available from: [10.3389/fmed.2020.00082](https://doi.org/10.3389/fmed.2020.00082) [Accessed 30 Nov 2020].
- Cirillo, D. and Valencia, A. (2019) Big data analytics for personalized medicine. Current Opinion in Biotechnology 2019, (58), p. 161–167.
- Coronato, A. (2018) Engineering high quality medical software: regulations, standards, methodologies and tools for certification, healthcare technologies. Inst Eng Technol 2018. <https://books.google.it/books?id=hPdIDwAAQBAJ>
- Coronato, A. et al., (2020) Reinforcement learning for intelligent healthcare applications: A survey. Available from: <https://doi.org/10.1016/j.artmed.2020.101964> [Accessed 29 Nov 2020].
- Cottle, M. et al., (2013) Transforming Health Care through Big Data. Institute for Health Technology Transformation, p. 5.
- Dagliati, A. et al., (2018) Big Data as a Driver for Clinical Decision Support Systems: Learning Health Systems Perspective. Available from: <https://doi.org/10.3389/fdigh.2018.00008> [Accessed 11 Oct 2020].
- Eetemadi, A. et al., (2020) The Computational Diet: A Review of Computational Methods Across Diet, Microbiome, and Health. Available from: <https://doi.org/10.3389/fmicb.2020.00393> [Accessed 04 Dec 2020].
- Elhoseny, M. et al., (2018) A hybrid model of Internet of Things and cloud computing to manage big data in health services applications. Future Gen. Comput. Syst. 86:1383–94, Available from: [10.1016/j.future.2018.03.005](https://doi.org/10.1016/j.future.2018.03.005) [Accessed 14 Nov 2020].
- Ghamari, M. et al., (2016) A survey on wireless body area networks for eHealthcare systems in residential environments. Available from: <https://doi.org/10.3390/s16060831> [Accessed 24 Dec 2020].

- Harper, E. (2014). Can big data transform electronic health records into learning health systems? *Stud. Health Technol. Inform.* 201, 470–475. Available from: [10.3233/978-1-61499-415-2-470](https://doi.org/10.3233/978-1-61499-415-2-470) [Accessed 04 Dec 2020].
- Hezarjaribi, N. et al., (2016) A Machine Learning Approach for Medication Adherence Monitoring Using Body-Worn Sensors: Proceedings of the 2016 Design, Automation & Test in Europe Conference & Exhibition (DATE).
- Hoos, H. et al., (2014) An efficient approach for assessing hyperparameter importance. *International Conference on Machine Learning*.
- Hulslen, T. et al., (2019) From Big Data to Precision Medicine. Available from: <https://doi.org/10.3389/fmed.2019.00034> [Accessed 04 Dec 2020].
- Hurley, R. (2019) *Big Data: A Guide to Big Data Trends, Artificial Intelligence, Machine Learning, Predictive Analytics, Internet of Things, Data Science, Data Analytics, Business Intelligence, and Data Mining*. Independently published, p. 27-29.
- Jaba Praba J. and Srividhya V. (2016) Big Data Analytics in Healthcare: Proceedings of 9th National Level Science Symposium, Rajkot, India, (3), p. 212-215.
- Jiang, F. et al., (2017) Artificial intelligence in healthcare: past, present and future, *BMJ*.
- Johnston, S. et al., (2019) Using Machine Learning Applied to Real-World Healthcare Data for Predictive Analytics: An Applied Example in Bariatric Surgery. Available from: <https://doi.org/10.1016/j.jval.2019.01.011> [Accessed 03 Jan 2020].
- Kaur, P. (2018) Big Data and Machine Learning Based Secure Healthcare Framework. *Procedia Computer Science* (132), p. 1049–1059. Available from: <https://doi.org/10.1016/j.procs.2018.05.020> [Accessed 03 Jan 2020].
- Khan, S. et al., (2019) Unbiased data analytic strategies to improve biomarker discovery in precision medicine. Available from: <https://doi.org/10.1016/j.drudis.2019.05.018> [Accessed 14 Nov 2020].
