



**ΤΜΗΜΑ ΟΡΓΑΝΩΣΗ ΚΑΙ ΔΙΟΙΚΗΣΗΣ  
ΕΠΙΧΕΙΡΗΣΕΩΝ**

**Π.Μ.Σ. ΔΙΟΙΚΗΣΗ ΥΠΗΡΕΣΙΩΝ ΥΓΕΙΑΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ:**

**«ΠΡΟΓΝΩΣΤΙΚΕΣ ΑΝΑΛΥΣΕΙΣ ΔΕΥΤΕΡΟΓΕΝΩΝ  
ΔΕΔΟΜΕΝΩΝ ΜΗ ΕΠΕΙΓΟΝΤΩΝ ΠΕΡΙΣΤΑΤΙΚΩΝ  
ΜΕΤΑΦΟΡΑΣ ΑΣΘΕΝΩΝ ΣΕ ΜΟΝΑΔΕΣ ΤΕΧΝΗΤΟΥ  
ΝΕΦΡΟΥ ΑΠΟ ΤΟ Ε.Κ.Α.Β. ΘΕΣΣΑΛΟΝΙΚΗΣ»**

Υποβλήθηκε ως αιτούμενο για την απόκτηση του  
μεταπτυχιακού διπλώματος

**ΕΥΣΤΑΘΙΟΣ ΦΕΛΕΣΑΚΗΣ  
ΘΕΣΣΑΛΟΝΙΚΗ 2022**

**ΕΠΙΒΛΕΠΟΝΤΕΣ ΚΑΘΗΓΗΤΕΣ:  
ΤΑΡΑΜΠΑΝΗΣ ΚΩΝΣΤΑΝΤΙΝΟΣ  
ΚΑΛΑΜΠΟΚΗΣ ΕΥΑΓΓΕΛΟΣ**

© Ευστάθιος Φελεσάκης, 2022

Με επιφύλαξη παντός δικαιώματος.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ' ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα.

Το περιεχόμενο αυτής της εργασίας δεν απηχεί απαραίτητα τις απόψεις του Τμήματος, του Επιβλέποντα, ή της Επιτροπής που την ενέκρινε.

## **Υπεύθυνη Δήλωση**

Βεβαιώνω ότι είμαι συγγραφέας αυτής της εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης έχω αναφέρει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επίσης βεβαιώνω ότι αυτή η εργασία προετοιμάστηκε από εμένα προσωπικά ειδικά για τις απαιτήσεις του Προγράμματος Μεταπτυχιακών Σπουδών στη Διοίκηση Υπηρεσιών Υγείας του Τμήματος Οργάνωσης και Διοίκησης Επιχειρήσεων του Πανεπιστημίου Μακεδονίας.

*Ευστάθιος Φελεσάκης*

## **Ευχαριστίες**

Ευχαριστώ όλους τους διδάσκοντες και ιδιαίτερα τον κύριο Βασίλειο Αλετρά υπεύθυνο του Μεταπτυχιακού Προγράμματος της Διοίκησης Υπηρεσιών Υγείας του τμήματος Οργάνωσης και Διοίκησης Επιχειρήσεων του Πανεπιστημίου Μακεδονίας για τις γνώσεις που μου μετέδωσαν κατά την διάρκεια των σπουδών μου. Ιδιαίτέρως ευχαριστώ τον κύριο Ταραμπάνη Κωνσταντίνο και τον κύριο Καλαμπόκη Ευάγγελο που με τίμησαν ως επιβλέποντες της παρούσας διπλωματικής εργασίας.

Επίσης, ευχαριστώ τον Ιατρό κύριο Ασλανίδη Θεόδωρο, καθώς η βοήθεια του σε θέματα ανάλυσης προνοσοκομειακών δεδομένων ήταν ουσιαστικής σημασίας για την εκπόνηση της παρούσας διπλωματικής εργασίας.

Ευχαριστίες θα ήθελα να απευθύνω στο Ε.Κ.Α.Β. Θεσσαλονίκης που μου έδωσε την δυνατότητα πρόσβασης στις πληροφορίες και τα δεδομένα που χρησιμοποιήθηκαν για την εκπόνηση της παρούσας διπλωματικής εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια μου που με υποστηρίζει σε κάθε μου προσπάθεια.

*Ευστάθιος Φελεσάκης 2022*

## *Περίληψη*

Η τεχνολογία αδιαμφισβήτητα αυξάνεται με τεράστιους ρυθμούς. Οι νέες τεχνολογίες δημιουργούν απαιτήσεις αποθήκευσης και κατηγοριοποίησης δεδομένων οι οποίες είναι αναγκαίες όσο ποτέ ξανά στην ιστορία της ανθρωπότητας. Στον τομέα της προνοσοκομειακής φροντίδας οι αναλύσεις και οι σωστές προβλέψεις μέσω δεδομένων μη επειγόντων περιστατικών όπως είναι αυτά των ασθενών που χρειάζονται μεταφορά για αιμοκάθαρση σε μονάδα τεχνητού νεφρού είναι πάρα πολύ σημαντική. Οι επιπλοκές της συγκεκριμένης ασθένειας είναι πολλές και σοβαρές, έτσι πρέπει να γίνει αντιληπτό ότι οι προβλέψεις με όσο το δυνατό μεγαλύτερη ακρίβεια είναι πάρα πολύ σημαντικές.

Στην παρούσα διπλωματική εργασία αναπτύχθηκε ένα μοντέλο χρονοσειρών για την πρόβλεψη ενδεχομένων κλήσεων στο Ε.Κ.Α.Β. Θεσσαλονίκης όπως επίσης και την ολοκλήρωση διακομιδών μη επειγόντων περιστατικών. Συγκεκριμένα, διερευνήθηκε στο κατά πόσο αυτού του είδους το μοντέλο είναι σημαντικό όσο αναφορά διακομιδές ασθενών που πάσχουν από νεφρική ανεπάρκεια και χρήζουν μεταφορά σε μονάδες τεχνητού νεφρού για αιμοκάθαρση με την μέθοδο της αιμοδιάλυσης, όπως επίσης και την σημαντικότητα της ολοκλήρωσης των διακομιδών επιστρέφοντας τους ασθενείς στην οικία τους.

*Λέξεις κλειδιά: μηχανική μάθηση, τεχνητή νοημοσύνη, εξόρυξη δεδομένων, προγνωστική ανάλυση, νευρωνικά δίκτυα, μηχανές διανυσμάτων στήριξης, δένδρα απόφασης, προνοσοκομειακή ιατρική, Εθνικό Κέντρο Άμεσης Βοήθειας.*

## Πίνακας περιεχομένων

Περίληψη .....	1
Εικόνες.....	4
Πίνακες .....	4
Γραφήματα .....	4
Συντομεύσεις .....	5
1.Εισαγωγή.....	6
1.1 Περιγραφή προβλήματος.....	6
1.2 Στόχος μελέτης .....	7
1.3 Δομή διπλωματικής εργασίας.....	7
2.Τι είναι η μηχανική μάθηση .....	9
2.1 Επισκόπηση .....	9
2.2 Προσεγγίσεις μηχανικής μάθησης.....	10
2.3 Τεχνητή νοημοσύνη.....	11
2.4 Βελτιστοποίηση.....	12
2.5 Στατιστική.....	12
2.6 Τύποι μαθησιακών αλγορίθμων .....	13
2.6.1 Εποπτευόμενη μάθηση .....	13
2.6.2 Μη εποπτευόμενη μάθηση .....	13
2.6.3 Μάθηση ενίσχυσης .....	14
2.6.4 Αυτομάθηση .....	14
2.6.5 Μάθηση χαρακτηριστικών .....	14
2.6.6 Μάθηση σποραδικών λεξικών.....	15
2.6.7 Ανίχνευση ανωμαλιών.....	16
2.7 Μοντέλα.....	16
2.7.1 Τεχνητά νευρωνικά δίκτυα.....	16
2.7.2 Δέντρα αποφάσεων.....	18
2.7.3 Μηχανές φορέα υποστήριξης .....	19
2.7.4 Ανάλυση παλινδρόμησης .....	20
2.7.5 Μπαγεζιανά Δίκτυα .....	20
2.7.6 Περιορισμοί .....	21
2.7.7 Αξιολογήσεις μοντέλου .....	22
2.7.8 Ηθική .....	22
3. Τι είναι η εξόρυξη δεδομένων .....	24
3.1 Ετυμολογία .....	24
3.2 Εξόρυξη Δεδομένων.....	24
3.3 Πρότυπα.....	24
3.4 Απόρρητο Και Ηθική .....	25
4.Τι είναι το Προγνωστικό Μοντέλο.....	27
4.1 Μοντέλα.....	27
4.2 Εφαρμογή στην υγειονομική περίθαλψη.....	27
5.Δεδομένα υγείας .....	28
5.1 Μεγάλα δεδομένα (Big Data) στον τομέα της υγείας .....	28
5.1.1 Κύριοι τύποι και πηγές μεγάλων δεδομένων στον τομέα της υγείας.....	28
5.1.2 Μεγάλα δεδομένα στην ιατρική και τις κλινικές.....	30
5.1.3 Μεγάλα δεδομένα στη δημόσια υγεία .....	32
5.1.4 Μεγάλα δεδομένα σε ιατρικά πειράματα .....	33

5.1.5 Μεγάλα δεδομένα στην ιατρική βιβλιογραφία των νοσοκομειακών πληροφοριακών συστημάτων και η εξέλιξή τους.....	34
5.2 Τα αναλυτικά δεδομένα.....	35
5.2.1 Όγκος (Volume) .....	36
5.2.2 Ταχύτητα (Velocity) .....	36
5.2.3 Ποικιλία (Variety) .....	37
5.2.4 Αξιοπιστία (Veracity).....	37
5.2.5 Ισχύς (Validity).....	37
5.2.6 Βιωσιμότητα (Viability) .....	38
5.2.7 Μεταβλητότητα (Volatility) .....	38
5.2.8 Ευπάθεια (Vulnerability).....	38
5.2.9 Οπτικοποίηση (Visualization) .....	39
6.Ιδιαιτερότητα των Μεγάλων Δεδομένων στην Υγεία .....	40
6.1 Μείωση κόστους υγειονομικής περίθαλψης.....	40
6.2 Ετερογένεια .....	40
6.3 Ατελή δεδομένα.....	41
6.4 Επικαιρότητα .....	41
6.5 Απόρρητο δεδομένων .....	42
6.6 Ιδιοκτησία.....	42
7. Μελέτη περίπτωσης: Προγνωστικές αναλύσεις δευτερογενών δεδομένων μη επειγόντων περιστατικών μεταφοράς ασθενών σε μονάδες τεχνητού νεφρού από το Ε.Κ.Α.Β. Θεσσαλονίκης.....	44
7.1 Το Εθνικό Κέντρο Άμεσης Βοήθειας (Ε.Κ.Α.Β.) .....	44
7.2 Ηλεκτρονική διαχείριση δεδομένων Ε.Κ.Α.Β.....	50
7.3 Μελέτη περίπτωσης.....	55
7.3.1 Υλικό και Μέθοδοι .....	55
7.3.2 Αποτελέσματα .....	57
8. Συζήτηση .....	68
9. Συμπεράσματα.....	70
Βιβλιογραφία.....	72

## Εικόνες

Εικόνα 1. Η ML ως μέρος της AI.....	11
Εικόνα 2. Βαθιά Νευρωνικά Δίκτυα.....	17
Εικόνα 3. Perceptron.....	18
Εικόνα 4. Παράδειγμα δέντρου αποφάσεων.....	19
Εικόνα 5. Παράδειγμα δικτύου Ασία.....	21
Εικόνα 6. Οι κυριότερες κατηγορίες Big Data στην ιατρική βιβλιογραφία.....	29
Εικόνα 7. Οι 12 μεγάλοι σταθμοί του Εθνικού Κέντρου Άμεσης Βοήθεια.....	45
Εικόνα 8. Παράδειγμα διαγράμματος ροής που εφαρμόζεται στην Ελλάδα για την αποστολή ασθενοφόρου ή κινητής ιατρικής μονάδας .....	48
Εικόνα 9. Μοντέλο προσομοίωσης αποστολής ασθενοφόρου ή κινητής ιατρικής μονάδας.....	49
Εικόνα 10. Τυχαίο παράδειγμα εισαγωγής δεδομένων σε κάρτα του προγράμματος...51	
Εικόνα 11. Τυχαίο παράδειγμα ανοιχτής κάρτας από την οπτική του ασυρματιστή.....	52
Εικόνα 12. Παράδειγμα συστήματος τηλεματικής τεχνολογίας στον τομέα των Αθηνών.....	54
Εικόνα 13. Μελέτη περίπτωσης Νομός Θεσσαλονίκης.....	56

## Πίνακες

Πίνακας 1. Μέσα Διακομιδής Εθνικού Κέντρου Άμεσης Βοήθειας.....	46
Πίνακας 2. Δεδομένα εισερχόμενων κλήσεων στο Ε.Κ.Α.Β. Θεσσαλονίκης.....	62
Πίνακας 3. Δεδομένα ολοκλήρωσης μεταφοράς διακομιδών από το Ε.Κ.Α.Β. Θεσσαλονίκης.....	62
Πίνακας 4. Αποτελέσματα υπολογισμού κινητού μέσου όρου, κεντραρισμένου κινητού μέσου όρου, όσο αναφορά τα δεδομένα κλήσεων στο Ε.Κ.Α.Β. Θεσσαλονίκης.....	64
4. Αποτελέσματα υπολογισμού κινητού μέσου όρου, κεντραρισμένου κινητού μέσου όρου, όσο αναφορά τα δεδομένα ολοκλήρωσης διακομιδών από το Ε.Κ.Α.Β. Θεσσαλονίκης.....	64

## Γραφήματα

Γράφημα 1. Σύνολο διακομιδών αιμοκαθαιρόμενων ασθενών.....	57
Γράφημα 2. Διάγραμμα διασποράς διακομιδών σε σχέση με την ηλικία των ασθενών από 9 έως 95 έτη για τον νομό Θεσσαλονίκης .....	58
Γράφημα 3. Ποσοστό έγκυρων και άκυρων διακομιδών.....	58
Γράφημα 4. Σύνολο διακομιδών ανάλογα με το φύλο.....	59
Γράφημα 5. Η γεωγραφική κατανομή των ασθενών που χρήζουν μεταφορά σε μονάδες τεχνητού νεφρού στον νομό Θεσσαλονίκης.....	59
Γράφημα 6. Σύνολο κλήσεων ανά βάρδια πίνακας για την περίοδο 01/08/2019 έως και 1/1/2021.....	60
Γράφημα 7. Σύνολο ολοκλήρωση διακομιδών ανά βάρδια από 01/08/2019 έως και 1/1/2021.....	62
Γράφημα 8. Διάγραμμα χρονοσειράς και προβλέψεων ενδεχομένων κλήσεων στο Ε.Κ.Α.Β. Θεσσαλονίκης για το τετράμηνο 01/12/2020 έως και 01/03/2021.....	65



Γράφημα 9. Διάγραμμα χρονοσειράς και προβλέψεων ολοκλήρωσης ενδεχόμενων διακομιδών από το Ε.Κ.Α.Β. Θεσσαλονίκης για το τετράμηνο 01/12/2020 έως και 01/03/2021.....66

## Συνομεύσεις

AI	Artificial intelligence
ANN	Artificial Neural Networks
ARIEN	AdapteR Interoperability Engine
AUC	Area Under the Curve
CAA	Crossbar Adaptive Array
CDO	Care Delivery Organization
CNN	Convolutional Neural networks
CRISP-DM	Cross-industry Standard Process for Data Mining
CT	Computed Tomography
EHR	Electronic Health Record
KDD	Knowledge Discovery and Data Mining
MDP	Markov Decision Process
LIS	Laboratory Information Systems
SNN	Simulated Neural Networks
SVM	Support Vector Machines
ML	Machine Learning
MLP	Multi-Layer Perceptron
MMSE	Mini-Mental State Examination
MRI	Magnetic Resonance Imaging
NLP	Natural Language Processing
NP	Network Propagation
PACS	Picture Archiving and Communications Systems
PCA	Principal Component Analysis
PHR	Personal Health Record
PMML	Predictive Model Markup
RELM	Regularized Extreme Machine Learning
RIS	Radiology Information Systems
RNN	Recurrent Neural Networks
SEMMA	Sample, Explore, Modify, Model, and Access
UIS	Ultrasound Information System

## **1. Εισαγωγή**

Οι αναλύσεις δεδομένων μέσω διαφόρων τεχνικών μπορούν να αποκαλύψουν χρήσιμες πληροφορίες. Το σύστημα υγείας που υπάγεται και το σύστημα προνοσοκομειακής φροντίδας είναι ένα από τα συστήματα που παράγει έναν τεράστιο όγκο δεδομένων σε καθημερινή βάση. Η χρήση αυτών των αναλύσεων μπορεί να βοηθήσει σημαντικά στην βελτίωση των υπηρεσιών που προσφέρει. (Pavlyshenko, 2016).

### **1.1 Περιγραφή προβλήματος**

Η νεφρική ανεπάρκεια είναι μια ασθένεια η οποία αυξάνεται ραγδαία σε παγκόσμιο επίπεδο (Carney, 2020). Η σοβαρότητα της κατάστασης των ασθενών που νοσούν από νεφρική ανεπάρκεια χρήζει άμεσης θεραπείας με την μέθοδο της αιμοδιάλυσης. Η αιμοδιάλυση πραγματοποιείται μέσω ειδικής μηχανικής υποστήριξης η οποία βρίσκεται σε μονάδα τεχνητού νεφρού εντός νοσοκομείου. Βάση του νόμου 2071 (ΦΕΚ 123, 1992) το Εθνικό Κέντρο Άμεσης Βοήθειας (Ε.Κ.Α.Β.) αναλαμβάνει τις διακομιδές των ασθενών που χρειάζονται μεταφορά σε μονάδα τεχνητού νεφρού. Συνήθως οι συνεδρίες είναι προγραμματισμένες, όμως αρκετές φορές παρατηρείται λόγω των επιπλοκών της νόσου, αύξηση των κλήσεων στο τηλεφωνικό κέντρο του Ε.Κ.Α.Β. για την πραγματοποίηση διακομιδών από και προς την οικία των ασθενών δεσμεύοντας ασθενοφόρα τα οποία προορίζονται για επείγοντα περιστατικά προκαλώντας μεγάλα κενά στην συνολική δύναμη των ασθενοφόρων και καθυστερήσεις, μειώνοντας την αποδοτικότητα των υπηρεσιών προνοσοκομειακής φροντίδας.

### **1.2 Στόχος μελέτης**

Στόχος της παρούσας διπλωματικής εργασίας είναι η διερεύνηση του κατά πόσο, μέσω ενός μοντέλου πρόβλεψης μπορούν να επιτευχθούν ορθά αποτελέσματα όσο αναφορά τις προβλέψεις κλήσεων που θα δεχθεί το τηλεφωνικό κέντρο του Ε.Κ.Α.Β. Θεσσαλονίκης όπως επίσης και των ολοκληρώσεων διακομιδών για μη επείγοντα περιστατικά τα οποία αφορούν ασθενείς που πάσχουν από νεφρική ανεπάρκεια.

### 1.3 Δομή διπλωματικής εργασίας

Στα πλαίσια της διπλωματικής εργασίας μελετήθηκαν διεθνείς δημοσιεύσεις στα κεφάλαια 2, 3 και 4 όπου παρουσιάζονται τεχνικές αναλύσεων δεδομένων (data analytics) με βασικούς πυλώνες την μηχανική μάθηση, την εξόρυξη δεδομένων και την προγνωστική μοντελοποίηση. Συγκεκριμένα στο κεφάλαιο 2 παρουσιάζονται οι προσεγγίσεις μηχανικής μάθησης, η τεχνητή νοημοσύνη, η βελτιστοποίηση, η στατιστική όπως επίσης οι τύποι μαθησιακών αλγορίθμων, η εποπτευόμενη μάθηση, η μη εποπτευόμενη μάθηση, η μάθηση ενίσχυσης, η μάθηση σποραδικών λεξικών και η ανίχνευση ανωμαλιών. Στη συνέχεια παρουσιάζονται τα μοντέλα των τεχνητών νευρωνικών δικτύων, τα δέντρα αποφάσεων, οι μηχανές φορέα υποστήριξης, η ανάλυση παλινδρόμησης και τα μπαγεζιανά δίκτυα. Σημαντική ανάλυση είναι οι περιορισμοί, οι αξιολογήσεις των μοντέλων όπως επίσης και η ηθική όσο αναφορά τις διαδικασίες μηχανικής μάθησης. Στο κεφάλαιο 3 ακολουθεί η ανάλυση της εξόρυξης δεδομένων και στο κεφάλαιο 4 του προγνωστικού μοντέλου.

Στο κεφάλαιο 5 πραγματοποιείται βιβλιογραφική ανασκόπηση των δεδομένων της υγείας, ξεκινώντας με τους κύριους τύπους και τις πηγές των δεδομένων στον τομέα της υγείας. Στην συνέχεια γίνεται ανάλυση των μεγάλων δεδομένων (Big Data) τα οποία παράγονται από τους ιατρούς και τις κλινικές, την δημόσια υγεία, τα ιατρικά πειράματα όπως επίσης τα μεγάλα δεδομένα στην ιατρική βιβλιογραφία των νοσοκομειακών πληροφοριακών συστημάτων και η εξέλιξη αυτών ενώ στο ίδιο κεφάλαιο ακολουθούν τα αναλυτικά δεδομένα και τα 9 V των μεγάλων δεδομένων (Big Data) στον τομέα της υγείας, ο όγκος (volume), η ταχύτητα (velocity), η ποικιλία (variety), η αξιοπιστία (veracity), η ισχύς (validity), η βιωσιμότητα (viability), η μεταβλητότητα (volatility), η ευπάθεια (vulnerability) και η οπτικοποίηση (visualization).

Στο κεφάλαιο 6 ακολουθούν οι ιδιαιτερότητες των μεγάλων δεδομένων στον τομέα της υγείας όπως είναι η μείωση κόστους υγειονομικής περίθαλψη, η ετερογένεια, τα ατελή δεδομένα, η επικαιρότητα, το απόρρητο των δεδομένων και η ιδιοκτησία.

Στο κεφάλαιο 7 παρουσιάζεται η μελέτη περίπτωσης με ονομασία: “Προγνωστικές αναλύσεις δευτερογενών δεδομένων μη επειγόντων περιστατικών μεταφοράς ασθενών σε μονάδες τεχνητού νεφρού από το Ε.Κ.Α.Β. Θεσσαλονίκης”. Η μελέτη ξεκίνα με την

ανάλυση του Εθνικού Κέντρου Άμεσης Βοήθειας (Ε.Κ.Α.Β.) ως βασικό πάροχο προνοσοκομειακής φροντίδας στην Ελλάδα. Στην συνέχεια παρουσιάζεται το πρόγραμμα καταχώρησης δεδομένων του Ε.Κ.Α.Β. με ονομασία «ηλεκτρονική διαχείριση ασθενών 166». Το κεφάλαιο 7 ολοκληρώνεται με την μελέτη περίπτωσης το υλικό και τις μεθόδους που χρησιμοποιήθηκαν όπως και τα αποτελέσματα τα οποία αφορούν προβλέψεις κλήσεων για διακομιδές ασθενών που χρειάζονται αιμοκάθαρση από την οικία τους προς μονάδα τεχνητού νεφρού, όπως επίσης προβλέψεις ολοκλήρωσης των διακομιδών με επιστροφή στην οικία τους.

Στο κεφάλαιο 8 ακολουθεί συζήτηση των αποτελεσμάτων ενώ στο κεφάλαιο 9 παρουσιάζονται τα συμπεράσματα της μελέτης μαζί με προτάσεις για μελλοντική μελέτη και έρευνα.

Τέλος παρουσιάζονται οι βιβλιογραφικές πηγές που χρησιμοποιήθηκαν για τη διεκπεραίωση της διπλωματικής εργασίας.

## 2. Τι είναι η μηχανική μάθηση

Ο T.Mitchell αναφέρει για την Μηχανική Μάθηση ML (Machine Learning) ότι “είναι η μελέτη λογισμικών αλγορίθμων, η οποία μπορεί να βελτιωθεί μέσω της εμπειρίας” (Mitchell, 1997) θεωρώντας την από πολλούς ερευνητές ένα τμήμα της Τεχνητής Νοημοσύνης AI (Artificial Intelligence). Ο όρος της ML δόθηκε το 1959 από τον Arthur Samuel. Το 1981 ο Bozinovski δίνει μια περιγραφή για το πως ένα νευρωνικό δίκτυο μαθαίνει να αναγνωρίζει 40 χαρακτήρες μέσω ενός τερματικού (terminal) Bozinovski (1981). Ο T.Mitchell δίνει επίσης τον εξής ορισμό: “Ένα πρόγραμμα υπολογιστή λέγεται ότι μαθαίνει από την εμπειρία E σε σχέση με κάποια κατηγορία εργασιών T και την απόδοση μέτρου P εάν η απόδοσή του σε εργασίες στο T, όπως μετριέται από το P, βελτιώνεται με την εμπειρία” (Mitchell, 1997).

Τα μοντέλα στα οποία χτίζονται οι αλγόριθμοι ML κάνουν χρήση συγκεκριμένων δεδομένων, τα ονομαζόμενα «δεδομένα εκπαίδευσης», έτσι μπορούν να κατασκευασθούν μοντέλα τα οποία μπορούν να προβλέπουν και να καταφέρνουν να παρουσιάζουν αποτελέσματα τα οποία βοηθούν στην λήψη αποφάσεων. Αλγόριθμοι ML χρησιμοποιούνται άμεσα σε αρκετές καθημερινές εφαρμογές, επί παραδείγματα ο διαχωρισμός επιθυμητών-ανεπιθύμητων e-mail (Koza et al, 1996).

Η στατιστική επιστήμη παρομοίως μπορεί και πετυχαίνει προβλέψεις με την χρήση ηλεκτρονικών υπολογιστών κάτι που χρησιμοποιείτε από την ML ως υποκατηγορία χωρίς αυτό να χαρακτηρίζει την ίδια παρακλάδι των στατιστικών επιστημών. Παράλληλα η μαθηματική βελτιστοποίηση συνεισφέρει στην ML με νέες μεθόδους και θεωρίες. Η συσχέτιση δε της εξόρυξης δεδομένων (Data Mining-DM) με την ML η οποία εστιάζεται σε αναλύσεις δεδομένων μέσω κάποιου είδους μάθησης θεωρείται άμεση. Σε θέματα που αφορούν επιχειρήσεις και στα προβλήματα που η εφαρμογή ML συνίσταται για την εύρεση λύσεων συχνά αναφέρεται ως προγνωστική ανάλυση (Predictive Analytics) (Friedman, 1998).

### 2.1 Επισκόπηση

Με σκοπό την αποπεράτωση συγκεκριμένων εργασιών ο εκάστοτε προγραμματιστής με εγκατεστημένα τα απαραίτητα προγράμματα και τροφοδοτώντας έναν ηλεκτρονικό υπολογιστή με τα κατάλληλα δεδομένα μπορεί να έχει ως αποτέλεσμα στην λειτουργία του υπολογιστή να αποκτήσει την ικανότητα να ανακαλύπτει τρόπους στο πως μπορεί

να κάνει αυτές τις διεργασίες αυτόματα. Σε περιπτώσεις που απαιτούνται απλές εργασίες τα απαραίτητα βήματα προγραμματίζονται μέσω αλγόριθμων, υποδεικνύοντας στον υπολογιστή χωρίς κάποιο είδος μάθησης το πώς να βρει την λύση. Για εργασίες με μεγαλύτερες απαιτήσεις είναι ποιο αποτελεσματικό να αναπτυχθεί μέσω ενός ηλεκτρονικού υπολογιστή αυτόματα ένας αλγόριθμος που θα αναπτύσσεται μόνος του βρίσκοντας τα απαραίτητα βήματα που απαιτούνται για την λύση ενός προβλήματος (Alpaydın, 2014).

Η ML έχει την ιδιότητα μάθησης με πολλές προσεγγίσεις εκτέλεσης εργασιών ακόμα και σε περιπτώσεις που ο αλγόριθμος που έχει προγραμματιστεί δεν ικανοποιεί πλήρως τις ανάγκες, όπως για παράδειγμα σε προβλήματα με μεγάλο αριθμό πιθανών απαντήσεων η υπόδειξη κάποιον από αυτών που είναι ήδη σωστές ως έγκυρες μπορεί να βελτίωση των αλγόριθμο κάνοντας τον να καταφέρει να βρει τρόπο να προσδιορίζει τις σωστές απαντήσεις, όπως συμβαίνει στα συστήματα εκπαίδευσης αναγνώρισης ψηφιακών χαρακτήρων με την χρήση της επέκτασης MNIST (Alpaydın, 2014).

## 2.2 Προσεγγίσεις μηχανικής μάθησης

Οι προσεγγίσεις μηχανικής μάθησης χωρίζονται σε τρεις γενικότερες κατηγορίες:

- Εποπτευόμενη μάθηση: Έχοντας ως στόχο να εγκατασταθεί ένας γενικός κανόνας που να χαρτογραφεί τις εισροές στις εξόδους στον ηλεκτρονικό υπολογιστή παρουσιάζονται παραδείγματα (όπου δίνονται από έναν «δάσκαλο») εισόδων με τις αντίστοιχα επιθυμητές εξόδους.
- Μη εποπτευόμενη μάθηση: Είναι το είδος της μάθησης όπου για να μπορέσει ο αλγόριθμος μάθησης, να βρει από μόνος του δομή στην είσοδο του δεν παρέχονται ετικέτες. Αυτό το είδος μάθησης μπορεί να είναι ένας στόχος/σκοπός από μόνος του (όπως είναι η ανακάλυψη κρυφών μοτίβων μέσα σε δεδομένα) ή το μέσο ως προς ένα τέλος (μάθηση χαρακτηριστικών).
- Μάθηση ενίσχυσης: Η μάθηση ενίσχυσης σχετίζεται με την αλληλεπίδραση ενός δυναμικού περιβάλλοντος το οποίο πρέπει να επιτελεί έναν συγκεκριμένο στόχο όπως για παράδειγμα τα παιχνίδια οδήγησης οχήματος. Καθώς πλοηγείτε κάποιος στον χώρο χωρίς να γνωρίζει ο χρήστης το τέλος της διαδρομής και ποιο θα είναι το αποτέλεσμα, το πρόγραμμα παρέχει

ανατροφοδότηση δίνοντας ανάλογες ανταμοιβές, τις οποίες προσπαθεί να μεγιστοποιήσει (Alpaydin, 2014).

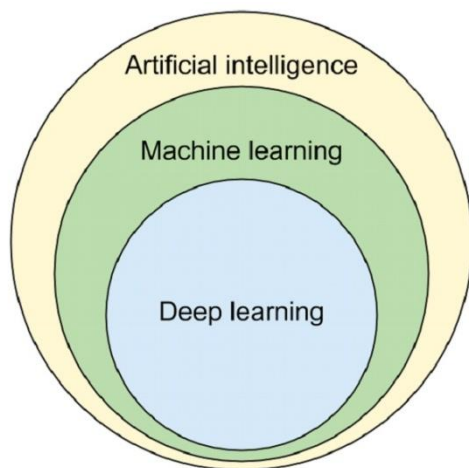
Επιπροσθέτως ένα άλλο είδος μάθησης το οποίο εντοπίζεται μεταξύ της εποπτευόμενης και της μη εποπτευόμενης μάθησης, είναι η ημι-εποπτευόμενη μάθηση όπου λόγω της έλλειψης σημάτων κατάρτισης στην εκπαίδευση πολλές από τις εκροές στόχους συχνά χάνονται. Αξίζει να σημειωθεί η ύπαρξη αναπτυξιακών μαθήσεων βασιζόμενων σε επαγωγικές προκαταλήψεις οι οποίες στηρίζονται σε προηγούμενες εμπειρίες. Παράδειγμα αναπτυξιακής μάθησης είναι η χρήση της στα ρομπότ όπου δημιουργεί τη δική της ακολουθία μέσω ενεργούς μάθησης, ωριμότητας, κινητής συνεργασίας και μίμησης μέσω της χρήσης μηχανισμών καθοδήγησης από τους εκπαιδευτές τους (Bishop, 2006).

### **2.3 Τεχνητή νοημοσύνη**

Ο τομέας της ΑΙ είναι πλέον ένας επιστημονικός κλάδος παγκόσμιου βεληνεκούς. Οι Haenlein και Kaplan, δίνουν τον δικό τους ορισμό ΑΙ “ικανότητα ενός συστήματος να ερμηνεύει σωστά τα εξωτερικά δεδομένα, μαθαίνει και χρησιμοποιεί αυτά τα για να επιτύχει συγκεκριμένους στόχους μέσω της ευέλικτης προσαρμογής” (Haenlein&Kaplan, 2019).

Από την μεριά τους οι Poole και Mackworth (2010) την αναφέρουν ως “το πεδίο που μελετά τη σύνθεση και την ανάλυση υπολογιστικών παραγόντων που ενεργούν έξυπνα”.

Η IBM έκανε προβλέψεις για το αν η ML έως το 2020 θα οδηγηθεί να βρίσκεται ως ένα υποπεδίο της ΑΙ (IBM 2020). Η σύγκυση που προκλήθηκε μεταξύ ΑΙ και ML εστιάζεται στο γεγονός ότι όποιος εργάζεται με εργαλεία ML θα μπορούσε αυτόματα να ταυτιστεί με τον τομέα της ΑΙ χωρίς να υπάρχει σαφήνεια στο αν εργάζεται για ML ή για ΑΙ ενώ παράλληλα διατυπώνεται η άποψη ότι η ML είναι ένα «έξυπνο» υποσύνολο της ΑΙ (Εικόνα 1) (Sindhu etal, 2020).



Εικόνα 1. Η ML ως μέρος της AI. Ανατύπωση από Sindhu etal (2020).

## 2.4 Βελτιστοποίηση

Σε ένα σύνολο δεδομένων που εκπαιδεύονται η βελτιστοποίηση βοηθά την ML στην ελαχιστοποίηση λειτουργικών απωλειών όσο αναφορά τις προβλέψεις μέσω συναρτήσεων οι οποίες εκφράζουν την ασυμφωνία των μοντέλων εκπαίδευσης σε περιπτώσεις που εμφανίζονται προβλήματα προγραμματισμού. Για παράδειγμα ένας προγραμματιστής θέλει να κάνει αντιστοιχίσεις από μια ετικέτα σε πολλές περιπτώσεις. Αυτό το οποίο μπορεί να κάνει είναι να δημιουργήσει μοντέλα που εκπαιδεύονται ούτως ώστε να μπορούν να προβλέπουν ορθά ποιες θα είναι οι προκαθορισμένες ετικέτες σε ένα σύνολο παραδειγμάτων (LeRoux etal, 2012).

## 2.5 Στατιστική

Η επιστήμη της στατιστικής χρησιμοποιεί δείγματα πληθυσμών για να μπορέσει να οδηγηθεί σε συμπεράσματα τα οποία μπορούν να χρησιμοποιηθούν για να γίνουν στοχευόμενες προβλέψεις κάτι το οποίο δεν συμβαίνει με τις προβλέψεις της ML αν και οι μέθοδοι που χρησιμοποιούνται έχουν πολλές ομοιότητες διότι συνήθως τα αποτελέσματα προβλέψεων της ML είναι γενικευμένα (Bzdok etal, 2018).

Ο Leo Breiman παρουσιάζοντας μοντέλα στατιστικής μοντελοποίησης ξεχώρισε τα μοντέλα με βάση τα δεδομένα και μοντέλα που προέρχονται από αλγόριθμους όπως είναι το Random Forest δηλαδή αλγόριθμους ML (Breiman, 2001).

Τα δεδομένα και η βελτίωση της απόδοσης των αναλύσεων τους βασίζονται σε μεγάλο βαθμό στο κατά πόσο η πολυπλοκότητα των υποθέσεων συμβαδίζουν με την



πολυπλοκότητα των συναρτήσεων και αυτό μπορούμε να το κατανοήσουμε με διάφορους τρόπους όπως για παράδειγμα για υποθέσεις που η πολυπλοκότητα τους είναι μικρή τότε θα μπορούσαμε να πούμε ότι τα δεδομένα έχουν προσαρμοστεί στο μοντέλο το οποίο δημιουργήθηκε ή στην περίπτωση που η πολυπλοκότητα του μοντέλου αυξάνεται, τότε αυτό που συμβαίνει είναι να παρατηρούμε μείωση των σφαλμάτων της εκπαίδευσης ή στην περίπτωση όπου έχουμε μεγάλη πολυπλοκότητα σε σχέση με την υπόθεση το μοντέλο που δημιουργήθηκε υπόκειται σε υπερβολική εφαρμογή και αυτό έχει σαν αποτέλεσμα μια φτωχότερη γενίκευση (Alpaydin, 2014).

## **2.6 Τύποι μαθησιακών αλγορίθμων**

### **2.6.1 Εποπτευόμενη μάθηση**

Οι εποπτευόμενοι μαθησιακοί αλγόριθμοι έχουν την ιδιότητα να δημιουργούν μοντέλα που μπορούν να μας βοηθήσουν στο να αναπαράξουμε επιθυμητά αποτελέσματα μέσω δεδομένων που τους παρέχουμε (Russell&Norvig, 2010). Για να μπορέσει να εκπαιδευθεί ο αλγόριθμος χρησιμοποιείται ένα παράδειγμα το οποίο αποτελείται από μια ή περισσότερες εισόδους δεδομένων όπως επίσης και μια έξοδο η οποία ονομάζεται σήμα εποπτείας από όπου αντλούμε τα επιθυμητά αποτελέσματα. Όσο αναφορά το μαθηματικό μοντέλο και τα προς εκπαίδευση δεδομένα τα παραδείγματα σε έναν διανυσματικό χώρο αντιπροσωπεύονται από έναν πίνακα. Επίσης μέσω των αλγορίθμων της εποπτευόμενης μάθησης μπορούν να προβλεφθούν παραγόμενες νέες εισοδοί μέσω επαναληπτικών μεθόδων βελτιστοποίησης μιας αντικειμενικής συνάρτησης (Mohri et al, 2012). Συνεπώς ο αλγόριθμος μέσω βέλτιστων λειτουργιών μπορεί πλέον να καθορίσει ορθά στην έξοδο άλλες εισόδους οι οποίες δεν είναι μέρος των εκπαιδευτικών δεδομένων κάτι το οποίο συνεπάγεται ότι βελτιώνει την ακρίβεια των προβλέψεων στις εξόδους συμπεραίνοντας ότι με την πάροδο του χρόνου έμαθε να εκτελεί αυτήν την εργασία (Mitchell, 1997).

Σε αυτούς τους τύπους των αλγορίθμων περιλαμβάνονται η ταξινόμηση και η παλινδρόμηση. Επί παραδείγματι, ένας αλγόριθμος ταξινόμησης που φιλτράρει τα email, η είσοδος θα είναι ένα εισερχόμενο ηλεκτρονικό μήνυμα (email) με έξοδο την ονομασία του φακέλου στον οποίο υποβάλλεται το ηλεκτρονικό μήνυμα (Alpaydin, 2014).

### **2.6.2 Μη εποπτευόμενη μάθηση**

Λαμβάνοντας δεδομένα εκπαίδευσης που έχουν μόνο εισόδους οι αλγόριθμοι της μη εποπτευόμενης μάθησης μαθαίνουν να βρίσκουν δομές μέσα σε αυτά που δεν έχουν επισημανθεί, ταξινομηθεί ή κατηγοριοποιηθεί. Συνεπώς οι συγκεκριμένοι αλγόριθμοι μπορούν να βρίσκουν κοινά χαρακτηριστικά στα δεδομένα και κάθε φορά που ένα νέο κομμάτι δεδομένων εμφανιστεί, ο αλγόριθμος αντιδρά σε σχέση με την απουσία ή παρουσία τους. Η συνάρτηση της πυκνότητας πιθανότητας για την εκτίμηση της πυκνότητας είναι μια από τις εφαρμογές της μη εποπτευόμενης μάθησης (Jordan et al, 2004).

### **2.6.3 Μάθηση ενίσχυσης**

Το χαρακτηριστικό αυτού του είδους μάθησης είναι η ενασχόληση παραγόντων λογισμικών που ασχολούνται με την διαδικασία λειτουργιών σε ένα περιβάλλον όπου κύριος στόχος είναι η αθροιστική έννοια της ανταμοιβής. Λόγου όμως της γενικότητας αυτών των διαδικασιών μάθησης μελετάται και από πολλούς άλλους κλάδους. Στην ML θα την συναντήσουμε συνήθως ως διαδικασία απόφασης Markov MDP (Markov decision process) όπου και χρησιμοποιούνται τεχνικές δυναμικού προγραμματισμού στους αλγόριθμους αυτού του είδους μάθησης (van Otterlo & Wiering, 2012).

### **2.6.4 Αυτομάθηση**

Η αυτομάθηση είναι το είδος της μάθησης όπου ο αλγόριθμος της είναι ο CAA (Crossbar adaptive array) ενός νευρωνικού δίκτυου που μπορεί και μαθαίνει από μόνο του μέσω παραδειγμάτων ML. Αυτός ο αλγόριθμος υπολογίζει καταστάσεις συνεπειών μέσω μιας εγκάρσιας γραμμής (Crossbar) και μιας αλληλεπίδρασης ενεργειών και αισθημάτων με σκοπό να παραχθεί ένα σύστημα που να μπορεί να παίρνει αποφάσεις σε θέματα που σχετίζονται με γνώσεις και αισθήματα και να καθοδηγείτε από αυτό (Bozinovski, 2014).

### **2.6.5 Μάθηση χαρακτηριστικών**

Κατά την διάρκεια της εκπαίδευσης αλγόριθμων είναι λογικό οι αλγόριθμοι να στοχεύουν στην καλύτερη ανακάλυψη αναπαραστάσεων που έχουν να κάνουν με τις εισροές που τους παρέχονται (Bengio et al, 2013). Για τα χαρακτηριστικά χρησιμοποιούνται ετικέτες εισόδου στα δεδομένα εισόδου όπου τα παραδείγματα

αυτών των ετικετών περιλαμβάνουν τεχνητά νευρικά δίκτυα και λεξικά όσο αναφορά την εποπτευόμενη μάθηση. Για τα χαρακτηριστικά της μη εποπτευόμενης μάθησης στα δεδομένα δε χρησιμοποιείται ετικέτα εισόδου. Τα παραδείγματα χρησιμοποιούν την μάθηση λεξικών, τους αυτόματους κωδικοποιητές και άλλες διάφορες μορφές ομαδοποίησης (Coates et al, 2011).

Οι πολλαπλοί μαθησιακοί αλγόριθμοι προσπαθούν να το πράξουν κάτω από τον περιορισμό ότι η μαθησιακή αναπαράσταση είναι χαμηλής διάστασης. Οι αλγόριθμοι που επεξεργάζονται υψηλού επιπέδου δεδομένα δέχονται τον περιορισμό αναπαραστάσεων χαμηλής διάστασης και αυτό έχει σαν αποτέλεσμα στην κωδικοποίηση να παρατηρείται αεριοποίηση της αναπαράστασης κάτι το οποίο οδηγεί στο συμπέρασμα ότι το μοντέλο έχει πολλά μηδενικά. Από την άλλη οι πολυεπίπεδοι αλγόριθμοι υποδιαστημάτων στοχεύουν την μάθηση τους κατευθείαν σε αναπαραστάσεις τανυστών χαμηλής διάστασης για δεδομένα με πολλές διαστάσεις χωρίς να μετατρέπονται και να αναδιαμορφώνουν αναπαραστάσεις υψηλότερης διάστασης (Lu et al, 2011). Όσο αναφορά την βαθιά μάθηση οι αλγόριθμοι έχουν την ιδιότητα να ανακαλύπτουν πολλαπλά επίπεδα αναπαραστάσεων με δεδομένα υψηλότερου επιπέδου έχοντας χαρακτηριστικά αφηρημένα τα οποία ορίζονται με όρους δεδομένων ως λειτουργίες χαμηλότερων επιπέδων. Μια μηχανή θα μπορούσε να θεωρηθεί έξυπνη όταν μαθαίνει και μπορεί να αποσυνδέσει τους υποκείμενους παράγοντες παραλλαγής που εξηγούν τα παρατηρούμενα δεδομένα που αναπαρίστανται (Bengio, 2009).

### **2.6.6 Μάθηση σποραδικών λεξικών**

Ένα ακόμη είδος μάθησης είναι αυτό των σποραδικών λεξικών. Είναι μια μέθοδος μάθησης χαρακτηριστικών όπου θεωρείται ότι είναι ένας αραιός πίνακας για τον λόγο του ότι ένα παράδειγμα εκπαίδευσης παρουσιάζεται ως γραμμικός συνδυασμός συναρτήσεων. Η μέθοδος που χρησιμοποιείται είναι έντονα NP (network propagation) και παράλληλα πολύ δύσκολο να λυθεί (Tillmann, 2015). Ο αλγόριθμος K-SVD (K-singular value decomposition) μπορεί να θεωρηθεί ως μια από τις πιο γνωστές ευρετικές μεθόδους όσο αναφορά την σποραδική μάθηση λεξικών. Στην ταξινόμηση, το πρόβλημα είναι να προσδιοριστεί η τάξη στην οποία ανήκει ένα προηγουμένως άορατο παράδειγμα εκπαίδευσης. Αυτό το είδος μάθησης έχει εφαρμοστεί επίσης και σε περιπτώσεις αποσυγκρότησης εικόνων. Η ιδέα είναι ότι ένα απόσπασμα εικόνας το

οποίο είναι κανονικό μπορεί μέσα από ένα λεξικό εικόνων να αραιώνει την μορφή που παρουσιάζεται (Aharon etal, 2006).