- Komer, B. et al., (2014) Hyperopt-sklearn: automatic hyperparameter configuration for scikit-learn. *ICML workshop on AutoML*, Citeseer.
- Kumar, P. M. et al., (2018) Cloud and IoT based disease prediction and diagnosis system for healthcare using Fuzzy neural classifier. *Future Generation Computer Systems* (86), p. 527-534. Available from: <https://doi.org/10.1016/j.future.2018.04.036> [Accessed 05 Jan 2020].
- Lee, I. (2019) The Internet of Things for enterprises: An ecosystem, architecture, and IoT service business model. *Internet of Things* 7 (2019) 100078. Available from: <https://doi.org/10.1016/j.iot.2019.100078> [Accessed 03 Jan 2020].
- Lepenioti, K. et al., (2020) Prescriptive analytics: Literature review and research challenges Available from: <https://doi.org/10.1016/j.ijinfomgt.2019.04.003> [Accessed 13 Sep 2020].
- Lopes, J. et al., (2020) Predictive and Prescriptive Analytics in Healthcare: A Survey. Available from: <https://doi.org/10.1016/j.procs.2020.03.078> [Accessed 03 Sep 2020].
- Majeed, A. and Lee, S. (2020) Attribute susceptibility and entropy based data anonymization to improve users community privacy and utility in publishing data. *Appl Intell* 50, 2555–2574 (2020). Available from: <https://doi.org/10.1007/s10489-020-01656-w> [Accessed 03 Jan 2021].
- Mamatjan, Y. et al., (2017) Molecular signatures for tumor classification: an analysis of the cancer genome atlas data. *J Mol Diagn* 2017; 19: 881–91.
- Mani, N. (2020) An IoT Guided Healthcare Monitoring System for Managing Real-Time Notifications by Fog Computing Services Available from: <https://doi.org/10.1016/j.procs.2020.03.424> [Accessed 04 Dec 2020].
- Mehta, S. et al., (2019) Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study. Available from: <https://doi.org/10.1016/j.jbi.2019.103311> [Accessed 03 Jan 2020].
- Mehtaa, N. et al., (2019) Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study Available from: <https://doi.org/10.1016/j.jbi.2019.103311> [Accessed 22 Oct 2020].
- Mrozek, D. et al., (2020) Fall detection in older adults with mobile IoT devices and machine learning in the cloud and on the edge Available from: <https://doi.org/10.1016/j.ins.2020.05.070> [Accessed 29 Nov 2020].
- National Research Council, Committee on A Framework for Developing a New Taxonomy of Disease. *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New*

- Taxonomy of Disease. National Academies Press; Washington DC: 2011. Available from: <https://pubmed.ncbi.nlm.nih.gov/22536618/> [Accessed 08 Nov 2020].
- Ngiam, K.Y. and Khor, I.W. (2019) Big data and machine learning algorithms for health-care delivery. *The Lancet Oncology*, (20:5), p. e262-e273. Available from: [https://doi.org/10.1016/S1470-2045\(19\)30149-4](https://doi.org/10.1016/S1470-2045(19)30149-4) [Accessed 03 Jan 2020].
 - Opalinski, A. (2019) et al., Medical data exploration based on the heterogeneous data sources aggregation system. *Ann Comput Sci InformSyst.* 18:591–7. Available from: <https://ieeexplore.ieee.org/document/8860014> [Accessed 03 Jan 2020].
 - Osseiran, A. et al., (2014) Scenarios for 5G mobile and wireless communications: The vision of the METIS project, *IEEE Commun. Mag.* 52 (5) (2014) 26–35.
 - Pearson, T. et al., (2020) Precision Health Analytics with Predictive Analytics and Implementation Research. Available from: <https://doi.org/10.1016/j.jacc.2020.05.043> [Accessed 09 Nov 2020].
 - Peng, R. (2015) The reproducibility crisis in science: A statistical counterattack. Available from: <https://doi.org/10.1111/j.1740-9713.2015.00827.x> [Accessed 30 Nov 2020].