### **2.6.7 Ανίχνευση ανωμαλιών**

Ένα σημαντικό θέμα που προκύπτει είναι η ανίχνευση ανωμαλιών σε δεδομένα που διαφέρουν στην πλειοψηφία τους και εγείρουν υποψίες όσο αναφορά γεγονότα και παρατηρήσεις (Zimek&Schubert, 2017). Κάποιο ζήτημα προκύπτει όταν παρουσιάζονται ακραίες τιμές και αποκλίσεις όπως για παράδειγμα σε ένα κείμενο που παρουσιάζει σφάλματα ακόμα και σε κάποιο είδος ιατρικού προβλήματος (Hodge&Austin, 2004).

Ποιο συγκεκριμένα στα πλαίσια ανίχνευσης διαδικτυακών εισβολών αυτό που παρατηρείται στην ανάλυση είναι ότι αντί να εμφανίζονται σπάνια αντικείμενα που θα έδειχναν ότι κάποιος κάνει επίθεση σε κάποιο συγκεκριμένο σύστημα ή πρόγραμμα, αντ' αυτού χρονικά έχουμε ένα είδος απροσδόκητης αδράνειας. Η μέθοδος outlier (αλγόριθμοι μη εποπτευόμενοι) η οποία χρησιμοποιείτε σε πολλές μεθόδους ανίχνευσης σπάνιων αντικειμένων αποτυχαίνει να πετύχει αναγνώριση τέτοιων δεδομένων εκτός εάν έχουν συγκεντρωθεί σωστά τα δεδομένα. Ένας αλγόριθμος συστάδων από την άλλη θα μπορούσε να ανιχνεύσει μοτίβα τα οποία προέρχονται από μικροσυστήματα. (Dokas etal 2002).

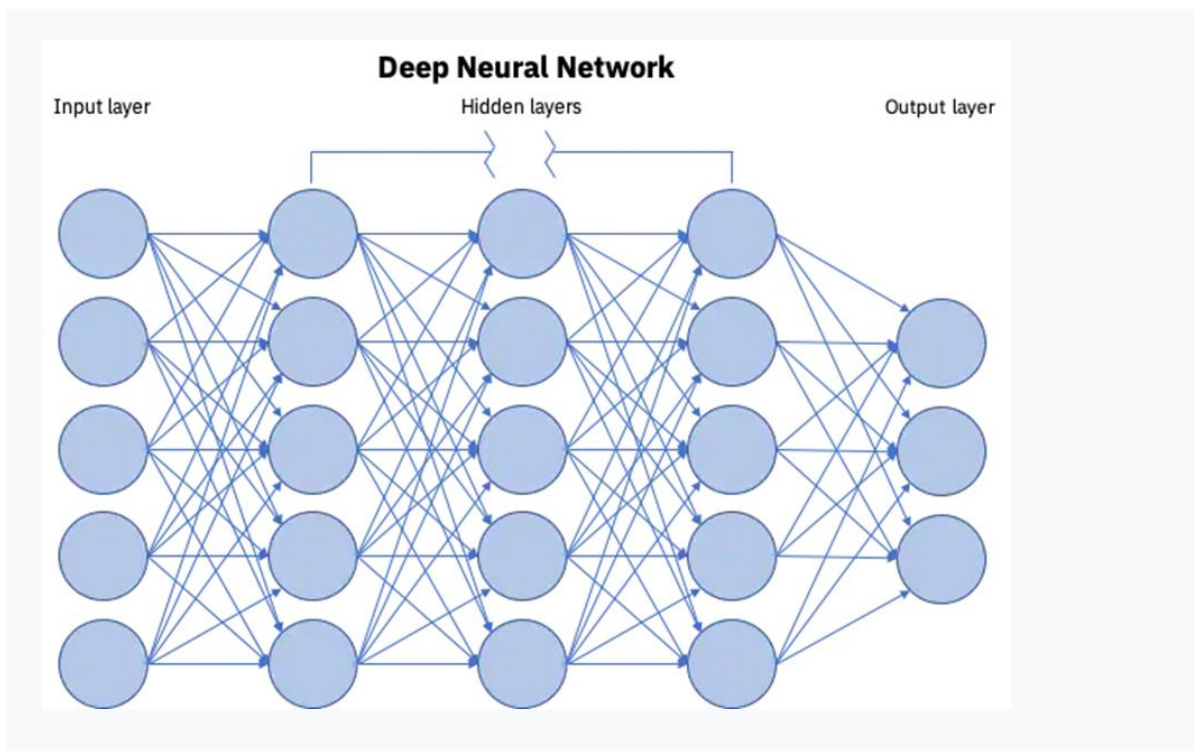
## **2.7 Μοντέλα**

### **2.7.1 Τεχνητά νευρωνικά δίκτυα**

Ένα υποσύνολο της ML και επίκεντρο αλγορίθμων βαθιάς μάθησης είναι τα τεχνητά νευρωνικά δίκτυα ANN (Artificial neural networks) ή αλλιώς προσομοιωμένα νευρωνικά δίκτυα SNN (Simulated neural networks). Η ονομασία και η δομή τους απορρέουν από μια λειτουργία του ανθρωπίνου εγκεφάλου την αλληλοενημέρωση των νευρώνων (IBM Cloud Education 2020).

Ένα ANN χωρίζεται σε ένα επίπεδο εισόδου, ένα ή περισσότερα κρυμμένα επίπεδα και μιας εξόδου. Κάθε τεχνητός νευρώνας συνδέεται με έναν άλλο κόμβο και έχει σχετικό βάρος (Εικόνα 2). Εάν η έξοδος ξεπεράσει τα όρια οποιουδήποτε κόμβου τότε αυτός ο κόμβος ενεργοποιείται με αποτέλεσμα να αποστέλλει τα δεδομένα στο

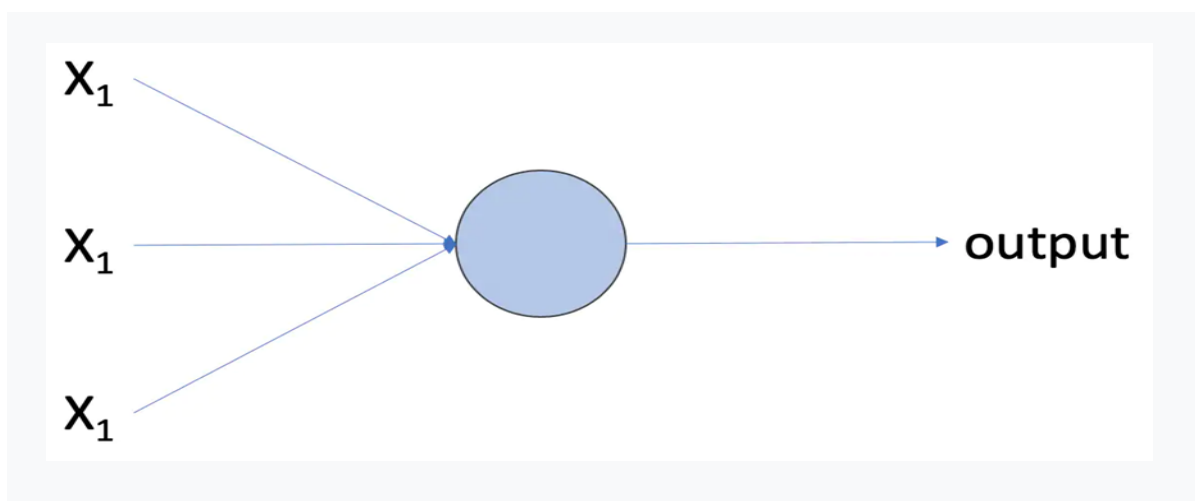
επόμενο επίπεδο. Αντίθετα αν δεν ξεπεράσει τα δεδομένα δεν μεταδίδονται στο επόμενο επίπεδο δικτύου (IBM Cloud Education 2020).



Εικόνα 2. Βαθιά Νευρωνικά Δίκτυα. Ανατύπωση από IBM Cloud Education (2020).

Τα νευρωνικά δίκτυα ανάλογα την χρησιμότητα τους ταξινομούνται σε διαφορετικούς τύπους:

Το perceptron είναι το παλαιότερο νευρωνικό δίκτυο, δημιουργήθηκε από τον Frank Rosenblatt το 1958. Αποτελείται από έναν νευρώνα και χαρακτηρίζεται ως την απλούστερη μορφή ενός νευρικού δικτύου (Εικόνα.3):



Εικόνα 3. Perceptron. Ανατύπωση από IBM Cloud Education (2020).

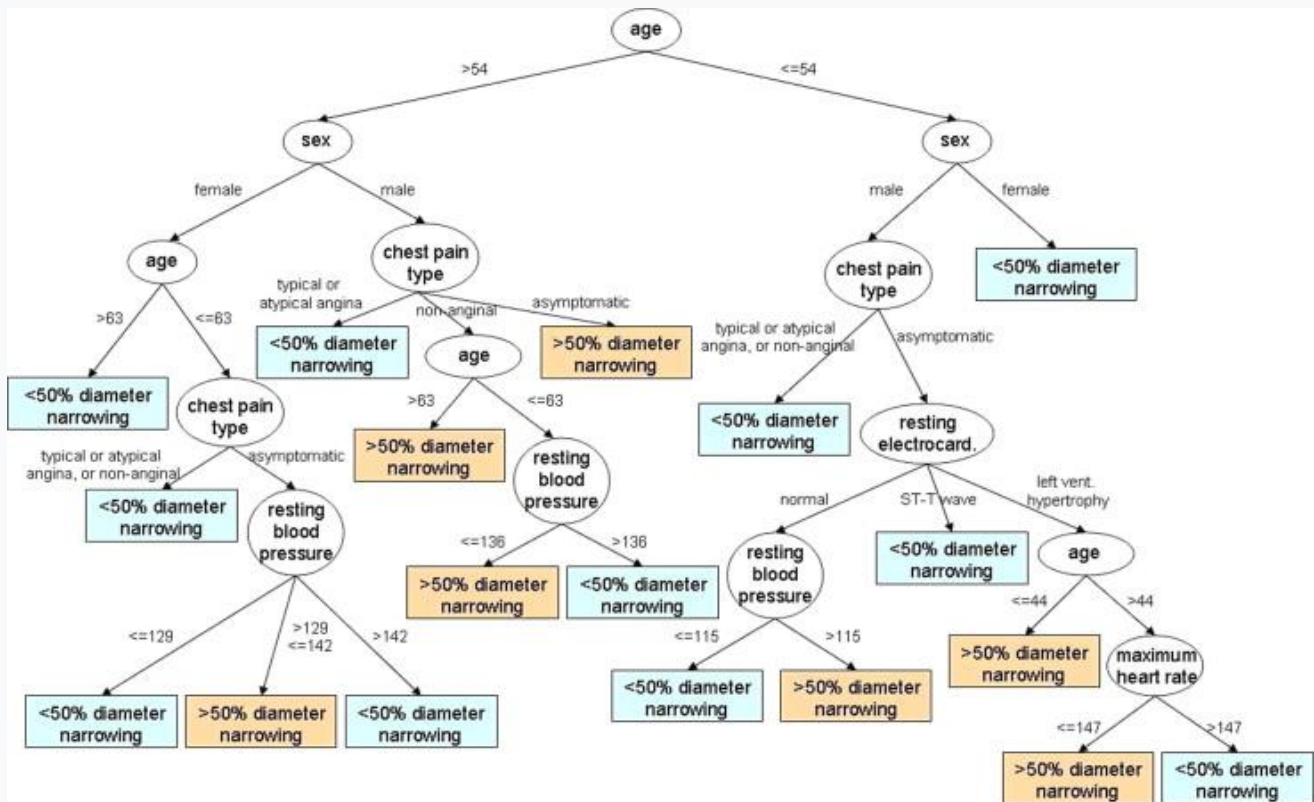
Τα τροφοδοτικά νευρικά δίκτυα (Feedforward) ή τα πολυεπίπεδα perceptrons MLP (Multi-layer perceptrons) αποτελούνται από μια είσοδο, ένα ή περισσότερα κρυφά στρώματα και μια έξοδο. Ενώ αυτά τα νευρωνικά δίκτυα αναφέρονται επίσης συνήθως ως MLP, είναι σημαντικό να σημειωθεί ότι στην πραγματικότητα αποτελούνται από σιγμοειδείς νευρώνες, όχι από perceptrons, καθώς τα περισσότερα προβλήματα στον πραγματικό κόσμο είναι μη γραμμικά. Πάνω σε αυτά τα μοντέλα τροφοδοτούνται και εκπαιδεύονται δεδομένα τα οποία αποτελούν την βάση για πολλές εφαρμογές όπως είναι η μηχανική όραση και η επεξεργασία της φυσικής γλώσσας (IBM Cloud Education 2020).

Ένα άλλο είδος νευρώνων είναι τα κυριαρχικά νευρωνικά δίκτυα CNN (Convolutional neural networks) τα οποία είναι πανομοιότυπα με τα τροφοδοτικά δίκτυα και συνήθως χρησιμοποιούνται για την αναγνώριση εικόνων σε επίπεδο αναγνώρισης προτύπων όπου αξιοποιώντας αρχές της γραμμική άλγεβρας μπορούν και προσδιορίζουν μοτίβα μέσα σε μια εικόνα (IBM Cloud Education 2020).

Τα επαναλαμβανόμενα νευρωνικά δίκτυα RNN (Recurrent Neural Networks) και οι αλγόριθμοι μάθησης που χρησιμοποιούνται βασίζονται πάνω σε δεδομένα χρονοσειρών ούτως ώστε να μπορούν και να κάνουν προβλέψεις, όπως για παράδειγμα τις τιμές των χρηματιστηριακών μετοχών ή τις προβλέψεις πωλήσεων (IBM Cloud Education 2020).

### **2.7.2 Δέντρα αποφάσεων**

Η αποτελεσματικότητα και η αξιοπιστία των ιατρικών αποφάσεων όπως είναι η διάγνωση ή η ταξινόμηση μπορεί να επιτευχθεί μέσω εννοιολογικών μοντέλων λήψης αποφάσεων τα οποία παρέχονται από τα δέντρα αποφάσεων (Εικόνα 4) λόγω της υψηλής ακριβείας που προσφέρει η ταξινόμηση του εν λόγω μοντέλου (Podgorelec et al, 2002).



Εικόνα 4. Παράδειγμα δέντρου αποφάσεων. Ανατύπωση από López-Vallverdú, etal, (2012).

### 2.7.3 Μηχανές φορέα υποστήριξης

Ένα μοντέλο το οποίο χρησιμοποιείται κυρίως στην παλινδρόμηση και στην ταξινόμηση είναι οι μηχανές φορέα υποστήριξης SVM (Support vector machines). Ένα SVM μέσω ενός συνόλου δεδομένων εκπαίδευσης που επισημαίνεται μια κατηγορία από δυο μπορεί να προβλέψει με τον κατάλληλο αλγόριθμο αν ταιριάζει σε μια από τις δυο κατηγορίες ένα νέο εισερχόμενο δείγμα δεδομένων (Cortes etal, 1995). Συνήθως για SVM χρησιμοποιείται ως αλγόριθμος εκπαίδευσης ένας μη πιθανός δυαδικός γραμμικός ταξινομητής όπως επίσης μπορεί να χρησιμοποιηθούν ταξινομητές σαν την κλιμάκωση Platt για την ρύθμιση πιθανοτήτων. Τα SVM πέραν της γραμμικής ταξινόμησης μπορούν να εκτελέσουν και μη γραμμική ταξινόμηση χαρτογραφώντας τις εισόδους τους σε υψηλών διαστάσεων δεδομένα αθόρυβα (Franc etal, 2011).

#### 2.7.4 Ανάλυση παλινδρόμησης

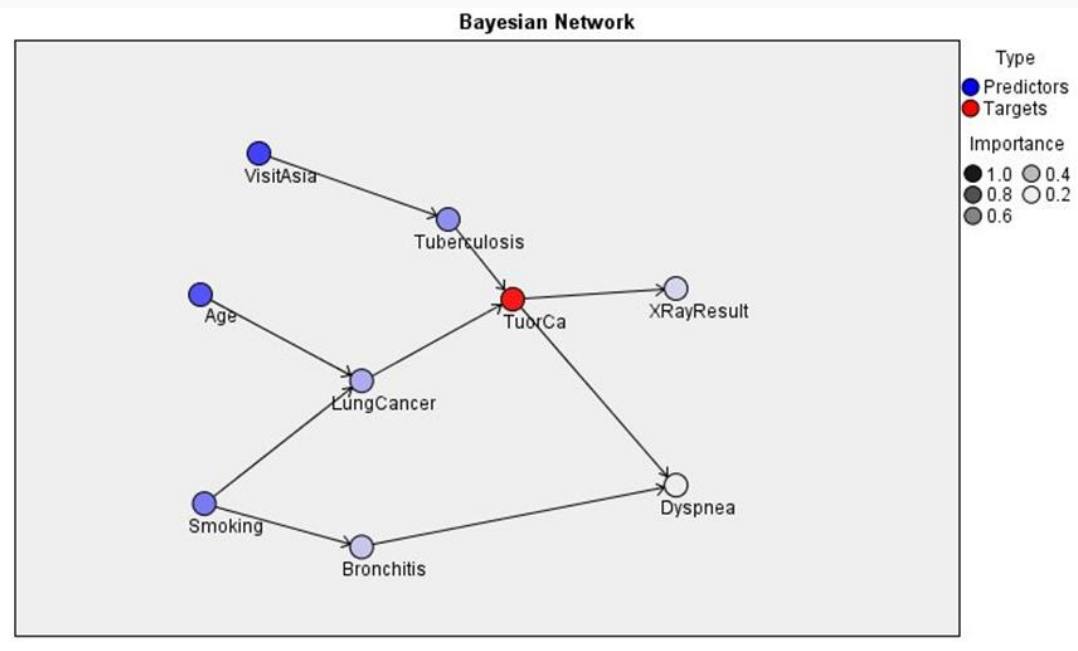
Η ανάλυση παλινδρόμησης περιλαμβάνει μια μεγάλη ποικιλία στατιστικών μεθόδων που σκοπό έχουν εκτιμήσεις σχέσεων για μεταβλητές εισόδου και των σχετικών χαρακτηριστικών τους. Η γραμμική παλινδρόμηση είναι η πιο συνήθης μορφή αυτού όπου σχεδιάζεται μια μονό γραμμή για να ταιριάζει καλύτερα στα δεδομένα σύμφωνα με ένα μαθηματικό κριτήριο τύπου ελαχίστων τετραγώνων. Σε προβλήματα μη γραμμικού χαρακτήρα όπως για παράδειγμα στο Microsoft Excel για την προσαρμογή της γραμμής τάσης προϋπάρχει μοντέλο πολυωνυμικής παλινδρόμησης το οποίο αντιστοιχεί τις μεταβλητές εισόδου (Stevenson, 2016).

#### 2.7.5 Μπαγεζιανά Δίκτυα

Το μπαγεζιανό δίκτυο (Bayesian Networks) είναι ένα γραφικό μοντέλο που εμφανίζει μεταβλητές σε ένα σύνολο δεδομένων το οποίο επεκτείνεται σε σχέσεις μεταξύ κόμβων επεξηγώντας αιτιώδες σχέσεις. Παράδειγμα ενός μπαγεζιανού δικτύου είναι να υπολογίζει ανάλογα τα συμπτώματα ενός ασθενούς το είδος της πάθησης που μπορεί να νοσήσει. Τα δίκτυα είναι πολύ ισχυρά όπου λείπουν πληροφορίες και κάνουν την καλύτερη δυνατή πρόβλεψη χρησιμοποιώντας οποιεσδήποτε πληροφορίες υπάρχουν (Stewart et al, 2015).

Οι Lauritzen & Spiegelhalter (1988) (Εικόνα 5) δημιούργησαν το μοντέλο «Ασία». Είναι μια έκδοση ενός δικτύου διάγνωσης νέων ασθενειών όπου η κατεύθυνση των συνδέσμων αντιστοιχεί περίπου στην αιτία. Κάθε κόμβος αντιπροσωπεύει συσχέτιση της κατάσταση ενός ασθενή όπως για παράδειγμα η όψη "Κάπνισμα" επιβεβαιώνει ότι ο ασθενής είναι καπνιστής και η εφαρμογή του "VisitAsia" παρουσιάζει το εάν ο ασθενής επισκέφτηκε πρόσφατα κάποια χώρα στην Ασία (IBM 2020). Οι σύνδεσμοι μεταξύ οποιουδήποτε κόμβου παρουσιάζουν την σχέση πιθανότητας όπως για παράδειγμα το κάπνισμα έχει αυξημένες πιθανότητες να εμφανίσει τόσο βρογχίτιδα όσο και καρκίνο του πνεύμονα ενώ αν ξεχωρίσουμε μόνο την ηλικία το δίκτυο παρουσιάζει συσχέτιση όσο αναφορά τον καρκίνο του πνεύμονα. Με την ίδια λογική οι ανωμαλίες που προκαλούνται από τις ακτινογραφίες των πνευμόνων μπορεί να έχουν βασικό παράγοντα, είτε την φυματίωση, είτε τον καρκίνο του πνεύμονα, ενώ οι πιθανότητες ενός ασθενή που πάσχει από δύσπνοια όπως επίσης εάν υποφέρουν από βρογχίτιδα ή από καρκίνο του πνεύμονα αυξάνονται (IBM 2020).





Εικόνα 5. Παράδειγμα δικτύου Ασία. Ανατύπωση από IBM(2020).

### 2.7.6 Περιορισμοί

Πολλά λογισμικά που έχουν δημιουργηθεί για την ML δεν παρουσιάζουν σωστά αποτελέσματα (Hall&Kumar, 2017) και ο λόγος είναι πολυπαραγοντικός. Είτε από έλλειψη δεδομένων, την πρόσβασης σε αυτά, τα προβλήματα απορρήτου όπως επίσης τίθεται θέμα εργασίας από την στιγμή που οι αλγόριθμοι που επιλέχθηκαν δεν είναι σωστοί και έχουν εισαχθεί με την λάθος σειρά (Artasanchez, 2018). Ένα από τα πιο τρανταχτά παραδείγματα είναι αυτό της IBM που δημιούργησε ένα πρόγραμμα εν ονόματι Watson που η αποτυχία του έγκειται στο γεγονός ότι παρόλο που επενδύθηκαν τεράστια χρηματικά ποσά δεν παρήγαγε ποτέ τα επιθήματα αποτελέσματα για τον τομέα της υγειονομικής φροντίδας. (Hernandez&Greenwald, 2018)

### 2.7.7 Αξιολογήσεις μοντέλου

Η αξιολόγηση μοντέλων ML εκτιμώνται με εκτιμήσεις ακριβείας οπου χωρίζοντας τα δεδομένα σε σύνολα δοκιμών δημιουργώντας εκτιμήσεις αποδόσεων των μοντέλων εκπαίδευσης. Άλλη μέθοδος είναι η K-fold-cross-validation η οποία μετά από τυχαίο διαχωρισμό των δεδομένων σε υποσύνολα K στα οποία στην συνέχεια πραγματοποιούνται πειράματα K. Το καθένα λαμβάνει υπόψη ένα υποσύνολο για

αξιολόγηση με τα υπόλοιπα υποσύνολα K-1 να λαμβάνονται υπόψη ξεχωριστά για την εκπαίδευση του μοντέλου. Επίσης το bootstrap λαμβάνοντας δείγματα  $n$  σε ένα σύνολο δεδομένων μπορεί να αξιολογήσει την ακρίβεια του μοντέλου (Kohavi, 1995).

Άλλοι παράμετροι είναι οι True Positive Rate (TPR) και True Negative Rate (TNR) για τον υπολογισμό της ευαισθησίας ενώ οι ερευνητές συχνά αναφέρουν και το ψευδώς θετικό ποσοστό False Positive Rate (FPR) καθώς και το ψευδώς αρνητικό ποσοστό False Negative Rate (FNR). Μια μέθοδος η οποία παρέχει εξαιρετικά αποτελέσματα είναι το συνολικό χαρακτηριστικό λειτουργίας Total Operating Characteristics (TOC) το οποίο μπορεί να διαγνώσει εάν ένα μοντέλο είναι αποτελεσματικό ή όχι. Το TOC εμφανίζει τους αριθμητές και τους παρονομαστές των τιμών παρέχοντας περισσότερες πληροφορίες μέσω του δέκτη receiver operating characteristic (ROC) και της σχετικής περιοχής κάτω από την καμπύλη area under the curve (AUC). (Pontius 2014).

### **2.7.8 Ηθική**

Τα ηθικά ερωτήματα που δημιουργούνται για το τομέα της ML εγείρονται κυρίως για συστήματα και αλγόριθμους που εκπαιδεύουν δεδομένα τα οποία περιέχουν προκαταλήψεις (Bostrom&Yudkowsky, 2011). Ένα λοιπόν κρίσιμο μέρος της ML είναι στο κατά πόσο θα συλλεχθούν υπεύθυνα και σωστά τα δεδομένα για να μπορούν να τεκμηριωθούν αλγοριθμικοί κανόνες ούτως ώστε χρησιμοποιηθούν από ένα σύστημα (Edionwe, 2017). Οι προκαταλήψεις που εκφράζονται από τους ανθρώπους μπορούν να οδηγήσουν σε ηθικούς περιορισμούς. Έτσι για οποιαδήποτε χρήση ML οι υπεύθυνοι των εταιρειών θα πρέπει να μάθουν στις μηχανές να εντοπίζουν αυτές τις ανθρώπινες προκαταλήψεις (Bostrom &Yudkowsky, 2011).

Ηθικές προκλήσεις εμφανίζονται και στον τομέα της υγειονομικής φροντίδας με την ανησυχία ότι μπορεί να σχεδιαστούν συστήματα τα οποία δεν θα λειτουργούν με άξονα το συμφέρον του κοινού καλού. Αυτό μπορεί να συμβεί σε χώρες όπως είναι η Αμερική όπου υπάρχουν αρκετοί λόγοι για κάποιους επαγγελματίες υγείας να θέλουν αυξήσουν τα κέρδη τους. Για παράδειγμα θα μπορούσαν κάποιιοι να σχεδιάσουν αλγορίθμους που να δίνουν στους ασθενείς περιττά φάρμακα ή εξετάσεις και στον κάτοχο του συστήματος μερίδια από αυτές τις αγορές. Η ML παρόλα αυτά παρέχει τεράστια βοήθεια στο υγειονομικό προσωπικό αφού χρησιμοποιούνται εργαλεία τα οποία βοηθούν τους επαγγελματίες υγείας να μπορούν να παίρνουν αποφάσεις π.χ. όσο

αφορά τις φαρμακευτικές αγωγές η πρόβλεψη αυτής της διαδικασίας θα μπορούσε να δώσει την δυνατότητα στους υπεύθυνους να διαλέξουν την πορεία αποκατάστασης μιας ασθένειας (Bostrom&Yudkowsky, 2011).

### **3. Τι είναι η εξόρυξη δεδομένων**

#### **3.1 Ετυμολογία**

Ως εξόρυξη δεδομένων Data Mining (DM) θα λέγαμε ότι είναι η ανακάλυψη προτύπων που διενεργούνται σε δεδομένα με μεγάλο όγκο μέσω διάφορων τεχνικών με στόχο να εξάγουμε πληροφορίες η οποίες είναι κατανοητές για τον κάθε χρήστη (Clifton, 2010) και βήματα ανάλυσης «ανακάλυψης γνώσεων σε βάσεις δεδομένων» ή αλλιώς KDD (Knowledge discovery and data mining) (Fayyad etal, 1996).

Οι Han&Kamber (2001) αναφέρουν ότι ο όρος είναι εσφαλμένος γιατί ο στόχος της DM είναι να μπορέσουμε να εξάγουμε γνώση από τα δεδομένα και όχι τα ίδια τα δεδομένα. Η DM παρουσιάζεται ως μια λέξη-κλειδί (Rome, 2005) για την επεξεργασία δεδομένων μεγάλης κλίμακας.

#### **3.2 Εξόρυξη δεδομένων**

Η εξόρυξη δεδομένων περιλαμβάνει έξι κοινές κατηγορίες εργασιών κατά τους Fayyad etal (1996):

Ανίχνευση ανωμαλιών: Αρχεία τα οποία παρουσιάζουν σφάλματα δεδομένων δημιουργώντας ενδιαφέρον περαιτέρω διερεύνησης.

Μάθηση κανόνων συσχετίσεων: Είναι η αναζήτηση σχέσεων μεταξύ μεταβλητών. Μια γνωστή διαδικασία αυτού του είδους είναι ανάλυση καλαθιού αγοράς που χρησιμοποιούν τα Super Market.

Συσταδοποίηση: Σε αυτό το είδος εργασίας δεν χρησιμοποιούνται γνωστές δομές στα δεδομένα με στόχο την εύρεση ομάδων και δομών σε αυτά τα δεδομένα.

Κατηγοριοποίηση: Μια εργασία η οποία εφαρμόζεται σε νέα δεδομένα με στόχο την γενίκευση γνωστών δομών π.χ. χαρακτηρισμός e-mail σε «επιθυμητό» ή «ανεπιθύμητο».

Παλινδρόμηση: Έχει ως στόχο την εύρεση συναρτήσεων στα δεδομένα με το μικρότερο σφάλμα.

Αυτόματη σύνοψη: Είναι η εργασία η οποία αναλαμβάνει να βρει όσο ποιο συμπαγές αναπαράσταση των δεδομένων με σκοπό την σωστή οπτικοποίηση τους με αναφορές.

### 3.3 Πρότυπα

Το 1999 εισήχθησαν δυο πρότυπα για την εξόρυξη δεδομένων Cross-industry standard process for data mining 1999 (CRISP-DM 1.0) και το πρότυπο Java Data Mining 2004 (JDM 1.0). Μέχρι το 2006 αυτά τα πρότυπα ήταν ενεργά (CRISP-DM 2.0 και JDM 2.0) μέχρι που αντικαταστάθηκαν από καινούργια πρότυπα (Günnemann et al, 2011).

Το βασικότερο πρότυπο προβλέψεων είναι η PMML (Predictive Model Markup Language) καλύπτοντας μοντέλα πρόβλεψης κύριος σε επιχειρηματικές εφαρμογές (Günnemann et al, 2011).

### 3.4 Απόρρητο και ηθική

Η ηθική στη DM προκύπτει από το γεγονός ότι εξάγονται πληροφορίες οι οποίες έχουν σχέση με τους ανθρώπους (Seltzer, 2005) κάτι το οποίο έχει άμεση σχέση με το απόρρητο και την ηθική. Το πρόγραμμα συνολικής ενημέρωσης πληροφοριών έχει εγείρει αρκετές φορές την ανησυχία σχετικά με το απόρρητο (Resig, 2004).

Η DM προϋποθέτει προετοιμασία δεδομένων, εκεί μπορεί κάποιος να έχει πρόσβαση σε πληροφορίες που μπορεί να θέσουν σε κίνδυνο το απόρρητο. Η συγκέντρωση δεδομένων συνεπάγεται συνδυασμό αυτών, με τρόπο που διευκολύνει την ανάλυση τους (NASCIO Research Brief, 2004). Το πρόβλημα εντοπίζεται όταν ένας ειδικός αναλυτής που έχει πρόσβαση σε νέα δεδομένα, μπορεί να αναγνωρίσει συγκεκριμένα άτομα από την στιγμή που αρχικά τα δεδομένα ήταν ανώνυμα (Ohm 2012).

Ο ειδικός αναλυτής μπορεί να αποκαλύψει προσωπικά στοιχεία παραβιάζοντας τις δίκαιες πρακτικές πληροφοριών προκαλώντας δυσάρεστες καταστάσεις στο υποδεικνυόμενο άτομο. Ένα παράδειγμα είναι η εταιρεία Walgreens όπου κάποιοι ειδικοί πούλησαν πληροφορίες συνταγών σε εταιρείες DM όπου οι ίδιες στην συνέχεια πούλησαν τα δεδομένα σε φαρμακευτικές εταιρείες (Kshetri, 2014).

Στα σύνολα των δεδομένων που δημοσιεύονται δεν αναφέρεται ποτέ η κατάσταση και η ταυτότητα του ασθενή. Παρόλα αυτά υπάρχει συζήτηση επί του θέματος να λαμβάνουν οι ειδικοί την συγκατάθεση του ασθενούς αφού πρώτα τον ενημερώσουν.

Για την χρήση ιατρικών δεδομένων υπάρχουν δυο μορφές η συστηματική εγγραφή επανεξέτασης και η καταγραφή σύνδεσης. Η συστηματική αναθεώρηση εγγραφών χρησιμοποιείται για τον έλεγχο εγγραφών μιας διαδοχικής σειράς κοινών χαρακτηριστικών ανάμεσα σε ασθενείς. Σύνδεση εγγραφής από την άλλη σημαίνει συλλογή ιατρικών πληροφοριών από ξεχωριστές πηγές σε μεμονωμένους ασθενείς που προσδιορίζονται με όλα τα προσωπικά στοιχεία και μεταξύ άλλων, οποιαδήποτε πιθανή συσχέτιση μεταξύ φαρμάκων και ασθενειών. Σε αυτή την περίπτωση είναι αναγκαία η προσωπική ταυτοποίηση στα αρχεία συλλογής δεδομένων κάτι το οποίο αυξάνει τον κίνδυνο απώλεια εμπιστευτικότητας (Ashwinkumar&Anandakumar, 2010).

## **4.Τι είναι το προγνωστικό μοντέλο**

### **4.1 Μοντέλα**

Οι προγνώσεις που προκύπτουν μέσω στατιστικών τεχνικών μπορούν να ονομαστούν προγνωστικά μοντέλα (Geisser, 1993). Η λογική βασίζεται εξ' ολοκλήρου στις μελλοντικές προβλέψεις ανεξάρτητα τότε πραγματοποιηθήκαν επί παραδείγματι σε δεύτερο χρόνο μετά την πράξη ενός εγκλήματος μπορούν να γίνουν προβλέψεις για τον εντοπισμό του ενόχου (Finlay, 2014).

Τα προγνωστικά μοντέλα χωρίζονται και αυτά σε δυο κατηγορίες τα παραμετρικά και τα μη παραμετρικά με τα ημι-παραμετρικά να βρίσκονται ενδιάμεσα συμπεριλαμβάνοντας χαρακτηριστικά και των δυο. Στην πρώτη περίπτωση των παραμετρικών μοντέλων οι παράμετροι που σχετίζονται με τις κατανομές είναι μια ή παραπάνω (Sheskin, 2011). Ενώ στα μη παραμετρικά μοντέλα οι παράμετροι είναι λιγότεροι και συνήθως περιλαμβάνουν παραδοχές που αφορούν την ανεξαρτησία (Cox, 2006).

### **4.2 Εφαρμογή στην υγειονομική περίθαλψη**

Ένα παράδειγμα προγνωστικής μοντελοποίησης είναι αυτό του Parkland Health & Hospital System το 2009 όπου έκανε χρήση προγνωστικών μοντέλων για τον εντοπισμό ασθενών με υψηλό κίνδυνο επανεισαγωγής. Ενώ στην αρχή ασχολήθηκε κυρίως με αναλύσεις για ασθενείς με συμφορητική καρδιακή ανεπάρκεια, ενώ στην συνέχεια επεκτάθηκε και σε άλλα είδη παθήσεων (U.S. Department of Health & Human Services, 2014).

## **5. Δεδομένα υγείας**

Τα δεδομένα της υγείας αφορούν γενικότερα την κατάσταση της υγείας ενός ατόμου ή πληθυσμού από την γέννηση έως τους λόγους θανάτου (McGraw-Hill, 2002). Αυτά τα δεδομένα αντλούνται κυρίως από κλινικές μετρήσεις. Πέραν τούτου μπορούμε να εντάξουμε μετρήσεις οι οποίες έχουν να κάνουν με την κατάσταση της κοινωνίας που ζούνε οι άνθρωποι το περιβάλλον κ.α. τα οποία έχουν άμεση σχέση με την υγεία και την ευεξία. Από την στιγμή που γίνεται η χρήση των υπηρεσιών υγείας τα δεδομένα αυτά συλλέγονται από τον εκάστοτε υγειονομικό σχηματισμό σε έναν φάκελο που περιλαμβάνει το αρχείο των παρεχόμενων υπηρεσιών τα κλινικά αποτελέσματα και γενικότερα πληροφορίες σχετικά με την εκάστοτε υπηρεσία (Tzourakis, 1996).

### **5.1 Μεγάλα δεδομένα (Big Data) στον τομέα της υγείας**

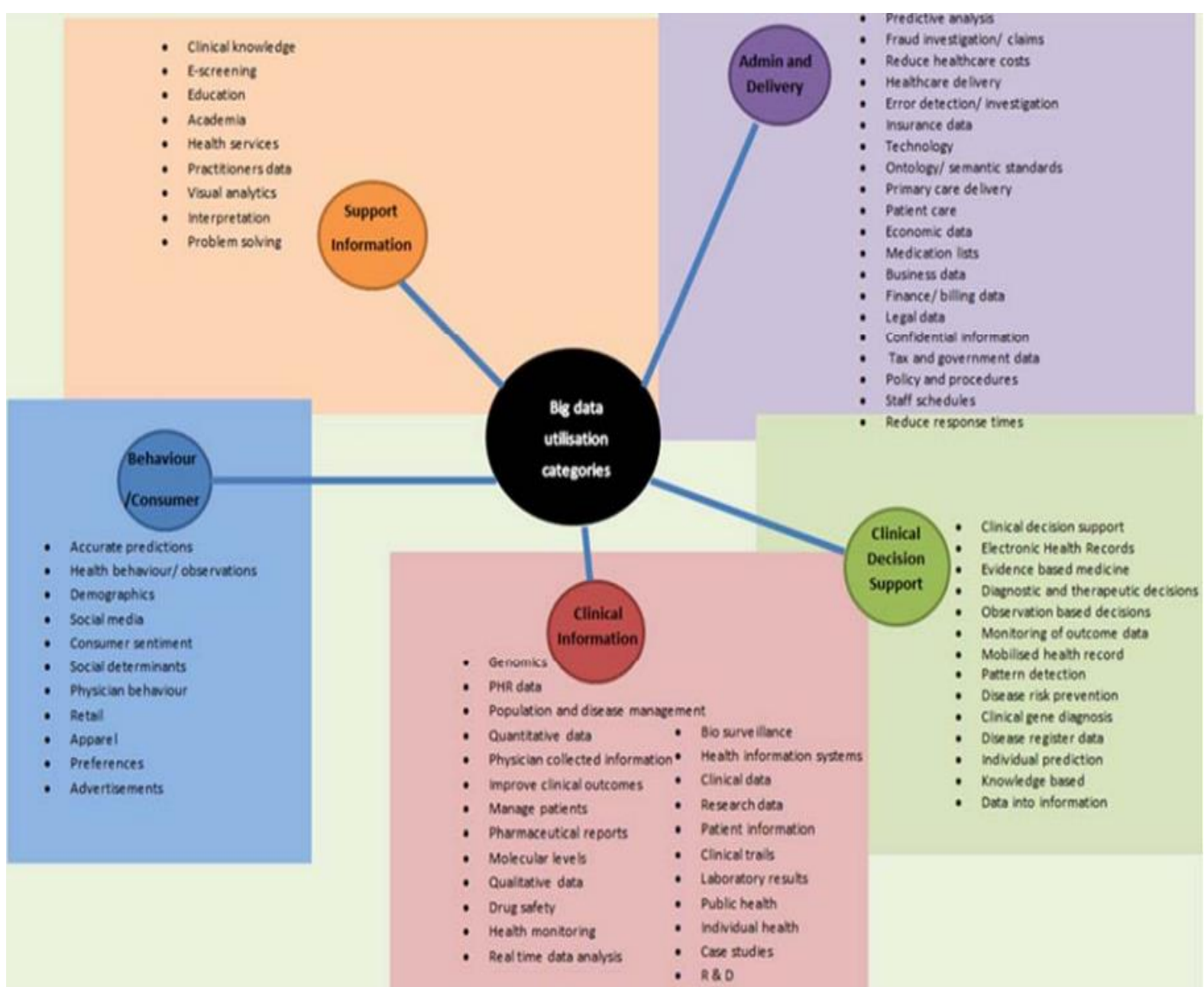
Τα Μεγάλα δεδομένα (Big Data) είναι ένα σύνολο δομημένων και μη δομημένων δεδομένων που από την ονομασία τους υποδηλώνουν ότι είναι πολλά και μεγάλα. Τα λογισμικά, οι αλγόριθμοι και οι βάσεις δεδομένων που είναι κατάλληλες για συλλογή, επεξεργασία, ανάλυση και αποθήκευση (Asante-Korang&Jacobs, 2016) μελετήθηκαν εκτενώς τα τελευταία χρόνια (Tan, Gao&Koch, 2015). Με όλη την ανάπτυξη που βιώνουμε καθημερινά ο όγκος των δεδομένων αυξήθηκε δραματικά. Τα Big Data δεν υπονοούν μόνο ότι είναι μεγάλα αλλά υποδηλώνουν την ικανότητα ταχείας επεξεργασίας δεδομένων και έναν νέο κόσμο τεχνολογικών προσεγγίσεων για το πως μπορούν να χειριστούν αυτά τα δεδομένα (Krumholz, 2014). Μέχρι στιγμής για τα Big Data έχουν δημιουργηθεί servers όπως είναι το γνωστό cloud το οποίο εκτός από την δυνατότητα αποθήκευσης διαθέτει και έναν τεράστιο εξοπλισμό δικτύωσης. Επίσης έχουν δημιουργηθεί λογισμικά τα οποία ανακτούν και αναλύουν τα δεδομένα με τεχνικές όπως η DM (Anderson & Chang, 2015).

#### **5.1.1 Κύριοι τύποι και πηγές μεγάλων δεδομένων στον τομέα της υγείας**

Εδώ και χρόνια έχει ξεκινήσει η συζήτηση του κατά πόσο ένα σύστημα υγείας είναι σημαντικό βάσει των λειτουργιών του όχι μόνο σε χώρες υψηλού εισοδήματος αλλά και σε χώρες μεσαίου εισοδήματος με προοπτικές επενδύσεων σε χώρες χαμηλού εισοδήματος (Kyoungyoung&Gang-Hoon,2013). Τα Big Data ταξινομούνται ανάλογα των πηγών όπου προέρχονται, για παράδειγμα τα Big Data της ιατρικής επιστήμης



φέρουν την ονομασία ιατρικά Big Data. Η Εικόνα 6 συνοψίζει τις πληροφορίες των κυριότερων τύπων Big Data στην υγεία.



Εικόνα 6. Οι κυριότερες κατηγορίες Big Data στην ιατρική βιβλιογραφία. Ανατύπωση από (Hermon&Williams, 2014).

### 5.1.2 Μεγάλα δεδομένα στην ιατρική και τις κλινικές

Τα Big Data έχουν μεγάλο όγκο δεδομένων τα οποία προέρχονται κυρίως από δομές υγείας και νοσοκομεία. Τα κλινικά δεδομένα και όλες οι απεικονίσεις δημιουργούν δεσμούς μεταξύ ασθενούς και ιατρού ενώ παράλληλα συνεισφέρουν άμεσα στην ιατρική βιομηχανία (Tsumoto et al, 2013). Η βοήθεια που προσφέρεται μέσω των Big Data είναι αξιοσημείωτη αφού δίνουν την δυνατότητα σε θέματα κλινικών αποφάσεων με εργαλεία επεξεργασίας κλινικών αποφάσεων CDS (Clinical decision support) όπως επίσης μπορούν να δώσουν την δυνατότητα στον ιατρό να επιλέξει και να σχεδιάσει την πορεία θεραπείας για μια ασθένεια βελτιώνοντας τα τεχνολογικά μέσα που χρησιμοποιούνται για αυτόν τον σκοπό (Kyoungyoung&Gang-Hoon, 2013).

Τα Big Data δημιουργούν μεγάλο όγκο πληροφοριών και αρχείων ανάλογα με τις πρακτικές που ακολουθούνται, για παράδειγμα για ένα προγραμματισμένο χειρουργείο τα δεδομένα που συλλέγονται είναι τα εξής:

- Αιματολογικές εξετάσεις
- Ακτινογραφίες
- Μαγνητικές τομογραφίες MRI (magnetic resonance imaging)
- Υπολογιστικές τομογραφίες CT (computed tomography)
- Πληροφορίες ασθενούς
- Φάρμακα που ήδη χρησιμοποιούνται ή θα χρησιμοποιηθούν
- Μελέτη θεραπείας
- Αποτελέσματα των ιατρικών πράξεων

(Tan et al, 2015, Wang & Alexander, 2013).

Τα δεδομένα προέρχονται από:

- το ηλεκτρονικό αρχείο υγείας EHR(Electronic health record)
- το ηλεκτρονικό ιατρικό αρχείο EMR(Electronic medical records)
- το προσωπικό αρχείο υγείας PHR (Personal health record) και μέσω των ιατρικών απεικονίσεων.