 - Pramanik, I. (2020) Healthcare informatics and analytics in big data. Available from: <https://doi.org/10.1016/j.eswa.2020.113388> [Accessed 09 Nov 2020].
 - Prospero, M. et al., (2018) Big data hurdles in precision medicine and precision public health. *BMC Med Inform Decis Mak* 2019;18:139. Available from: <https://doi.org/10.1186/s12911-018-0719-2> [Accessed 09 Nov 2020].
 - Rajabion, L. et al., (2019) Healthcare big data processing mechanisms: The role of cloud computing. *International Journal of Information Management*, (49), p. 271–289.
 - Reis, M. and Gins, G. (2017) Industrial Process Monitoring in the Big Data/Industry 4.0 Era: from Detection, to Diagnosis, to Prognosis Available from: <https://doi.org/10.3390/pr5030035> [Accessed 19 Nov 2020].
 - Rodrigues, J.J. et al., (2018) Enabling Technologies for the Internet of Health Things. Available from: <https://www.researchgate.net/publication/322261039> [Accessed 29 Nov 2020].
 - Rodriguez, J.D. et al., (2009) Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Trans Pattern Anal Mach Intell.* 32:569–75, Available from: <https://ieeexplore.ieee.org/document/5342427> [Accessed 09 Nov 2020].
 - Saheb, T. and Izadi, L. (2019) Paradigm of IoT big data analytics in the healthcare industry: A review of scientific literature and mapping of research trends. *Telematics and Informatics* (41), p. 70–85.
 - Savjani, R. (2017) 5 things to consider before choosing an IoT platform company (2017). Available from: <https://www.softwebsolutions.com/resources/IoT-platform-company.html> [Accessed 12 Dec 2020].
 - Schork, N.J. (2015) Personalized medicine: time for one-person trials. *Nature.* 520 (7549):609-11.
 - Serrano, L. (2020) *Exploring Machine Learning Basics*, Manning Publications, p. 2-28.
 - Siriwardhana, Y. et al., (2020) The role of 5G for digital healthcare against COVID-19 pandemic: Opportunities and challenges. Available from: <https://doi.org/10.1016/j.ict.2020.10.002> [Accessed 12 Dec 2020].
 - Song, C. et al., (2020) Big data-driven precision medicine: Starting the custom-made era of iatrolology. Available from: <https://doi.org/10.1016/j.biopha.2020.110445> [Accessed 19 Nov 2020].
 - Stopa, B. et al., (2019) Variance Reduction in Neurosurgical Practice: The Case for Analytics-Driven Decision Support in the Era of Big Data. Available from: <https://doi.org/10.1016/j.wneu.2019.01.292> [Accessed 11 Sep 2020].
 - Sweeney, L. (2002) k-Anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10 (5), 2002; 557-570. Available from https://epic.org/privacy/reidentification/Sweeney_Article.pdf [Accessed 13 Sep 2020].
 - Tao, H. et al., (2019) Economic perspective analysis of protecting big data security and privacy Available from: <https://doi.org/10.1016/j.future.2019.03.042> [Accessed 20 Oct 2020].
 - Thibaud, M. et al., (2018) Internet of Things (IoT) in high-risk Environment, Health and Safety (EHS) industries: A comprehensive review. *Decision Support Systems*, (108), p. 79-95. Available from: <https://doi.org/10.1016/j.dss.2018.02.005> [Accessed 03 Jan 2020].
 - Toh, T. et al., (2019) Looking beyond the hype: Applied AI and machine learning in translational medicine. Available from: <https://doi.org/10.1016/j.ebiom.2019.08.027> [Accessed 29 Nov 2020].

- Van Mens, K. et al., (2020) Applying machine learning on health record data from general practitioners to predict suicidality Available from: <https://bigdata-madesimple.com/machine-learning-explained-understanding-supervised-unsupervised-and-reinforcement-learning/> [Accessed 30 Nov 2020].
- Wang, Y. et al., (2018) Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. Available from: <https://doi.org/10.1016/j.techfore.2015.12.019> [Accessed 03 Jan 2020].