Η συνεισφορά του EMR έγκειται στο γεγονός ότι παρέχει πληροφορίες για οποιαδήποτε δραστηριότητα ιατρού και ασθενή σε ότι αναφορά την θεραπεία ενώ στο EHR εμπεριέχονται πληροφορίες που έχουν σχέση με την υγεία των ατόμων όπως επίσης ασφαλιστικές και οικονομικές πληροφορίες που έχουν να κάνουν γενικότερα με την υγειονομική φροντίδα των ασθενών (Wang&Alexander, 2013, Wu et al, 2016). Η βασικότερη διαφορά ανάμεσα στο EHR και EMR είναι ότι το EHR έχει την δυνατότητα να παρέχει και να διανέμει ιστορικές πληροφορίες που αφορούν τον ασθενή από την γέννηση μέχρι και τον θάνατο του και σε άλλους υγειονομικούς σχηματισμούς όπως επίσης και άλλα συστήματα υγείας ενώ το EMR είναι το αρχείο νόσησης του ασθενούς το οποίο είναι τοπικά αποθηκευμένο σε ένα νοσοκομείο. (Heart etal, 2017, Wang&Alexander, 2013). Έτσι συμπεραίνεται ότι το EHR επικεντρώνεται σε θέματα που αφορούν την υγεία κατοίκων μιας περιοχής μέσω δημογραφικών στοιχείων, ιατρικού ιστορικού κ.α. ενώ το EMR είναι στην ουσία το αρχείο που υπάρχει σε κάθε κράτος στην οργάνωση παροχής φροντίδας CDO (Care delivery organization) (Garets&Davis, 2006). Το 2012 είχε γίνει μια πρόβλεψη από τους Feldman etal ότι μέχρι το 2020 τα EHR από 500 petabytes θα φθάσουν τα 25.000 petabytes (Feldman etal, 2012).

Το PHR επίσης είναι πολύ σημαντικό μιας και παράγει ένα μεγάλο μέρος πληροφοριών τα οποία αφορούν την υγεία και την κοινωνία γενικότερα. Τα δεδομένα που παράγονται μέσω των PHR χρησιμοποιούνται κυρίως για ιατρικές αναλύσεις όπως επίσης λειτουργούν ως φορέας υποστηρίξεως λήψης κλινικών αποφάσεων για την πορεία κάποιας θεραπείας (Poulymenopoulou etal, 2015). Στο PHR περιλαμβάνονται δεδομένα ανεπιθύμητων ενεργειών ADRs (Allergies and adverse drug reactions) παρατηρήσεις της καθημερινής ζωής ODL (Observations of daily living) κ.α.. Σε αντίθεση με άλλα έγγραφα, οι ιατρικές απεικονίσεις προέρχονται κυρίως από ακτινογραφίες, CT, ιστολογία, PET, μαγνητικές τομογραφίες MRI, την φωτοακουστική

απεικόνιση την ηχοκαρδιογραφία κ.α. κάτι το οποίο συνεπάγεται ότι τα δεδομένα που παράγονται είναι συνήθως πολύ μεγάλα (Kovalev&Kalinovsky, 2015).

### **5.1.3 Μεγάλα δεδομένα στη δημόσια υγεία**

Στην αναπτυσσόμενη τεχνολογική περίοδο που ζούμε τα μεγάλα δεδομένα στην δημόσια υγεία συλλέγονται από φορητές συσκευές όπως για παράδειγμα οι διάφορες Health εφαρμογές που μπορεί να κατεβάσει οποιοσδήποτε στο έξυπνο κινητό του τηλέφωνο. Τέτοιες εφαρμογές είναι το Health Kit της Apple, το Google Fit της Google και το S Health της Samsung. Υπάρχουν όμως και συσκευές που με την χρήση αισθητήρων μπορούν να καταγράψουν δεδομένα τα οποία μπορούν να χρησιμοποιηθούν για την υγεία του χρήστη επί παραδείγματι τα Google Glasses και τα ειδικά ρολόγια Android. Το ηλεκτροκαρδιογράφημα είναι ένα άλλο παράδειγμα φορητής συσκευής που αποτελεί παράγοντα παροχής πληροφοριών για ένα ζωτικό μέλος του σώματος την καρδιά. Η δραστηριότητα της καρδιάς καταγράφεται με την μέτρηση του καρδιακού παλμού, έπειτα από τοποθέτηση ηλεκτροδίων στο δέρμα. Κατά την εγγραφή μετريέται σε πόσο χρόνο, πόσοι παλμοί δηλαδή σε ένα λεπτό κάνει η καρδιά. Από αυτές τις μετρήσεις μπορεί ο γιατρός να αντλήσει πληροφορίες και να εκτιμήσει την κατάσταση του ασθενούς (Safavi&Shukur, 2014).

Το ενδιαφέρον των ανθρώπων για την κατάσταση της προσωπικής τους υγείας αυξάνεται όσο αυξάνεται και το βιοτικό τους επίπεδο και αυτό συνεπάγεται σε ποιότητα της ζωής τους στην καθημερινότητα. Οι συσκευές ODL έχουν μεγάλο ρόλο στην καθημερινή καταγραφή δεδομένων τα οποία αφορούν την κατάσταση της υγείας και της συμπεριφοράς του χρήστη. Μέσω της καταγραφής που γίνεται από αυτές τις συσκευές μπορεί ο γιατρός να εκτιμήσει από τα συμπτώματα που εκδηλώνονται αν οδηγείται ο χρήστης σε κάποιο είδος ασθένειας (Backonja etal, 2012). Άλλες εφαρμογές που αφορούν την διατροφή ή τον αθλητισμό δημιουργούν επίσης Big Data για την δημόσια υγεία αρκεί μονό να σκεφτεί κανείς ότι στο κατάστημα Apple iTunes υπάρχουν παραπάνω από 40.000 εφαρμογές που έχουν άμεση σχέση με το θέμα της υγείας (Aitken&Gauntlett, 2013).

#### 5.1.4 Μεγάλα δεδομένα σε ιατρικά πειράματα

Τα ιατρικά πειράματα είναι καθοριστικά για την εξέλιξη της ιατρικής επιστήμης. Τα Big Data που δημιουργούνται βασίζονται κυρίως στην μοριακή βιολογία και την έρευνα των γονιδιωμάτων στο σύνολο του ανθρώπινου σώματος εξετάζοντας αλληλεπιδράσεις DNA, RNA και πρωτεϊνών (Lodish, 2008).

Κάποιες από τις βασικές τεχνικές που χρησιμοποιεί η μοριακή βιολογία είναι η μοριακή κλωνοποίηση, η αλυσιδωτή αντίδραση πολυμεράσης γνωστή και ως PCR (Polymerase Chain Reaction) κ.α.. Όλες οι τεχνικές και οι εξετάσεις βοηθούν τον θεράπων ιατρό να αξιολογήσει ορθά την κατάσταση του κάθε ασθενή (Vesna, 2000). Είναι γνωστό στην ιατρική βιβλιογραφία ότι για κάθε νέο φάρμακο που δημιουργείται χρειάζονται κλινικές δοκιμές. Είναι στην ουσία μια διαδικασία αξιολόγησης για την αποτελεσματικότητα οποιασδήποτε νέας μορφής θεραπείας μέσω της παρατήρησης και μελέτης των ευρημάτων που εφαρμόζεται κυρίως σε εθελοντές ασθενείς (DerSimonian&Laird, 1986).

Στα πλαίσια της γονιδιακής έρευνας και της ανάλυσης των νουκλεοτιδίων εντός του DNA δημιουργούνται μεγάλες ποσότητες δεδομένων για τον λόγο ότι οι καταγραφές των αλληλουχιών είναι πάρα πολλές. Για αυτό η έρευνα θα πρέπει να γίνεται σε ερευνητικά ιδρύματα με στόχο την ανακάλυψη βιολογικών μηχανισμών βελτιώνοντας το σύστημα υγείας (Obenshain, 2004). Τα γονιδιακά δεδομένα μοριακής ανάλυσης χωρίζονται ως εξής:

1. Γονιδιωματική (genomics)
2. Πρωτεϊμική (proteomics)
3. Επιγονιδιωματική (epigenomics)
4. Λιπιδωμική (lipidomics)
5. Ανοσονομική (immunomics)
6. Γλυκομική (glycomics)
7. Rnomics

(Wu etal, 2016).

### 5.1.5 Μεγάλα δεδομένα στην ιατρική βιβλιογραφία των νοσοκομειακών πληροφοριακών συστημάτων και η εξέλιξή τους

Τα Big Data αποθηκεύονται και επεξεργάζονται σε βάσεις δεδομένων όπως είναι η Cassandra της εταιρείας Apache. Μέσω της Cassandra μπορούμε να επεξεργασθούμε μεγάλους όγκους δεδομένων για τον λόγο ότι σε μια σειρά μπορούν να καταχωρηθούν δυο εκατομμύρια στήλες (Kyoungyoung&Gang-Hoon, 2013). Το πιο δημοφιλές εργαλείο επεξεργασίας που αφορά την υγειονομική περίθαλψη προέρχεται πάλι από την εταιρεία Apache και ονομάζεται Hadoop (Asante-Korang&Jacobs, 2016). Η διαδικτυακή επεξεργασία συναλλαγών OLTP (Online transaction processing) και η διαδικτυακή επεξεργασία ανάλυσης OLAP (Online analytical processing) χρησιμοποιείται κυρίως για ανάλυση και λήψη αποφάσεων (Sheta&Eldeen, 2013). Επίσης η πιο δημοφιλής τεχνολογική προσέγγιση όσο αναφορά τις αναλύσεις των Big Data είναι η ML στην DM. Για παράδειγμα στην διαδικασία ομαλοποίησης γονιδίων η πιο σημαντική τεχνική της BioNLP (Biomedical language processing) είναι η αναγνώριση οντοτήτων (Usami et al, 2011). Μια ακόμα τεχνική είναι ο πολυμορφισμός μήκους ενισχυμένων τμημάτων AFLP (Amplified fragment length polymorfism) η οποία χρησιμοποιείται για την ανάλυση και την εύρεση δακτυλικών αποτυπωμάτων και γονιδιωμάτων DNA (Hassani, 2010). Η DanTE και η DanTE R που είναι η επέκταση της πρώτης, μπορούν να κάνουν επίσης αναλύσεις γονιδιωμάτων μέσω στατιστικών αναλύσεων (Taverner et al, 2012). Εσωτερικά στα νοσοκομεία για την άντληση πληροφοριών μπορούν να χρησιμοποιηθούν νοσοκομειακά πληροφοριακά σύστημα HIS (Hospital Information System) τα οποία κύρια προτεραιότητα τους θα είναι η διαχείριση και η κοινή χρήση δεδομένων από τους ενδιαφερομένους (Abernethy et al, 2011). Για την διαχείριση επιστημονικών και ερευνητικών προβλημάτων οι τεχνολογίες και τα μοντέλα που έχουν ανακαλυφθεί παρατηρήθηκε βάσει αποτελεσμάτων ότι η αποτελεσματικά τους είναι εξαιρετικά πολύτιμη όσο αναφορά δομημένα και μη δομημένα Big Data τα οποία αφορούν την υγεία (Asante-Korang&Jacobs, 2016).

Αυτού του είδους HIS βοηθούν την ορθή διαχείριση νοσοκομείων και δομών υγείας λειτουργώντας ως κοινωνικοτεχνικό υποσύστημα (Tsumoto et al, 2013). Τα HIS είχαν ως αρχή την διοίκηση μονάδων υγείας ενώ αργότερα υπήρξε επέκταση στο κομμάτι της διαχείρισης πληροφοριών (Pai&Huang, 2011). Λόγο των πολυάριθμων

νοσοκομειακών λειτουργιών είναι λογικό να προκύπτουν αρκετά είδη HIS όπως για παράδειγμα το PACS (Picture archiving and communications systems) ενός συστήματος του οποίου ο ρόλος είναι η αποθήκευση και η μεταφορά ψηφιακών εικόνων (Joshi&Yesha, 2012). Άλλοι τύποι HIS είναι τα εργαστηριακά συστήματα πληροφοριών LIS (Laboratory information systems), ακτινολογίας RIS (Radiology information systems), υπερήχων UIS (Ultrasound information systems) και γενικότερα συστήματα EHR, EMR και PHR (He etal, 2010, Joshi&Yesha, 2012). Για δεδομένα τύπου HL7 (Health level seven), ενός συστήματος διαχειρίσεις δεδομένων ασθενών εφαρμόστηκε μοντέλο συστήματος πληροφοριών ανοικτού κώδικα (Celesti etal, 2016).

Η χρήση του cloud στα HIS βοηθά στην επεξεργασία δεδομένων, την δημιουργία αντιγράφων ασφαλείας και την ανταλλαγή πληροφοριών μεταξύ οργανισμών υγείας όπως το σύστημα PACS που προαναφέρθηκε αφού και τα EHR βασίζονται στο cloud (He etal, 2010, Joshi&Yesha, 2012). Η ασφάλεια των δεδομένων και η διαχείριση αυτών στο cloud εξασφαλίζεται μέσω μιας πλατφόρμας υψηλής ποιότητας. Επιπρόσθετος η βελτίωση του HIS είναι συνεχόμενη για παράδειγμα ένα προσαρμοστικό σύστημα διαμεσολάβησης το ARIEN (Adapter interoperability engine) εισήχθη προκειμένου να επιτευχθεί η διαλειτουργικότητα σε ένα σύνολο διαφόρων προτύπων υγειονομικής περίθαλψης όπως επίσης και διάφορα λειτουργικά ανοικτού κώδικα για την υποστήριξη και την ανάπτυξη αυτών των Πληροφοριακών Συστημάτων Νοσοκομείου (HIS) (Khan etal, 2014). Σύμφωνα με τους Bagayoko και Dufour (2010), όλες αυτές οι υποδομές, χρησιμοποιούνται συνήθως στην Ευρώπη και τη Βόρεια Αμερική.

## **5.2 Τα αναλυτικά δεδομένα**

Μια από τις μεγαλύτερες προκλήσεις που καλούνται σήμερα οι οργανισμοί υγειονομικής φροντίδας όσο αναφορά τα Big Data είναι η ορθή εξαγωγή πληροφοριών. Αυτό γίνεται εμφανές στο κατά πόσο πλέον τα Big Data έχουν καταστεί απαραίτητα για κάθε σχεδόν λειτουργία σε όλους του τομείς εργασίας των επαγγελματιών υγείας. Τα τρία μεγάλα V εισήχθησαν από τον αναλυτή Gartner Doug Laney το 2001 όγκος, ταχύτητα και ποικιλία (Volume, velocity, and variety). Με το πέρασ του χρόνου προέκυψαν προβλήματα όπως ελλείψεις δεδομένων αναδεικνύοντας την ανάγκη προσθήκης περισσότερων παραμέτρων όπως είναι η αξία, η οπτικοποίηση,

η βιωσιμότητα, η ευπάθεια, η μεταβλητότητα και η εγκυρότητα περιγράφοντας νέες ιδιότητες οι οποίες πρέπει να ληφθούν υπόψη στην εξελισσόμενη εποχή που ζούμε (Bresnick, 2017).

### **5.2.1 Όγκος (Volume)**

Μια αναφορά από το EMC το 2013 ανέφερε ότι 4,4 zettabytes δεδομένων υπήρχαν παγκοσμίως. Η πρόβλεψη ήταν ότι αυτός ο αριθμός θα αυξάνεται εκθετικά και θα φτάσει σαν όγκος δεδομένων να δεσμεύει 44 zettabytes ή αλλιώς 44 τρισεκατομμύρια gigabytes έως το 2020 (IBM, 2020).

Τα δεδομένα της υγείας αποτελούν μέρος αυτού του όγκου μέσω των πολυάριθμων νοσοκομειακών λειτουργιών και των εισερχόμενων δεδομένων που προκύπτουν από αυτές. Αυτά τα δεδομένα παρέχουν πολλές πληροφορίες που μπορούν να γίνουν ακόμη πιο χρήσιμες όταν διασταυρωθούν με τον σωστό τρόπο ούτως ώστε να παράξουν καινούργιες πληροφορίες. Οι υγειονομικοί οργανισμοί θα πρέπει στο μέλλον να βρουν τρόπο να διασφαλίσουν την υποδομή την ασφάλεια και τον όγκο αποθήκευσης αυτών των δεδομένων λόγω της ανοδικής αύξηση τους χωρίς να παρατρέπεται κρίσιμότερες λειτουργίες όπως είναι η πρόσβαση στα EHR (Bresnick, 2017).

### **5.2.2 Ταχύτητα (Velocity)**

Η ταχύτητα των δεδομένων θα συνεχίσει να αυξάνεται στο μέλλον λόγω των νέων τεχνολογιών που εισάγονται στην καθημερινότητα των ανθρώπων με νέες φορητές συσκευές ελέγχου ζωτικών σημείων κ.α.. Παρόλα αυτά ο χρόνος απόκρισης ενημέρωσης δεδομένων εντός των νοσοκομείων είναι ζωτικής σημασίας να γίνεται σε πραγματικό χρόνο όπως για παράδειγμα θα πρέπει να συμβαίνει στις Μ.Ε.Θ (Μονάδα εντατικής θεραπείας). Σε άλλες περιπτώσεις η συλλογή αυτού του είδους δεδομένων όπως αυτή της επανεισδοχής ασθενούς στο νοσοκομείο, μπορεί να θεωρηθεί ως μη αναγκαία χωρίς αυτό να μειώνει την σημαντικότητα της συλλογής τους (Bresnick, 2017).



### **5.2.3 Ποικιλία (Variety)**

Η ποικιλία των δεδομένων σε παγκόσμιο επίπεδο είναι δεδομένη. Για παράδειγμα από χώρα σε χώρα διαφέρει η γλώσσα που χρησιμοποιείτε από τους χρήστες δημιουργώντας ένα πολυγλωσσικό περιβάλλον καθιστώντας όλα αυτά τα δεδομένα ως μη δομημένα λόγω του μεγάλου όγκου των λέξεων που χρησιμοποιούνται. Για αυτόν τον λόγο έχουν δημιουργηθεί βάσεις δεδομένων NoSQL βοηθώντας στην αποθήκευση και την εξόρυξη αυτών (Kumar, 2018).

Εκτιμάται ότι υπάρχουν 225+ NoSQL βάσεις δεδομένων ανά τον κόσμο. Στον χώρο που βρίσκονται αυτά τα Big Data έχουν την ιδιαιτερότητα της καταχώρησης από πολλές και διαφορετικές πλατφόρμες δημιουργώντας απαιτήσεις πρόσβασης και μορφοποίησης ιδιωτικού έναντι δημόσιου cloud (Sun, 2018).

### **5.2.4 Αξιοπιστία (Veracity)**

Το πρόβλημα που συνήθως προκύπτει για τα δεδομένα στον τομέα της υγειονομικής περίθαλψης είναι το κατά πόσο μπορούμε να τα εμπιστευθούμε. Πολλές φορές προκύπτει να υπάρχουν ελλιπή δεδομένα από το όνομα του ασθενούς μέχρι και το είδος της ασθένειας που νοσεί. Για την χρήση αυτών των δεδομένων έχει παρατηρηθεί ότι χρειάζεται αρκετός χρόνος για τον καθαρισμό τους ούτως ώστε να εξασφαλισθεί η ακεραιότητα, η τυποποίηση και η πληρότητα τους. Η στρατηγική για την διακυβέρνηση αυτών των δεδομένων πρέπει να υιοθετηθεί στο μέλλον για να υπάρχουν σε χώρους εργασίας υψηλής ποιότητας και έτοιμα για χρήση δεδομένα (Bresnick, 2017).

### **5.2.5 Ισχύς (Validity)**

Υπάρχουν περιπτώσεις που τα δεδομένα φαίνονται πλήρη. Στην πραγματικότητα δεν είναι γιατί μπορεί να έχουν προκύψει ενημερώσεις με αλλαγές στα ήδη υπάρχοντα δεδομένα. Για αυτόν τον λόγο είναι αναγκαίο να δημιουργούνται μεταδεδομένα τα οποία θα βοηθήσουν στο μέλλον επιστήμονες που ασχολούνται με την επιστήμη των δεδομένων να ανατρέξουν σε παλαιότερα δεδομένα ούτως ώστε να βρουν αυτό που ψάχνουν (Bresnick, 2017).

### **5.2.6 Βιωσιμότητα (Viability)**

Είναι αδιαμφισβήτητο για να μπορέσουμε να έχουμε επιθυμητά αποτελέσματα όσο αναφορά τις προβλέψεις, να μπορεί και να κατανοεί τα δεδομένα που έχει στην κατοχή του ο εκάστοτε ερευνητής. Για παράδειγμα στο twitter αναφέρθηκαν περιστατικά ατμοσφαιρικής ρύπανσης οδηγώντας σε αύξηση των εισροών σε νοσοκομεία από ασθενείς που νοσούσαν με άσθμα. Από μόνο του αυτό σαν γεγονός μας οδηγεί στο συμπέρασμα ότι πολλές προγνωστικές αναλύσεις επικεντρώνονται στα δεδομένα του παρόντος όπου μέσω αναζήτησης καινοτόμων μεταβλητών μπορούν να περιγράψουν συμπεριφορές ασθενών προβλέποντας κλινικά αποτελέσματα (Bresnick, 2017).

### **5.2.7 Μεταβλητότητα (Volatility)**

Η μεταβολή των δεδομένων της υγειονομικής περίθαλψης είναι αρκετά συχνή εγείροντας πολλά ερωτήματα διαχείρισης από την αρχειοθέτηση τους, έως και την διαγραφή τους. Καθώς παρατηρείται αύξηση του όγκου αυτών των δεδομένων οι αποφάσεις σε σχέση με αυτά ερωτήματα αυξάνεται σε επίπεδο σημαντικότητας. Ένα από τα προβλήματα που θα πρέπει να ανησυχεί τους παρόχους υγείας είναι το κόστος αποθήκευσης των ιατρικών δεδομένων διότι σε πολλές περιπτώσεις αυτά παραμένουν σταθερά και επαναχρησιμοποιήσιμα για μεγάλα χρονικά διαστήματα όπως για παράδειγμα οι γονιδιωματικές εξετάσεις ενός ασθενούς. Σε κάποιες χώρες έχει παρατηρηθεί απαίτηση από τους παρόχους υγείας να αποθηκεύονται τα ιατρικά δεδομένα με έξοδα του ασθενούς για έξι έτη (Bresnick, 2017).

### **5.2.8 Ευπάθεια (Vulnerability)**

Το ιατρικό απόρρητο υποδεικνύει την τεράστια σημασία που έχει η ασφάλεια του ασθενούς. Όσο αυξάνεται ο όγκος των δεδομένων, η αποθήκευση γίνεται πλέον μέσω cloud. Παράλληλα η ασφάλεια αυτών των δεδομένων είναι πλέον υψίστης σημασίας λόγω της μετακίνησης τους ανάμεσα στους οργανισμούς υγείας ως αποτέλεσμα της διαλειτουργικότητας. Οι υγειονομικοί οργανισμοί πρέπει πλέον να δείξουν μεγαλύτερο ενδιαφέρον για την ευπάθεια αυτών των δεδομένων με την συνεχή εκπαίδευση του προσωπικού ασφαλείας δεδομένων σε νέες τεχνικές διασφάλισης. Επιπλέον για την διατήρηση της ιδιωτικότητας των αρχείων προτείνεται η υπογραφή συμφωνιών επιχειρηματικών συνεργασιών οι οποίες θα λειτουργούν με τους κανόνες

συμμόρφωσης και της αυστηρής προστασίας της ιδιωτικής ζωής και ασφάλειας (Bresnick, 2017).

### **5.2.9 Οπτικοποίηση (Visualization)**

Σε τμήματα όπως τα τμήματα επειγόντων περιστατικών (ΤΕΠ) ή των μονάδων εντατικής θεραπείας (Μ.Ε.Θ.) η πολυπλοκότητα επεξεργασίας πληροφοριών φέρνει σε σύγκρουση καθημερινά τους επαγγελματίες υγείας μέσω την χρήση των ηλεκτρονικών αρχείων υγείας. Η δυσκολία έγκειται στο γεγονός του κατά πόσο είναι ευκολονόητες οι οπτικοποιήσεις για την διεξαγωγή προβλέψεων ή για την άντληση κάποιας βασικής γνώσης. Οι πολλές και δυσνόητες αναφορές αποθαρρύνουν τους υγειονομικούς υπάλληλους να κάνουν πλήρη χρήση των εργαλείων πληροφορικής που τους δίνονται. Οι προγραμματιστές που είναι υπεύθυνοι για τον έλεγχο των προγραμμάτων που χρησιμοποιούνται θα πρέπει να εξετάσουν την προσθήκη κατανοητών απεικονίσεων με αναγνωρίσιμα χρώματα και γραφήματα που θα επισημαίνουν πληροφορίες και δεν θα κουράζουν τον αναγνώστη. Οι ευκολονόητες πληροφορίες θα έχουν σαν αποτέλεσμα την αποφυγή υπερφόρτωσης πληροφοριών κάτι το οποίο θα οδηγήσει στον μετριασμό της εξάντλησης που προκαλείται μεταξύ των απασχολούμενων κλινικών. Όσο αναφορά τις οικονομικές αναφορές όπως επίσης τις λειτουργικές και κλινικές μετρήσεις μπορούν χρησιμοποιηθούν διαδραστικοί πίνακες. Ενώ όσο αναφορά την δημόσια υγεία μέσω εργαλείων χαρτογράφησης διαδικτύου μπορούν να επισημανθούν προβληματισμοί σε τοπική και εθνική κλίμακα λόγω της ποικιλίας εφαρμογών που έχουν εμφανιστεί για υπολογιστές tablet και smartphone οι οποίες αλληλεπιδρούν με δεδομένα ποιο ουσιαστικά από ποτέ άλλοτε (Bresnick, 2017).

## **6. Ιδιαιτερότητα των μεγάλων δεδομένων στην υγεία**

Με την ανάλυση των Big Data οι ενδιαφερόμενοι επωφελούνται καθημερινά βελτιώνοντας την ποιότητα των υγειονομικών υπηρεσιών λόγω των σημαντικών πληροφοριών που δημιουργούνται. Αυτές οι πληροφορίες βοηθούν τους επαγγελματίες υγείας στη λήψη τεκμηριωμένων αποφάσεων και την παρακολούθηση ασθενειών. Οι υπεύθυνοι που παρακολουθούν μαζικά θέματα υγείας θα πρέπει να προσαρμόσουν τις γενικότερες προβλέψεις προβλημάτων υγείας σε ανάλυση ενός συγκεκριμένου είδους θεραπείας κατάλληλης για ένα συγκεκριμένο άτομο. Μέσω αυτών των αναλύσεων μπορούν να διατυπωθούν σχέδια θεραπείας ενημερώνοντας τον ασθενή για την τρέχουσα και τη μελλοντική κατάσταση της υγείας του, γεγονός που θα του επιτρέπει να λάβει ορθές αποφάσεις αυξάνοντας την ποιότητα των υπηρεσιών υγείας (Raghupathi&Raghupathi, 2014).

### **6.1 Μείωση κόστους υγειονομικής περίθαλψης**

Τα υγειονομικά συστήματα σήμερα για την θεραπεία ασθενειών βασίζονται σε μοντέλα και πρωτόκολλα όπου βάση συμπτωμάτων και εργαστηριακών αναφορών προτείνεται η ανάλογη θεραπεία. Μέσω της κλινικής εμπειρογνωμοσύνης σε περιπτώσεις που η θεραπεία δεν έχει τα επιθυμητά αποτελέσματα μπορούν να προταθούν νέες εναλλακτικές μέθοδοι θεραπείας. Οι διαδικασίες αυτές συνήθως είναι χρονοβόρες ενώ σε περιπτώσεις σοβαρών ασθενειών αποδεικνύονται δαπανηρές. Η χρήση των Big Data με τον ήδη τεράστιο όγκο δεδομένων μπορούν να οδηγήσουν σε θεραπείες οι οποίες επικεντρώνονται στον ασθενή επιτρέποντας την ορθή διάγνωση και πρόβλεψη της θεραπείας του με μεγαλύτερη ακρίβεια. Επίσης μέσω αναλύσεων στα γονίδια του ασθενούς εξασφαλίζεται η δυνατότητα εξατομικευμένης θεραπείας. Έτσι οδηγούμαστε στο συμπέρασμα ότι τα Big Data μέσω των αναλύσεων μπορούν να μειώσουν τις περιττές ιατρικές θεραπευτικές δαπάνες (Srinivasan&Arunasalam, 2013).

### **6.2 Ετερογένεια**

Τα Big Data όπως αναφέρθηκε, μπορούν να ταξινομηθούν σε δομημένα και μη δομημένα. Για παράδειγμα ορισμένα EHR έχουν δεδομένα σε δομημένες μορφές όπως επίσης η Διεθνής Ταξινόμηση των Νοσημάτων (ICD-10) κατέχει δομημένα δεδομένα (Asante-Korang&Jacobs, 2016). Ωστόσο το μεγαλύτερο ποσοστό των Big Data που αφορούν την υγεία δεν περιέχουν δομημένα δεδομένα (Swan, 2013) ενώ παράλληλα

υπάρχει έλλειψη εργαλείων όσο αναφορά αναλύσεις πληροφοριών από ετερογενείς πηγές δεδομένων. Ένα εργαλείο ανάλυσης ετερογενών δεδομένων αναπτύχθηκε από Γερμανούς ειδικούς και είχε ως στόχο την βελτίωση γνώσεων για σπάνιες ασθένειες ενσωματώνοντας ιατρικά αρχεία και δεδομένα απεικόνισης (Deserno et al, 2014). Οι Windridge και Bober (2014) κάνοντας χρήση του πυρήνα (kernel) για ανάλυση ετερογενών δεδομένων επέτυχαν παλινδρόμηση και ταξινόμηση ετερογενών ιατρικών πληροφοριών. Αντιμετώπισαν όμως σωρεία προβλημάτων εξαιτίας ελλείψεων πληροφοριών που αφορούσαν τους ασθενείς.

### **6.3 Ατελή δεδομένα**

Οι συσκευές παρακολούθησης ζωτικών οργάνων φορητές ή μη αποτελούν ένα εργαλείο εισερχόμενων δεδομένων όπως είναι για παράδειγμα το ηλεκτροκαρδιογράφημα (Kruse et al, 2018). Εξαιτίας του κόστους αποθήκευσης των Big Data πολλές φορές δεν αποθηκεύονται κάποια από τα δεδομένα που συλλέγονται οδηγώντας σε ελλείψεις καταγραφής τους. Επιπροσθέτως τα EHR οδηγούν το υγειονομικό προσωπικό να καταγράφει αυτές τις πληροφορίες χειρωνακτικά κάτι το οποίο επίσης μπορεί να οδηγήσει σε ελλείψεις δεδομένων για τον λόγο ότι κάποιος μπορεί να ξεχάσει να προσθέσει πληροφορίες στο ιστορικό (Hong et al, 2015). Σε μια έρευνα που διεξήχθη στο Menelik II Referral Hospital, παρατηρήθηκε ότι το ιστορικό και η πληρότητα πληροφοριών σε σχέση με τους ασθενείς και τις ιατρικές πράξεις ήταν στο 73% ένα ποσοστό πολύ χαμηλό σε σχέση με τα πρότυπα νοσηλείας. Επομένως η βελτίωση της πληρότητας των ιατρικών αρχείων είναι σημαντικότερη για την βελτίωση της ποιότητας των υπηρεσιών υγείας (Tola et al, 2017).

### **6.4 Επικαιρότητα**

Οι ιατρικές εξετάσεις όπως το ηλεκτροκαρδιογράφημα (ΗΚΓ), οι μονές φωτονικές εκπομπές υπολογιστικής τομογραφίας SPECT (Single photon emission computed tomography), η μαγνητική τομογραφία και το ηλεκτροεγκεφαλογράφημα EEG (Electroencephalogram), επηρεάζονται από την συνάρτηση του χρόνου και επομένως πρέπει να διατηρούνται όσο το δυνατόν ενημερωμένες. Η ενημέρωση των ιατρικών πληροφοριών αποτελεί μείζονα πρόκληση για τα Big Data της υγειονομικής φροντίδας ενώ τα σύστημα υγείας θα πρέπει να εργασθούν στο κομμάτι τις άμεσης επικαιροποίησης των δεδομένων υγείας για τον λόγο ότι ο χρόνος αποθήκευσης και

ενημέρωσης αυτών των δεδομένων είναι διαφορετικός μεταξύ των νοσοκομείων. Για την υποστήριξη λήψης αποφάσεων όσο αναφορά ασθενείς που πάσχουν από γενετικές ασθένειες είναι χρήσιμη η ενημέρωση πληροφοριών που σχετίζονται με το οικογενειακό ιστορικό του ασθενή (Hongetal, 2019).

## **6.5 Απόρρητο δεδομένων**

Η ιδιαιτερότητα και η ευαισθησία των δεδομένων που αφορούν την υγεία εγείρουν τεράστιες ανησυχίες οι οποίες αφορούν το απόρρητο και την ασφάλεια τους (Naito, 2014). Η προστασία της ιδιωτικότητας των δεδομένων είναι άμεσα συνυφασμένη με την εμπειρία των ασθενών που κάνουν χρήση των δομών υγείας οδηγώντας στην βελτίωση της εμπειρίας παροχής υπηρεσιών υγείας (Yuen-Reed&Mojsilović, 2016). Επίσης πρέπει να αναφερθεί ότι λόγω του συγκεντρωμένου όγκου πληροφοριών τα δεδομένα είναι ευάλωτα σε επιθέσεις χάκερ οι οποίοι μπορούν να αντλήσουν κρίσιμες πληροφορίες από το αρχείο κάποιου ασθενή (Mohr etal, 2013) για αυτόν το λόγο οι Herland etal (2014) με τη βοήθεια ενός υγειονομικού επαγγελματία χρησιμοποίησαν συνθεμένα EMR, EHR και PHRs για την διεξαγωγή έρευνας όσο αναφορά τα συγκεκριμένα ζητήματα ασφάλειας.

Εφαρμογές όπως το Google Health υπόσχεται «τον πλήρη έλεγχο των δεδομένων» τα οποία δεν θα πωληθούν ή κοινοποιηθούν χωρίς τη ρητή άδεια του καταναλωτή (Steinbrook, 2008). Σε άλλες χώρες, υπάρχουν πρότυπα πολιτικών και κανονισμών για την προστασία των δεδομένων υγείας. Σε ένα μοτίβο, με βάση τους βασικούς νόμους περί απορρήτου, οι κυβερνήσεις θεσπίζουν πρόσθετους νόμους, πολιτικές και κανονισμούς για την προστασία των προσωπικών πληροφοριών που αφορούν την υγεία, όπως το HIPAA στις ΗΠΑ, ο νόμος περί προστασίας προσωπικών δεδομένων και πληροφοριών του 2002 στην Αυστραλία όπως επίσης ο Ασφαλιστικός νόμος στη Γαλλία (Hongetal, 2019).

## **6.6 Ιδιοκτησία**

Αν και οι ασθενείς κατέχουν νόμιμα τα δεδομένα της υγείας τους, ενδέχεται να μην έχουν πρόσβαση σε αυτά. Για αυτόν τον λόγο προτείνεται ένας συνεταιρισμός ο οποίος θα λειτουργεί ως μια μορφή εταιρείας και θα ανήκει εξ ολόκληρου στους πολίτες. Ο ασθενής θα μπορεί να δημιουργεί έναν λογαριασμό όπου θα αποθηκεύει και θα διαχειρίζεται όλα τα αρχεία τα οποία τον αφορούν όπως επίσης θα μπορεί να

διαμοιράζει πληροφορίες και δεδομένα για ερευνητικούς σκοπούς (Pentland, Reid&Heibeck, 2013).

## **7. Μελέτη περίπτωσης : Προγνωστικές αναλύσεις δευτερογενών δεδομένων μη επειγόντων περιστατικών μεταφοράς ασθενών σε μονάδες τεχνητού νεφρού από το Ε.Κ.Α.Β. Θεσσαλονίκης**

### **7.1 Το Εθνικό Κέντρο Άμεσης Βοήθειας (Ε.Κ.Α.Β.)**

Με τον Νόμο 1579/1985 ιδρύεται στην Ελλάδα το Εθνικό Κέντρο Άμεσης Βοήθειας (Ε.Κ.Α.Β.). Θεωρείται από τα παλαιότερα συστήματα προνοσοκομειακής ιατρικής στην Ευρώπη (WHO, 2008) και έχει ως αποστολή την παροχή προνοσοκομειακής φροντίδας μέσω ενός συνόλου υπηρεσιών όπως αυτές έχουν οριστεί από τον οργανισμό (Ε.Κ.Α.Β., 2021).

Με τον όρο προνοσοκομειακή ιατρική ή αλλιώς επείγουσα ιατρική θεωρείται η ιατρική ειδικότητα κατά την οποία απαιτούνται γνώσεις για την διαχείριση επειγόντων περιστατικών για παράδειγμα ενός τροχαίου ατυχήματος. Το προνοσοκομειακό σύστημα όπου υπάγεται η ειδικότητα της προνοσοκομειακής ιατρικής φροντίδας ορίζεται ως το σύστημα το οποίο περιλαμβάνει μεταφορά κινητής ιατρικής μονάδας στο σημείο περιστατικού το οποίο χρήζει άμεσης προνοσοκομειακής φροντίδας (Spraitte, 1995).

Οι υπηρεσίες αυτού του είδους αποτελούνται από μια ομάδα ειδικού ενεργού ιατρικού και παραϊατρικού προσωπικού όπως επίσης ενός κέντρου ελέγχου που περιλαμβάνει τις διοικητικές νοσηλευτικές υπηρεσίες, το τηλεφωνικό-συντονιστικό κέντρο όπου εκτός των διοικητικών υπηρεσιών λειτουργούν 24 ώρες κάθε μέρα καθ' όλη την διάρκεια του έτους σε πλήρη ετοιμότητα για την διαχείριση επειγόντων περιστατικών (Papasprou et al, 2004).

Με τον εκσυγχρονιστικό νόμο 2071 (ΦΕΚ 123, 1992) ο σκοπός, η δομή και οι λειτουργίες του Ε.Κ.Α.Β. αναβαθμίζονται. Στο Κεφάλαιο Γ άρθρο 39 ο σκοπός του οργανισμού αναφέρεται ως εξής:

- Εφαρμογή σωστικών μέτρων για την υγεία και τη ζωή σε άτομα που βρίσκονται σε κίνδυνο και μεταφορά τους, με παροχή ειδικής φροντίδας, σε κατάλληλο υγειονομικό σχηματισμό.



- Μεταφορά περιστατικών που δεν χαρακτηρίζονται ως επείγοντα, αλλά που δεν είναι δυνατό να μεταφερθούν με άλλο τρόπο.
- Εκπαίδευση σε θέματα παροχής πρώτων βοηθειών και επείγουσας προνοσοκομειακής ιατρικής, στη πληθυσμό, στα πληρώματα των ασθενοφόρων και σε λοιπές άλλες κοινωνικές ομάδες.
- Άμεση συμμετοχή και συντονισμός για την παροχή ιατρικής βοήθειας σε περιπτώσεις μαζικών ατυχημάτων ή καταστροφών.

Το Ε.Κ.Α.Β. χωρίζεται σε 12 κύριους σταθμούς: Αθήνας, Θεσσαλονίκης, Πάτρας, Ηράκλειου, Λάρισας, Καβάλας, Ιωαννίνων, Αλεξανδρούπολης, Λαμίας, Μυτιλήνης, Τρίπολης και Κοζάνης (Εικόνα 7) Όσο αναφορά τον συντονισμό υποπεριοχών έχουν δημιουργηθεί μικροί υποσταθμοί οι οποίοι είναι υπό την εποπτεία των 12 μεγαλύτερων σταθμών διασφαλίζοντας το 96,2% των αναγκών προνοσοκομειακής φροντίδας σε όλες τις περιοχές της Ελλάδος. Συνήθως οι τοπικοί σταθμοί βρίσκονται μέσα σε νοσοκομεία ή σταθμούς δομών υγείας όπως για παράδειγμα τα Κέντρα Υγείας (Papasprou etal, 2004).



Εικόνα 7. Οι 12 μεγάλοι σταθμοί του Εθνικού Κέντρου Άμεσης Βοήθειας ανά την Ελλάδα. Ανατύπωση από Ε.Κ.Α.Β. (2021).

Τα μέσα διακομιδής του Ε.Κ.Α.Β. είναι το πολυτιμότερο εργαλείο διαχείρισης κρίσεων. Στον πίνακα 1 παρουσιάζονται τα μέσα που χρησιμοποιούνται συνολικά σε όλη την Ελλάδα από το Ε.Κ.Α.Β.:

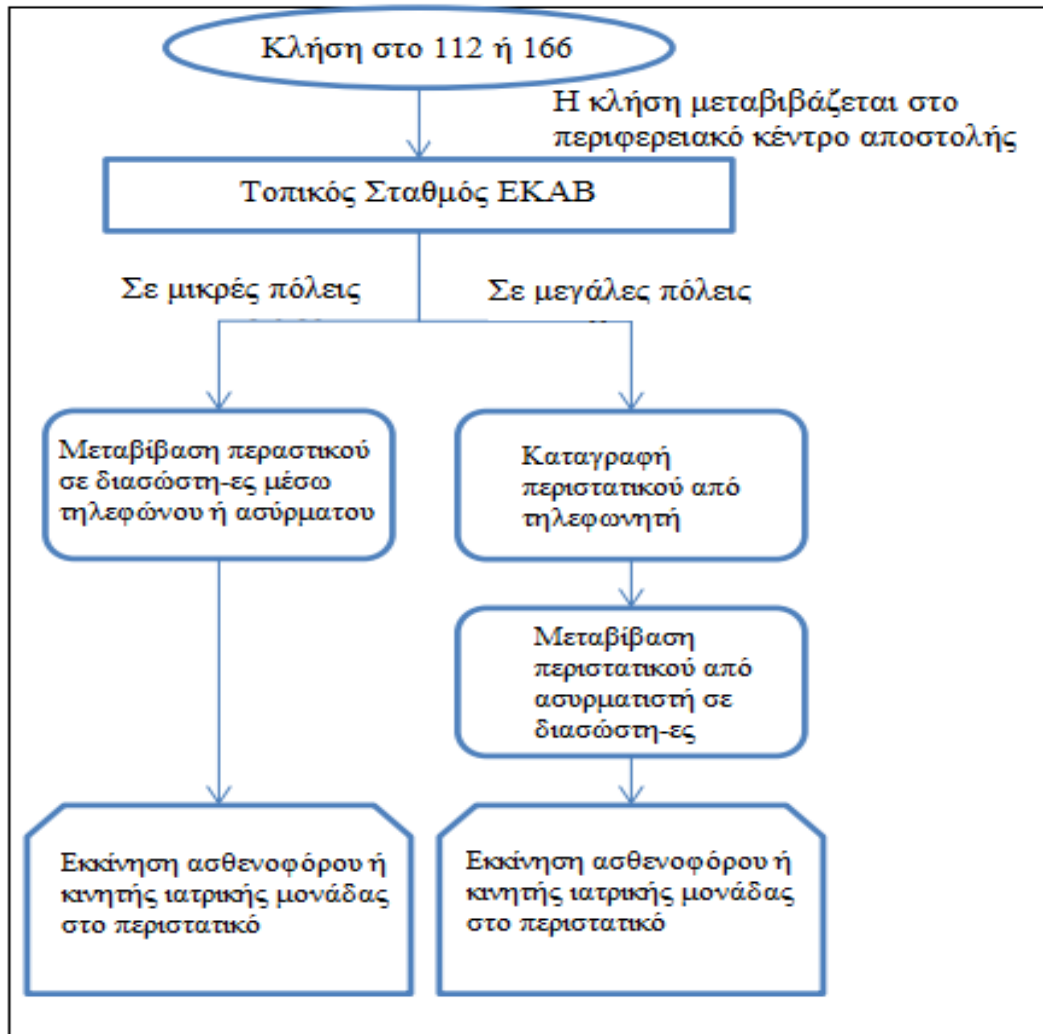
<b>Ασθενοφόρα</b>	735
<b>Ειδικές Κινητές Μονάδες Επείγουσας Προνοσοκομειακής Ιατρικής</b>	102
<b>Μοτοσυκλέτες</b>	25
<b>Μικρά οχήματα ταχείας πρόσβασης (smart, saxo)</b>	4
<b>Ελικόπτερα</b>	3
<b>Αυτοκινούμενα Συντονιστικά Κέντρα</b>	2
<b>Οχήματα Αντιμετώπισης Καταστροφών</b>	2
<b>Ασθενοφόρα που έχουν παραχωρηθεί από το ΕΚΑΒ και χρησιμοποιούνται από Νοσοκομεία και Κέντρα Υγείας</b>	350

Πίνακας 1. Μέσα Διακομιδής Εθνικού Κέντρου Άμεσης Βοήθειας. Ανατύπωση από Ε.Κ.Α.Β. (2021).

Η χρήση των υπηρεσιών προνοσοκομειακής έκτακτης φροντίδας του Ε.Κ.Α.Β. πραγματοποιούνται μέσω του τριψήφιου αριθμού "166" όπως επίσης και του αριθμού "112" (Paraspyrou et al, 2004) οι οποίοι θεωρούνται βάσει κανονισμού της Ευρωπαϊκής Ένωσης αριθμοί κλήσης που αφορούν επείγοντα περιστατικά (WHO, 2008).