- Waring, J. et al., (2020) Automated machine learning: Review of the state-of-the-art and opportunities for healthcare. Available from: <https://doi.org/10.1016/j.artmed.2020.101822> [Accessed 29 Nov 2020].
- Willems, S. et al., (2019) The potential use of big data in oncology. Oral Oncology, (98), p. 8-12. Available from: <https://doi.org/10.1016/j.oraloncology.2019.09.003> [Accessed 03 Jan 2020].
- Zhang, Y.D. and Zhou, Q. (2019) Grand Challenges for Medtech Data Analytics. Available from: <https://doi.org/10.3389/fmedt.2019.00002> [Accessed 24 Sep 2020].
- Zhang, Z. et al., (2020) Predictive analytics in the era of big data: opportunities and challenges. Available from: [10.21037/atm.2019.10.97](https://doi.org/10.21037/atm.2019.10.97) [Accessed 03 Nov 2020].

12.2 Διαδίκτυο

- Αρχή Προστασίας Δεδομένων Προσωπικού Χαρακτήρα (2020) Απόφαση 20/2020. Available from: https://www.dpa.gr/sites/default/files/2020-07/20_2020anonym.pdf [Accessed 30 Jan 2021].
- Αρχή Προστασίας Δεδομένων Προσωπικού Χαρακτήρα (2020) Γνωμοδότηση 1/2020. Available from: https://www.dpa.gr/sites/default/files/2020-01/gnomodotisi%201_2020.pdf [Accessed 30 Ιαν 2021].
- Αρχή Προστασίας Δεδομένων Προσωπικού Χαρακτήρα (2021) Ασφάλεια προσωπικών δεδομένων. Available from: https://www.dpa.gr/enimerwtiko/thematikes_enotites/asfaleia [Accessed 19 Ιαν 2021].
- Βουλή των Ελλήνων (2019) Σύνταγμα της Ελλάδας. Available from: <https://www.hellenicparliament.gr/Vouli-ton-Ellinon/To-Politevma/Syntagma/> [Accessed 19 Dec 2020].
- Γενικό Νοσοκομείο Καβάλας (2020) Γενικά για τον Κανονισμό της Αρχής Προστασίας Προσωπικών Δεδομένων. Available from: <https://kavalahospital.gr/Patients/Προσωπικα-Δεδομενα/GDPR> [Accessed 17 Jan 2021].
- Ελληνική Δημοκρατία - Υπουργείο Υγείας (2018) mHealth - Υγεία εν Κινήσει. Available from: <https://www.moh.gov.gr/articles/ehealth/5688-mhealth> [Accessed 17 Jan 2021].
- Ελληνική Δημοκρατία (2019) Ν.4600/2019, ΦΕΚ Α 43/9-3-2109, Εκσυγχρονισμός και Αναμόρφωση Θεσμικού Πλαισίου Ιδιωτικών Κλινικών, Σύσταση Εθνικού Οργανισμού Δημόσιας Υγείας, Σύσταση Εθνικού Ινστιτούτου Νεοπλασιών και λοιπές διατάξεις. Άρθρο 84, σ. 1135.
- Ευρωπαϊκή Ένωση (1992) Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης, Συνθήκη για την Ευρωπαϊκή Ένωση (ΣΕΕ), Άρθρο 4. Available from: <https://eur-lex.europa.eu/legal-content/EL/TXT/HTML/?uri=CELEX:12016M004> [Accessed 19 Dec 2020].
- Anon (2018) \$11.45 Bn Big Data in Healthcare Market, 2025. PRNewswire. Available from: <https://www.prnewswire.com/news-releases/1145-bn-big-data-in-healthcare-market-2025-300623544.html>
- Apple (2018) Accessing Health Records with HealthKit. Available from: <https://developer.apple.com/videos/play/wwdc2018/706> [Accessed 16 Dec 2019].
- Ashton, K. (2009) That 'Internet of Things' Thing. Available from: <https://www.rfidjournal.com/articles/view?4986> [Accessed 17 Dec 2019].