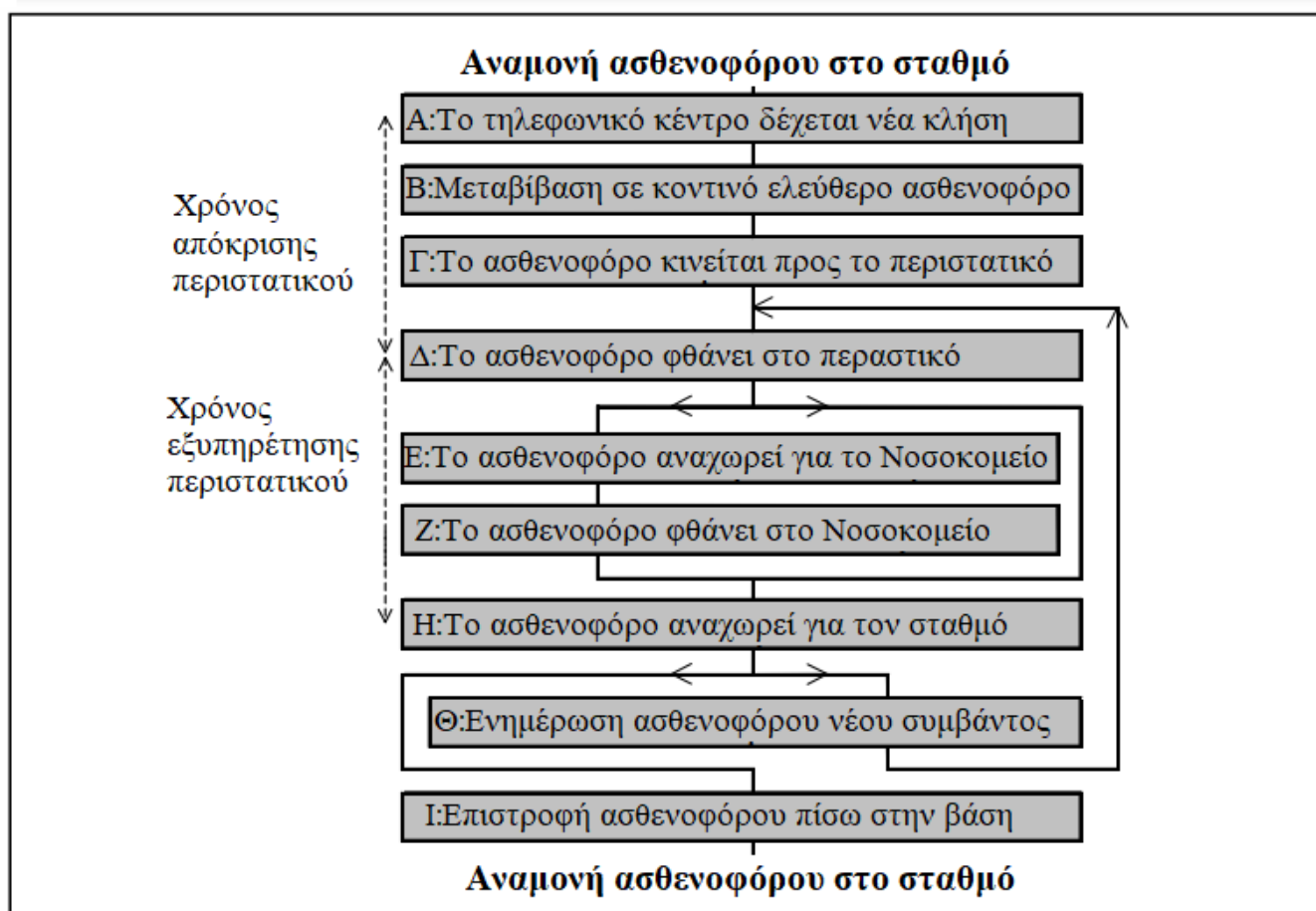
Η διαδικασία αποστολής κινητής ιατρικής μονάδας σε ένα περιστατικό ξεκινά από το τηλεφωνικό κέντρο που βρίσκεται σε έναν από τους 12 βασικούς σταθμούς. Μέσω ενός εξειδικευμένου προσωπικού ειδικό σε θέματα προνοσοκομειακής φροντίδας μπορεί να γίνει η γρήγορη και σαφής καταγραφή των περιστατικών με όλες τις πληροφορίες που είναι αναγκαίες για την προετοιμασία αποστολής κινητής ιατρικής μονάδας. Οι κινητές ιατρικές μονάδες ενημερώνονται μέσω ασύρματου που τον χειρίζεται ειδικός ασυρματιστής ο οποίος έχει λάβει ήδη τα δεδομένα που αφορούν το περιστατικό μέσω των καταγραφών που έχουν γίνει από το τηλεφωνικό κέντρο. Στην συνέχεια ο ασυρματιστής μέσω ασυρμάτου μεταβιβάζει την πληροφορία σε ασθενοφόρο που βρίσκεται σε πόστο ανάλογου τοπογραφικού διαμερίσματος για την άμεση και γρήγορη αποστολή του. Όσο αναφορά τους υποσταθμούς που βρίσκονται

σε μικρότερες περιοχές η επικοινωνία γίνεται μέσω ασυρμάτου και τηλεφώνου αναλόγως την περίπτωση (Εικόνα 8) (Papaspyrou et al. 2004).



Εικόνα 8. Παράδειγμα διαγράμματος ροής που εφαρμόζεται στην Ελλάδα για την αποστολή κινητής ιατρικής μονάδας. Ανατύπωση από Page et al (2013).

Για την πλήρη κατανόηση των υπηρεσιών προνοσοκομειακής ιατρικής φροντίδας στην Εικόνα 9 παρουσιάζεται ένα μοντέλο προσομοίωσης με τα βήματα που ακολουθεί ένα ασθενοφόρο ή μια κινητή ιατρική μονάδα από την έναρξη έως και την ολοκλήρωση της αποστολής τους.



Εικόνα 9. Μοντέλο προσομοίωσης αποστολής ασθενοφόρου ή κινητής ιατρικής μονάδας. Ανατύπωση από Henderson etal (2004).

Όταν φτάνει μια κλήση στο τηλεφωνικό κέντρο, ο ασυρματιστής στο κέντρο ελέγχου καλεί το πρώτο διαθέσιμο ασθενοφόρο που βρίσκεται στην βάση του ή οποιοδήποτε διαθέσιμο ασθενοφόρο που επιστρέφει από προηγούμενη αποστολή. Το προσωπικό του ασθενοφόρου καθώς φθάνει στο σημείο του συμβάντος παρέχει τις πρώτες βοήθειες και στην συνέχεια συνήθως μεταφέρει τον ασθενή σε νοσοκομείο. Στην συνέχεια ο ασθενής διακομίζεται στο τμήμα επειγόντων περιστατικών του νοσοκομείου και αφού γίνει η παράδοση του το ασθενοφόρο επιστρέφει στην βάση του. Σε περιπτώσεις που δεν απαιτείται μεταφορά του ασθενή στο νοσοκομείο το ασθενοφόρο επιστρέφει απευθείας στη βάση του. Σε κάθε περίπτωση, το ασθενοφόρο θεωρείται διαθέσιμο από την στιγμή που έχει αρχίσει ήδη την επιστροφή στη βάση του (Henderson etal, 2004). Ο χρόνος απόκρισης και εξυπηρέτησης του περιστατικού από το Ε.Κ.Α.Β. θεωρείται ο σημαντικότερος παράγοντας του μοντέλου και υπολογίζεται στα 6 με 8 λεπτά (Aslanidis etal 2016, Γιαννόπουλος etal 2018).

## **7.2 Ηλεκτρονική διαχείριση δεδομένων Ε.Κ.Α.Β.**

Για την κατανόηση της ιδιαιτερότητας των δεδομένων του Ε.Κ.Α.Β. σε αυτό το κεφάλαιο γίνεται η παρουσίαση του προγράμματος καταχώρησης δεδομένων που χρησιμοποιεί η υπηρεσία του Ε.Κ.Α.Β. η οποία φέρει την ονομασία «ηλεκτρονική διαχείριση ασθενών 166» και έχει ως βασική λειτουργία την καταγραφή περιστατικών.

Το πρόγραμμα «ηλεκτρονική διαχείριση ασθενών 166» παρέχει τέσσερα επίπεδα χρηστών προσαρμόζοντας λειτουργίες ανάλογα την εργασία του χρήστη ως τηλεφωνητής, ασυρματιστής, διαχειριστή ασθενοφόρων και διαχειριστή συστήματος «ηλεκτρονικής διαχείρισης ασθενών 166» σε συνεργασία με την εταιρεία Link Technologies (2008).

Με την έναρξη της τηλεφωνικής επικοινωνίας ο τηλεφωνητής έχει ήδη μπροστά του ανοιχτή μια κάρτα η οποία είναι έτοιμη για χρήση όπως φαίνεται στην εικόνα 10. Το σημαντικότερο χαρακτηριστικό της λειτουργίας αυτής της κάρτας είναι η διαμόρφωση του περιβάλλοντος καταχώρησης που έχει σκοπό την ταχύτερη εισαγωγή δεδομένων όπως είναι το ονοματεπώνυμο του πάσχοντος ή του καλούντος (σε αρκετές περιπτώσεις ο παθών είναι αδύνατο να επικοινωνήσει ο ίδιος με το Ε.Κ.Α.Β.), περιοχή, διεύθυνση, είδος πάθησης, ηλικία, φύλο, σημαντικότητα περιστατικού κ.α..

**Εισαγωγή Καρτέλας**

**ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ  
ΥΠΟΥΡΓΕΙΟ ΥΓΕΙΑΣ ΚΑΙ ΠΡΟΝΟΙΑΣ  
ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΑΜΕΣΗΣ ΒΟΗΘΕΙΑΣ  
ΠΑΡΑΡΤΗΜΑ ΕΚΑΒ-2 ΘΕΣΣΑΛΟΝΙΚΗΣ**

**Ε.Σ.Υ.**

Α/Α ΚΑΡΤΑΣ: 294  
 ΗΜΕΡΟΜΗΝΙΑ: 07/05/2008  
 ΘΕΣΗ ΤΗΛΕΦΩΝΗΤΗ: 1  
 ΩΡΑ Τ/Φ ΚΛΗΣΗΣ: 13:00:47

ΠΡΟΓΡΑΜΜΑΤΙΣΜΕΝΟ:  Όχι  Ναι

ΤΗΛΕΦΩΝΑ ΚΑΛΟΥΝΤΟΣ: \*\*\*\*\*  
 ΟΝΟΜΑ ΚΑΛΟΥΝΤΟΣ: \*\*\*\*\*

ΕΠΩΝΥΜΟ ΠΑΣΧΟΝΤΟΣ: \*\*\*\*\*  
 ΟΝΟΜΑ ΠΑΣΧΟΝΤΟΣ: \*\*\*\*\*

ΔΙΕΥΘΥΝΣΗ: \*\*\*\*\* \*\*  
 ΔΗΜΟΣ/ΠΕΡΙΟΧΗ: \*\*\*\*\*

ΧΑΡΑΚΤΗΡΙΣΤΙΚΟ ΣΗΜΕΙΟ ΑΝΑΜΟΝΗΣ ΣΥΝΟΔΟΥ: \*\*\*\*\* \*\*

ΠΕΡΙΓΡΑΦΗ ΠΑΘΗΣΗΣ: ΔΗΛΗΤΗΡΙΑΣΕΙΣ - ΑΠΟΡΡΥΠΑΝΤΙΚΑ  
 ΔΙΑΠΙΣΤΩΜΕΝΕΣ ΝΟΣΟΙ:

ΣΗΜΕΙΩΣΕΙΣ:

ΤΗΛΕΦΩΝΗΣΕ Η ΚΑΘΗΓΗΤΡΙΑ ΤΗΣ

ΣΗΜΑΝΤΙΚΟΤΗΤΑ:  Μη Επείγον  Επείγον  Υπερεπείγον  
 ΕΚΓΥΡΟ:  Ναι  Όχι

Επισύναψη Καταχώρηση

Εικόνα 10. Τυχαίο παράδειγμα εισαγωγής δεδομένων σε κάρτα του προγράμματος «ηλεκτρονικής διαχείρισης ασθενών 166». Ανατύπωση από «ηλεκτρονική διαχείριση ασθενών 166» σε συνεργασία με την εταιρεία Link Technologies (2008).

Εφόσον η κάρτα καταχωρηθεί στο σύστημα θεωρείται ενεργό περιστατικό. Από την στιγμή που το περιστατικό έχει χαρακτηριστεί (για παράδειγμα στην εικόνα 10 φέρει χαρακτηρισμό υπερεπείγον) η κάρτα στο σύστημα εμφανίζεται με ανάλογο χρώμα. Στην συνέχεια ειδικοί διασώστες ασυρματιστές λαμβάνουν τις πληροφορίες που χρειάζονται για την διαχείριση του περιστατικού μέσω του προγράμματος «ηλεκτρονική διαχείριση ασθενών 166». Το περιβάλλον του προγράμματος από την μεριά του ασυρματιστή είναι διαφορετικό λόγο των σύνθετων πληροφοριών που καλείται να διαχειριστεί. Για αυτόν τον λόγο έχουν δημιουργηθεί αρκετά υπό-παράθυρα τα οποία συνοψίζουν περιστατικά επείγοντα ή μη, διαθεσιμότητα ασθενοφόρων και άλλες πολλές λειτουργίες (Εικόνα 11). Στην περίπτωση που ο ασυρματιστής ανοίγει μια κάρτα την βλέπει πλέον μόνο αυτός. Ο σκοπός αυτής της

λειτουργίας είναι η απομόνωση του περιστατικού από το περιβάλλον των καταχωρητών, τηλεφωνητών και συντονιστών ασθενοφόρων ούτως ώστε να μπορεί να προγραμματίσει την αποστολή ασθενοφόρου ή κινητής ιατρικής μονάδας σε κάποιο περιστατικό «ηλεκτρονική διαχείριση ασθενών 166» σε συνεργασία με την εταιρεία Link Technologies (2008).

The screenshot displays the Link Technologies software interface, which is used for managing ambulance operations. The interface is divided into several sections:

- Top Bar:** Shows the application name "Ασθενοφόρο" and navigation options like "Από Τηλεφωνητή" and "Καρτέλες".
- Left Panel:** Contains a "Προγραμματισμένο" (Scheduled) table and an "Εκκρεμεί" (Outstanding) table. The "Εκκρεμεί" table lists ambulance IDs (A/A), names (Όνομα), status (Προτεραιότητα), and phone numbers (Ανοχή).
 

A/A	Όνομα	Προτεραιότητα	Ανοχή
294	ΘΕΡΜΗΣ	Υπερσίγου	+23-27-22
291	ΚΑΛΑΜΑΡΙΑΣ	Επίγου	+23-42-35
- Main Form:** A detailed form for patient and ambulance information. It includes fields for:
  - ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ** (Hellenic Republic)
  - ΥΠΟΥΡΓΕΙΟ ΥΓΕΙΑΣ ΚΑΙ ΠΡΟΝΟΙΑΣ** (Ministry of Health and Social Solidarity)
  - ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΑΜΕΣΗΣ ΒΟΗΘΕΙΑΣ ΠΑΡΑΡΤΗΜΑ ΕΚΑΒ-2 ΘΕΣΣΑΛΟΝΙΚΗΣ** (National Center for Emergency Medical Services - Thessaloniki Branch)
  - Α/Α ΚΑΡΤΑΣ** (Ambulance ID): 292
  - ΗΜΕΡΟΜΗΝΙΑ** (Date): 07/05/2008
  - ΘΕΣΗ ΤΗΛΕΦΩΝΗΤΗ** (Phone Location): 1
  - ΩΡΑ Τ/Φ ΚΛΗΣΗΣ** (Call Time): 12:48:51
  - ΦΥΛΟ** (Gender):  Άντρας  Γυναίκα
  - ΗΛΙΚΙΑ** (Age): 65
  - ΑΣΦ. ΦΟΡΕΑΣ** (Insurance Provider): ΟΚΑ
  - ΑΡΘΡΟΣΜΟΣ** (Arrested): 12
  - ΠΡΟΓΡΑΜΜΑΤΙΣΜΕΝΟ** (Scheduled):  Όχι  Ναι
  - ΤΗΛΕΦΩΝΙΑ ΚΑΛΟΥΝΤΟΣ** (Caller Phone Number): \*\*\*\*\*
  - ΟΝΟΜΑ ΚΑΛΟΥΝΤΟΣ** (Caller Name): \*\*\*\*\*
  - ΕΠΩΝΥΜΟ ΠΑΣΧΟΝΤΟΣ** (Patient Surname): \*\*\*\*\*
  - ΟΝΟΜΑ ΠΑΣΧΟΝΤΟΣ** (Patient Name): \*\*\*\*\*
  - ΔΙΕΥΘΥΝΣΗ** (District): \*\*\*\*\* \*\*
  - ΔΗΜΟΣ/ΠΕΡΙΟΧΗ** (Municipality/Area): ΘΕΣΣΑΛΟΝΙΚΗΣ
  - ΧΑΡΑΚΤΗΡΙΣΤΙΚΟ ΣΗΜΕΙΟ ΑΝΑΜΟΝΗΣ ΣΥΝΟΔΟΥ** (Waiting Point Characteristics):
  - ΠΕΡΙΓΡΑΦΗ ΠΑΘΗΣΗΣ** (Disease Description): ΠΑΘΟΛΟΓΙΚΕΣ ΠΑΘΗΣΕΙΣ - ΑΓΓΕΙΑ
  - ΔΙΑΓΝΩΣΤΟΜΕΝΕΣ ΝΟΣΟΙ** (Diagnosed Diseases):
  - ΣΗΜΕΙΩΣΕΙΣ** (Notes):
  - ΣΗΜΑΝΤΙΚΟΤΗΤΑ** (Priority):  Μη Επίγου  Επίγου  Υπερσίγου
  - Ο ΙΑΤΡΟΣ ΚΕΝΤΡΟΥ** (Center Doctor):
- Right Panel:** Contains information about ambulance status and location.
  - ΑΠΟΣΤΟΛΗ ΚΑΡΤΑΣ** (Card Dispatch): 12:51:36
  - ΠΑΡΑΛΑΒΗ ΚΑΡΤΑΣ** (Card Reception): 12:27:43
  - ΚΩΔ. ΑΣΘΕΝΟΦΟΡΟΥ** (Ambulance Code):
  - ΘΕΣΗ ΑΣΘΕΝΟΦΟΡΟΥ** (Ambulance Location):
  - ΔΙΑΣΒΑΣΗ** (Reading): 00:00:00
  - ΑΦΕΣΗ ΤΟΠΟ ΑΤΥΧ** (Location Release): 00:00:00
  - ΑΝΑΓΡΟΦΗ** (Recording): 00:00:00
  - ΑΦΕΣΗ ΝΟΣΟΚΟΜΕΙΟΥ** (Hospital Release): 00:00:00
  - ΤΕΛΟΣ** (End): 00:00:00
  - ΒΑΣΗ** (Base): 00:00:00
  - ΤΟΠΟΣ ΠΑΡΑΔΟΣΗΣ ΑΣΘΕΝΟΥΣ** (Patient Handover Location):
  - ΕΙΔΟΣ ΑΤΥΧΗΜΑΤΟΣ** (Accident Type):
  - ΕΚΤΙΜΩΜΕΝΟ ΠΑΡΩΝΟΤΟΣ** (Estimated Present):
  - ΛΟΓΟΙ ΑΚΥΡΩΣΗΣ** (Cancellation Reasons):
    - ΑΚΥΡΩΣΗ
    - ΔΕΝ ΑΝΕΥΡΕΘΗ
    - ΑΚΥΡΗ ΚΑΘΟΣΟΝ
    - ΑΡΝΗΤΗ ΜΕΤΑΦΟΡΑΣ
    - ΑΝΑΧΩΡΗΣΗ ΜΕ ΙΧ
    - ΕΛΛΗΝΗ ΑΣΘΕΝΟΦΟΡΟΥ
    - ΚΑΛΕΝΑ
    - ΑΝΑΒΟΛΗ
  - ΣΗΜΕΙΩΣΕΙΣ** (Notes):
  - ΙΑΤΡΟΣ ΜΟΝΑΔΑΣ** (Unit Doctor): Προσθήκη Ασθενοφόρου
  - Buttons:** Επισύναψη, Προσωπικό Κλείσιμο, Ολοκλήρωση και Κλείσιμο
- Bottom Panel:**
  - Ανοχτές Καρτέλες** (Outstanding Cards) table:
 

A/A	Όνομα	Διεύθυνση	Διάρκεια
292	ΘΕΣΣΑΛΟΝΙΚΗΣ	ΕΠΙΝΑΤΙΑΣ	+23-39-15
  - Ασθενοφόρο** (Ambulance) table:
 

Διαθέσιμο	Εκτός κίνησης
A02	A01 Σε Περιστατικό
A03	A05 Μη διαθέσιμο
A04	A06 Μη διαθέσιμο
	A07 Μη διαθέσιμο
	A08 Μη διαθέσιμο
	A09 Μη διαθέσιμο
  - Υπόμνημα** (Legend) table:
 

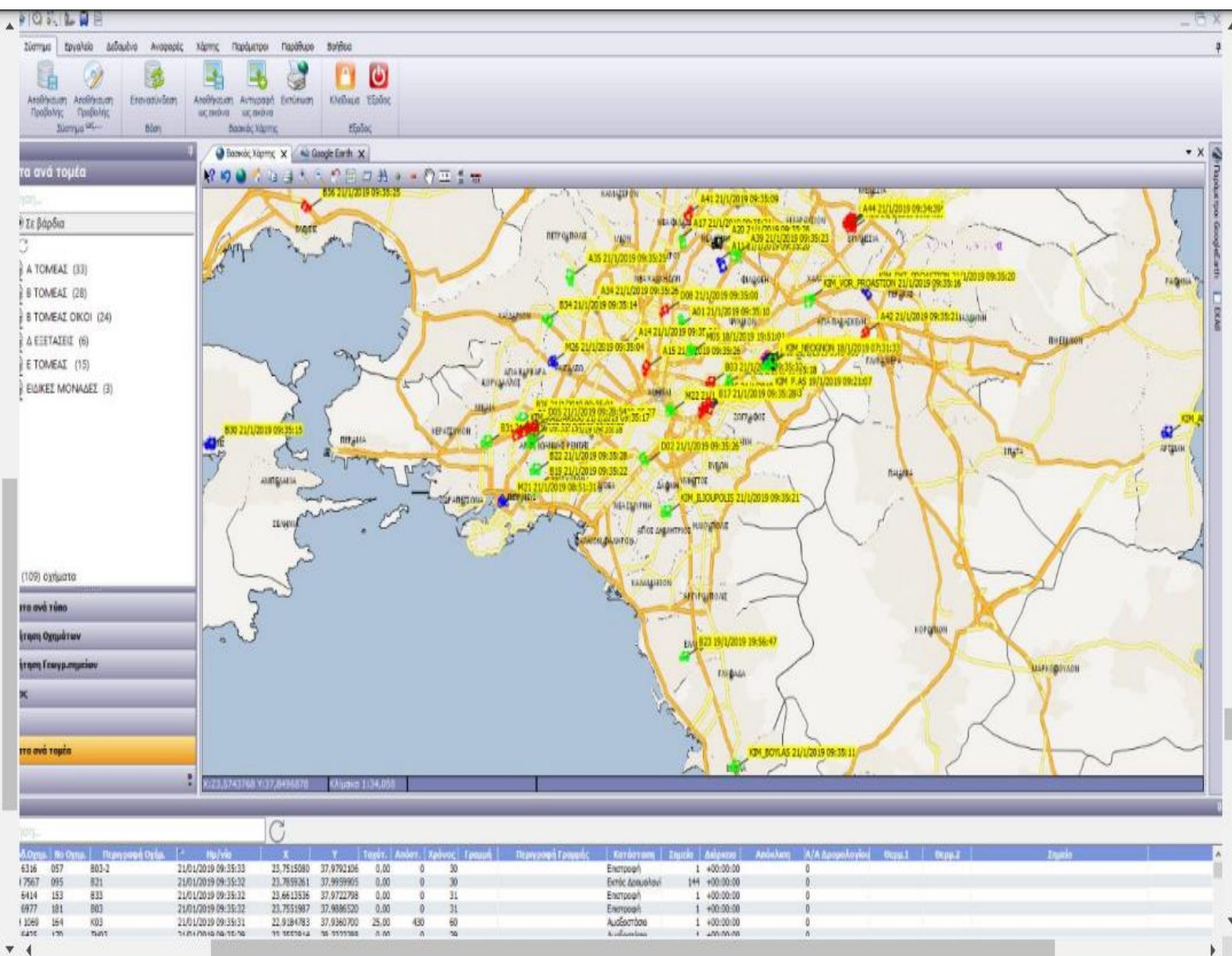
Από Τηλεφ.	Παράκληση	Διατίθεται	Με-Οι Νέοι

An orange callout box points to the "Εκκρεμεί" table with the text: "Το περιστατικό μεταφέρθηκε στις ανοχτές καρτέλες" (The case was transferred to the outstanding cards).