- BBC News, Tech (2014) Community Health Systems data hack hits 4.5 million. Available from: <https://www.bbc.com/news/technology-28838661> [Accessed 10 Dec 2020].
- Brown, N. (2015) Healthcare Data Growth: An Exponential Problem. Available from: <https://www.nextech.com/blog/healthcare-data-growth-an-exponential-problem> [Accessed 17 Nov 2020].
- CISCO (2019) What Is 5G? Available from: <https://www.cisco.com/c/en/us/solutions/what-is-5g.html> [Accessed 03 Jan 2020].
- Comarch (2019) Internet of Healthcare Things. Available from: <https://www.comarch.com/iot-ecosystem/internet-of-healthcare-things/> [Accessed 02 Jan 2020].
- De Capitani di Vimercati, S. and Foresti, S. (2011) Encyclopedia of Cryptography and Security. Quasi-Identifier. Available from: https://link.springer.com/referenceworkentry/10.1007%2F978-1-4419-5906-5_763#howtocite [Accessed 19 Dec 2020].

- Desjardins, J. (2018) How Big Data Will Unlock the Potential of Healthcare. Available from: https://www.visualcapitalist.com/big-data-healthcare/?fbclid=IwAR3uT3CsdV_0JAO5reug-8ogOIGupvIFkBCj56P_U06j6HFOPk0d4jk_ey4 [Accessed 17 Nov 2020].
- Dua, D. and Graff, C. (2019) UCI Machine Learning Repository. Irvine, CA: University of California, School of Information and Computer Science. Available from: <http://archive.ics.uci.edu/ml> [Accessed 05 Jan 2020].
- Edwards, J. (2019) CIO.com - What is predictive analytics? Transforming data into future insights. Available from: <https://www.cio.com/article/3273114/what-is-predictive-analytics-transforming-data-into-future-insights.html> [Accessed 24 Dec 2019].
- European Commission (2014) Green Paper on mobile health ("mHealth"). Available from: <https://ec.europa.eu/digital-single-market/en/news/green-paper-mobile-health-mhealth> [Accessed 17 Jan 2021].
- European Commission (2019) eHealth : Digital health and care. Available from: https://ec.europa.eu/health/ehealth/overview_en [Accessed 19 Dec 2019].
- European Commission (2020) White Paper on Artificial Intelligence - A European approach to excellence and trust. Available from: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf [Accessed 17 Jan 2021].
- European Parliament News (2020) Artificial intelligence: threats and opportunities. Available from: <https://www.europarl.europa.eu/news/en/headlines/society/20200918STO87404/artificial-intelligence-threats-and-opportunities> [Accessed 19 Nov 2020].
- Forbes (2019) 7 trends that healthtech innovators expect to shape 2019. Available from: <https://www.forbes.com/sites/clairenovorol/2019/01/09/7-trends-that-healthtech-innovators-expect-to-shape-2019/#d8b7e862a2f9> [Accessed 09 Nov 2020].
- Friedman, U. (2012) Big Data: A Short History. FP Website. Available from: <https://foreignpolicy.com/2012/10/08/big-data-a-short-history/> [Accessed 01 Jan 2020].
- Gartner (2019) The CIO's Guide to Artificial Intelligence. Available from: <https://www.gartner.com/smarterwithgartner/the-cios-guide-to-artificial-intelligence/> [Accessed 02 Jan 2020].
- Gartner Glossary (2019) Big Data. Available from: <https://www.gartner.com/en/information-technology/glossary/big-data> [Accessed 17 Dec 2019].
- Glen, S (2014) Least Squares Regression Line: Ordinary and Partial. Available from: StatisticsHowTo.com: Elementary Statistics for the rest of us! <https://www.statisticshowto.com/least-squares-regression-line/> [Accessed 13 Jan 2021].
- Gordon J.I (2012) Honor thy gut symbionts redux. Science 2012, 336:1251–1253.
- Great Learning Team (2020), Use of Reinforcement Learning in Healthcare. Available from: <https://www.mygreatlearning.com/blog/reinforcement-learning-in-healthcare/> [Accessed 08 Nov 2020].
- Hit Consultant (2016) What is the Relationship Between Precision Medicine & Predictive Analytics? Available from: <https://hitconsultant.net/2016/02/22/31535/> [Accessed 17 Jan 2021].
- Jain, A. (2016) The 5 V's of big data. Available from: <https://www.ibm.com/blogs/watson-health/the-5-vs-of-big-data/> [Accessed 17 Dec 2019].
- Kolhe, A. (2020) What is Prescriptive Analytics and How Can it Help You? Available from: <https://www.vovia.com/blog/analytics/what-is-prescriptive-analytics/> [Accessed 17 Dec 2020].
- Lehman, S. and Lehman, J. (2015) NNT for Statins vs. the Mediterranean Diet. Available from: <https://www.acatoday.org/News-Publications/ACA-News-Archive/ArtMID/5721/ArticleID/106/NNT-for-Statins-vs-the-Mediterranean-Diet> [Accessed 10 Nov 2020].
- Microsoft Azure (2019) What is cloud computing? Available from: https://www.sas.com/en_us/insights/analytics/machine-learning.html [Accessed 02 Jan 2020].
- Microsoft Azure Machine Learning AutoML (2020) Tutorial: Use automated machine learning to predict taxi fares. Available from <https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-auto-train-models> [Accessed 10 Jan 2021].
- Net Over World (2020) Machine Learning. Available from: <https://nowmag.gr/machine-learning/> [Accessed 26 Jan 2021].
- Nurseslabs (2021) Common Laboratory Values Cheat Sheet. Available from: <https://nurseslabs.com/> [Accessed 31 Jan 2021].

- Patrizio, A. (2018) Networkworld, IDC: Expect 175 zettabytes of data worldwide by 2025. Available from: <https://www.networkworld.com/article/3325397/idc-expect-175-zettabytes-of-data-worldwide-by-2025.html> [Accessed 31 Jan 2021].
- Ross, C. and Sweltitz, I. (2018) IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show. STAT, 2018. Available from: <https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/> [Accessed 02 Jan 2020].
- SAS (2019) Machine Learning. What it is and why it matters. Available from: https://www.sas.com/en_us/insights/analytics/machine-learning.html [Accessed 02 Jan 2020].
- Shendre, S. (2020) Model Drift in Machine Learning, How and when should machine learning models be retrained. Available from: <https://towardsdatascience.com/model-drift-in-machine-learning-models-8f7e7413b563> [Accessed 30 Jan 2021].
- SR Cloud Solutions (2021) Imperva Camouflage. Available from: <https://srcloudsolutions.co.uk/web-application-firewall/> [Accessed 30 Jan 2021].
- Szegner, P. (2019) International Congress of Health Workforce Education and Research, Nicosia, SEPEN, WHO CC. Available from: https://inhwe.org/sites/default/files/documents/Peter%20Szegner_Nicosia_Digital%20skills%20.pdf [Accessed 17 Nov 2020].
- Technosoft (2019) Internet of Healthcare Things (IoHT). Available from: <https://techno-soft.com/internet-of-healthcare-things-ioht.html/> [Accessed 27 Sep 2020].
- TRŪATA Data Anonymization Solution. Available from: <https://www.truata.com/solution/how-it-works/> [Accessed 27 Jan 2021].
- Van Loon, R. (2020) Machine learning explained: Understanding supervised, unsupervised, and reinforcement learning. Available from: <https://bigdata-madesimple.com/machine-learning-explained-understanding-supervised-unsupervised-and-reinforcement-learning/> [Accessed 30 Nov 2020].
- VIU Library (2020) Available from: <https://library.viu.ca/c.php?g=189013&p=1247410> [Accessed 07 Sep 2020].
- WHO (2019) eHealth at WHO. Available from: <https://www.who.int/ehealth/about/en/> [Accessed 19 Dec 2019].