Εικόνα 11. Τυχαίο παράδειγμα ανοιχτής κάρτας από την οπτική του ασυρματιστή. Στο λειτουργικό χρησιμοποιείται μενού πολλαπλών ανοιχτών καρτελών για να μπορεί ο εκάστοτε ασυρματιστής να γνωρίζει ποιο περιστατικό είναι υπό την εποπτεία του ανατύπωση από «ηλεκτρονική διαχείριση ασθενών 166» σε συνεργασία με την εταιρεία Link Technologies (2008).

Εφόσον τα δεδομένα καταγραφούν και η κινητή ιατρική μονάδα ή το ασθενοφόρο ξεκινήσει την πορεία προς ένα περιστατικό μέσω του συστήματος τηλεματικής τεχνολογίας και της χρήσης χάρτη WGS84 δίνεται η δυνατότητα εντοπισμού της μονάδος. «Ηλεκτρονική διαχείριση ασθενών 166» σε συνεργασία με την εταιρεία Link Technologies (2008) (Εικόνα 12).





Εικόνα 12. Παράδειγμα συστήματος τηλεματικής τεχνολογίας στον τομέα των Αθηνών. Ανατύπωση από «ηλεκτρονική διαχείριση ασθενών 166» σε συνεργασία με την εταιρεία Link Technologies (2008).

Συνοψίζοντας μέσω του λειτουργικού «ηλεκτρονική διαχείριση ασθενών 166» δημιουργούνται ιατρικά δεδομένα τα οποία καταχωρούνται σε μια βάση δεδομένων. Το πρόγραμμα «ηλεκτρονική διαχείριση ασθενών 166» διαθέτει πλατφόρμα αναζήτησης που είναι συνδεδεμένη με την βάση δεδομένων από όπου μπορούν να αντληθούν πληροφορίες όπως είναι το ιστορικό των περιστατικών με όλες τις πληροφορίες που καταγράφηκαν από το τηλεφωνικό κέντρο τους ασυρματιστές και

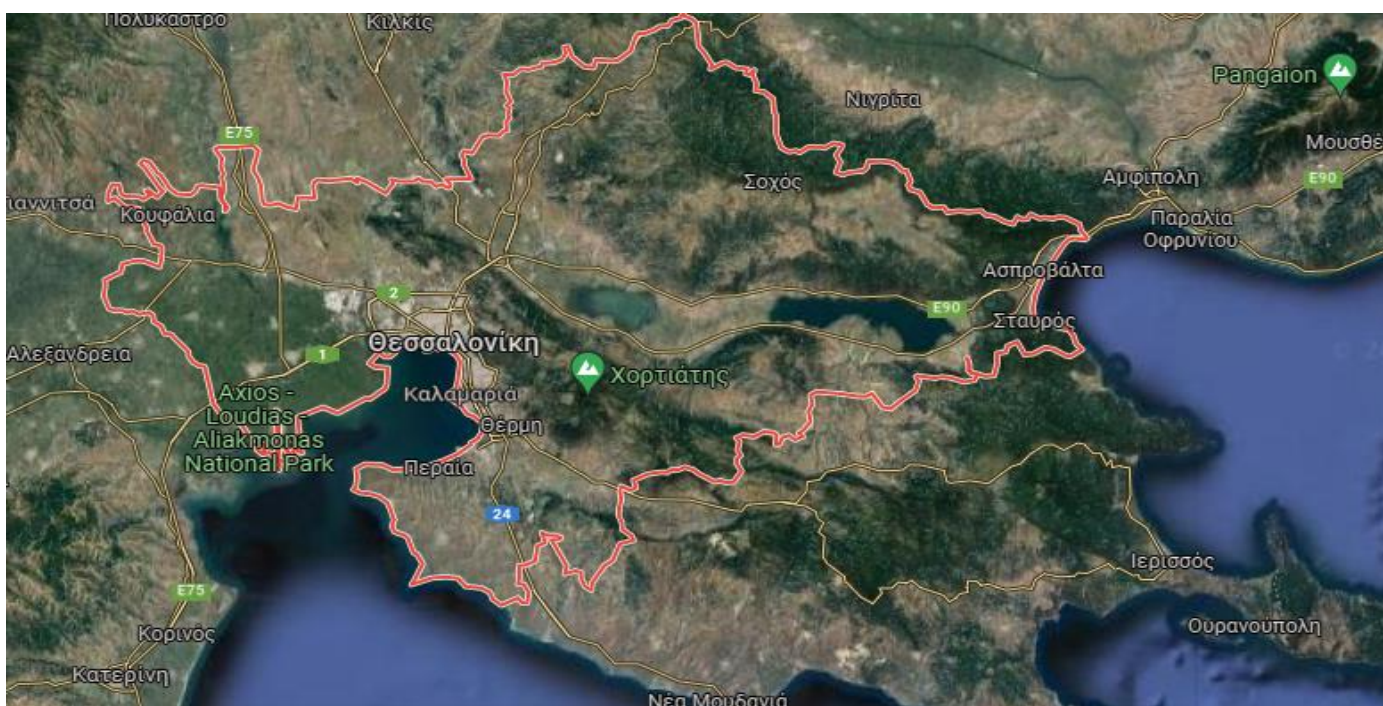
τους συντονιστές, βοηθώντας στην επεξεργασία συμβάντων και την λήψη αποφάσεων όσο αναφορά μελλοντικά περιστατικά. «Ηλεκτρονική διαχείριση ασθενών 166» σε συνεργασία με την εταιρεία Link Technologies (2008).

## 7.3 Μελέτη περίπτωσης

### 7.3.1 Υλικό και Μέθοδοι

Η μελέτη βασίστηκε σε δεδομένα που χορηγήθηκαν από το Ε.Κ.Α.Β. Θεσσαλονίκης κατόπιν ειδικής άδειας με αριθμό αίτησης 13205/08-10-21, αριθμό πρωτοκόλλου: 13518 και τίτλο διαβίβασης: Δευτερογενής διακομιδές (έγκυρες και άκυρες) του τομέα Θεσσαλονίκης Κιλκίς και Χαλκιδικής για το χρονικό διάστημα από 22/07/2019 έως και 01/01/2021.

Το σύνολο των δεδομένων αρχικά ανερχόταν στα 20340. Από αυτά επιλέχθηκαν 2132 περίπου το 8,6% του συνόλου των δεδομένων τα οποία αφορούν διακομιδές ασθενών σε μονάδες αιμοκάθαρσης του νομού Θεσσαλονίκης (Εικόνα 13) για το διάστημα 01/08/2019 έως και 01/01/2021.



Εικόνα 13. Μελέτη περίπτωσης νομός Θεσσαλονίκης. Ανατύπωση από Google Maps (2021).

Η περιφερειακή ενότητα Θεσσαλονίκης, έχει έκταση 3.683 km<sup>2</sup> ενώ βάσει των δημογραφικών στοιχείων του 2011 ο πληθυσμός της περιφέρειας ανέρχεται στους 1.101.312 κάτοικους αντιπροσωπεύοντας περίπου το 10% του συνολικού πληθυσμού της Ελλάδας (Hellenic Statistical Authority, 2011).

Το τμήμα του Ε.Κ.Α.Β. Θεσσαλονίκης περιλαμβάνει ένα εξειδικευμένο τηλεφωνικό κέντρο, που δέχεται 31000-34000 περίπου κλήσεις ανά μήνα. Για κάθε 8ωρη πρωινή και μεσημεριανή βάρδια το Ε.Κ.Α.Β. Θεσσαλονίκης έχει στην διάθεση του 20-24 ασθενοφόρα και 1-2 μοτοσυκλέτες, ενώ κατά τις νυχτερινές βάρδιες και τα σαββατοκύριακα ο αριθμός των ασθενοφόρων μειώνεται στα 16-18. Υπάρχει επίσης κινητή μονάδα αναζωογόνησης έκτακτης ανάγκης που συνοδεύεται από Ιατρό διαθέσιμη 7 μέρες την εβδομάδα όλο τον χρόνο (Μ.Ρ.1), ενώ υπάρχει και δεύτερη κινητή μονάδα (Μ.Ρ.2) η οποία είναι διαθέσιμη 5 ημέρες/εβδομάδα από Δευτέρα έως Παρασκευή. Το προσωπικό των ασθενοφόρων των μοτοσυκλετών και του τηλεφωνικού κέντρου απαρτίζεται από ειδικά εκπαιδευμένους διασώστες οι οποίοι έχουν ολοκληρώσει επιτυχώς την 2ετή εκπαίδευση τους στο ΙΕΚ διασώστη του Ε.Κ.Α.Β. (Aslanidis 2016, Γιαννόπουλος et al 2018).

Στην παρούσα μελέτη οι διακομιδές σε μονάδες τεχνητού νεφρού που πραγματοποιήθηκαν αφορούσαν τα νοσοκομεία Α.Χ.Ε.Π.Α., Άγιος Πάυλος, Παπανικολάου, Παπαγεωργίου, Ιπποκράτειο και το 424 Στρατιωτικό νοσοκομείο.

Οι αναλύσεις των δεδομένων πραγματοποιήθηκαν με το Microsoft Excel 2016.

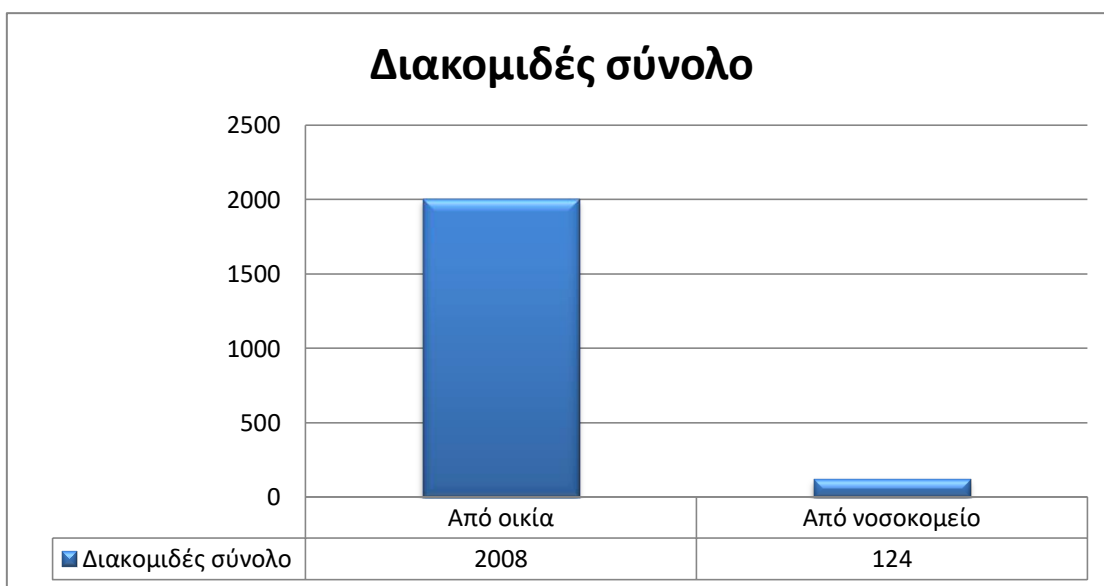
Η μελέτη πραγματοποιήθηκε: 1) με την διαδικασία της ανωνυμοποίησης, όπως αυτό ορίζεται από τον Γενικό Κανονισμό Προσωπικών Δεδομένων (GDPR) ΕΕ 2016/679, που καθιστά αδύνατο να προσδιοριστεί το υποκείμενο των δεδομένων και 2) με την διαδικασία της ψευδονυμοποίησης επομένως δεν χρειάστηκε γραπτή ενημέρωση και συγκατάθεση από τους ασθενείς.

### **7.3.2 Αποτελέσματα**

Είναι σύνηθες φαινόμενο για τους ασθενείς που χρίζουν αιμοκάθαρση να χρησιμοποιούν ιατρικές υπηρεσίες έκτακτης ανάγκης όπως επίσης συχνή είναι και η

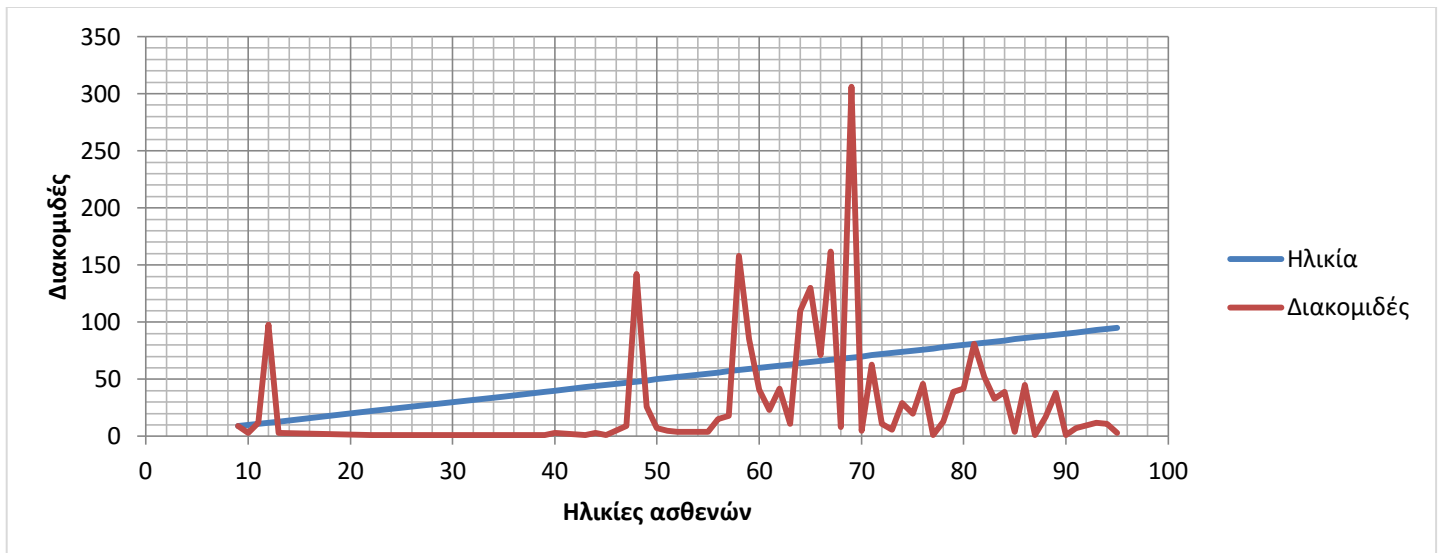
μεταφορά τους στο τμήμα επειγόντων περιστατικών λόγω πολλών παραγόντων (Chow etal 2014, Harel etal 2015, Loran etal 2002, Ronksley etal 2017).

Λόγω της σοβαρότητας της κατάστασης των αιμοκαθαιρόμενων ασθενών, δεν δύναται να μετακινηθούν με δικό τους όχημα, έτσι βάσει του νόμου 2071 (ΦΕΚ 123, 1992) το Ε.Κ.Α.Β. συμβάλει καθοριστικά στην μεταφορά μη επειγόντων περιστατικών, τα οποία δεν δύναται να μεταφερθούν με άλλο τρόπο. Από το σύνολο των διακομιδών οι 2008 κλήσεις αφορούσαν την μεταφορά ασθενών από την οικία τους σε μονάδες αιμοκάθαρσης ενώ 124 κλήσεις την μεταφορά ασθενών από τις μονάδες αιμοκάθαρσης στην οικία τους (Γράφημα 1).



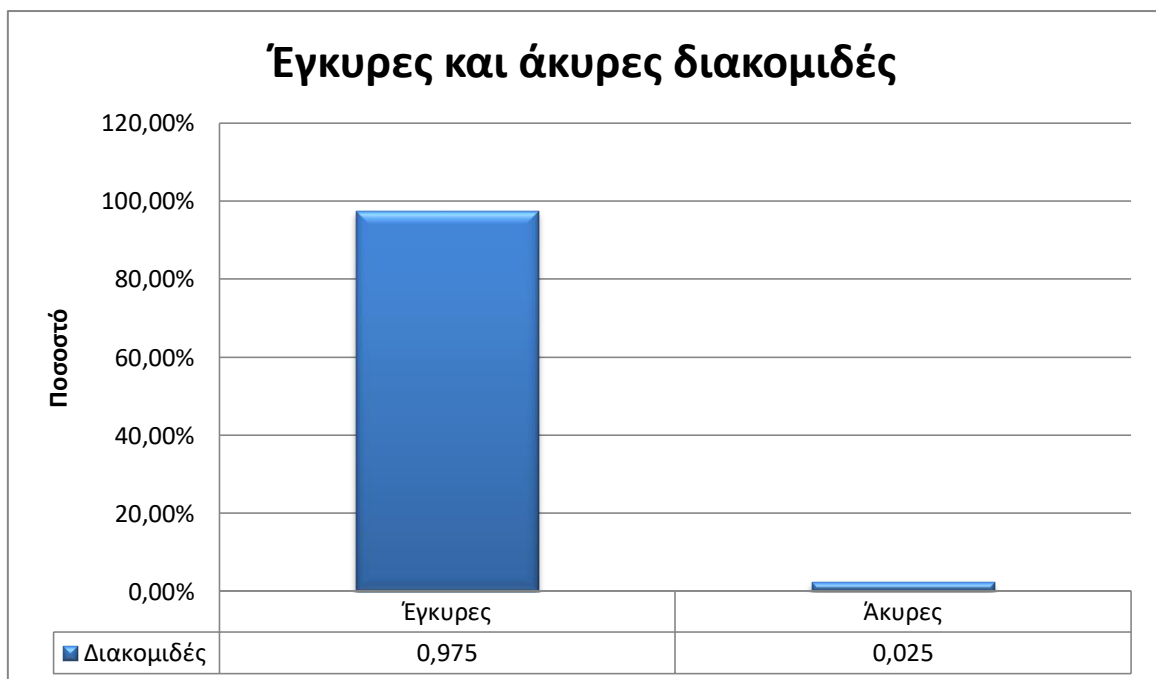
Γράφημα 1. Σύνολο διακομιδών αιμοκαθαιρόμενων ασθενών. Στις περισσότερες των περιπτώσεων το ασθενοφόρο περιμένει τους ασθενείς να ολοκληρώσουν την αιμοκάθαρση επιστρέφοντας τους στην οικία τους.

Η μέση ηλικία των ασθενών υπολογίστηκε στα 65,5 έτη ενώ το εύρος των ηλικιών παρουσιάζεται στο Γράφημα 2.



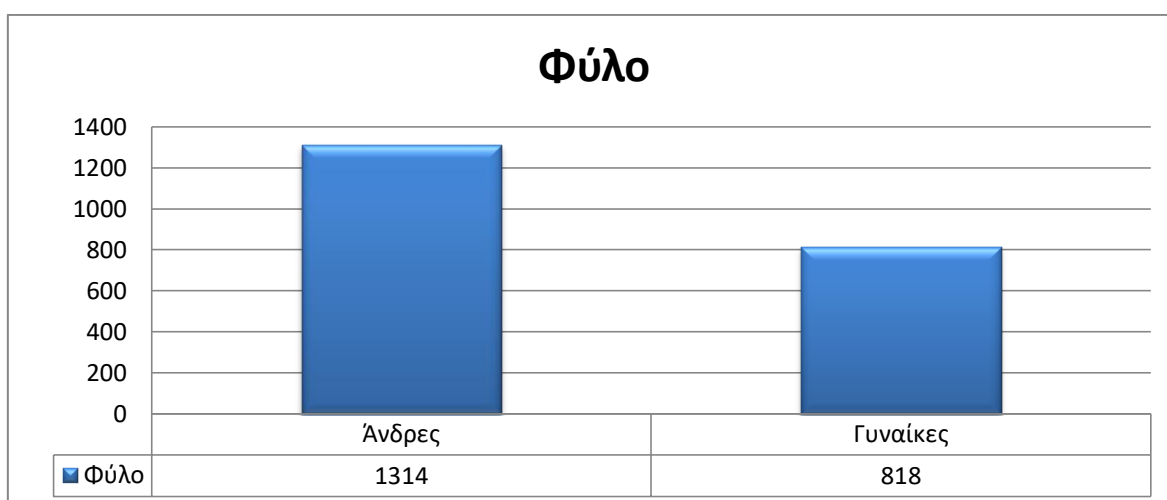
Γράφημα 2. Διάγραμμα διασποράς διακομιδών σε σχέση με την ηλικία των ασθενών από 9 έως 95 έτη για τον νομό Θεσσαλονίκης.

Σε κάποιες περιπτώσεις μπορεί να δημιουργηθούν επιπλοκές και ο ασθενής να χρειαστεί να παραμείνει στην μονάδα αιμοκάθαρσης. Σε αυτή την περίπτωση είναι αναγκαίο να γίνει κλήση ξανά στο Ε.Κ.Α.Β. από τον θεράποντα ιατρό ο οποίος θα κρίνει πότε είναι δυνατό να προγραμματιστή νέο εξιτήριο για την διακομιδή του ασθενούς στην οικία του (Γράφημα 3).



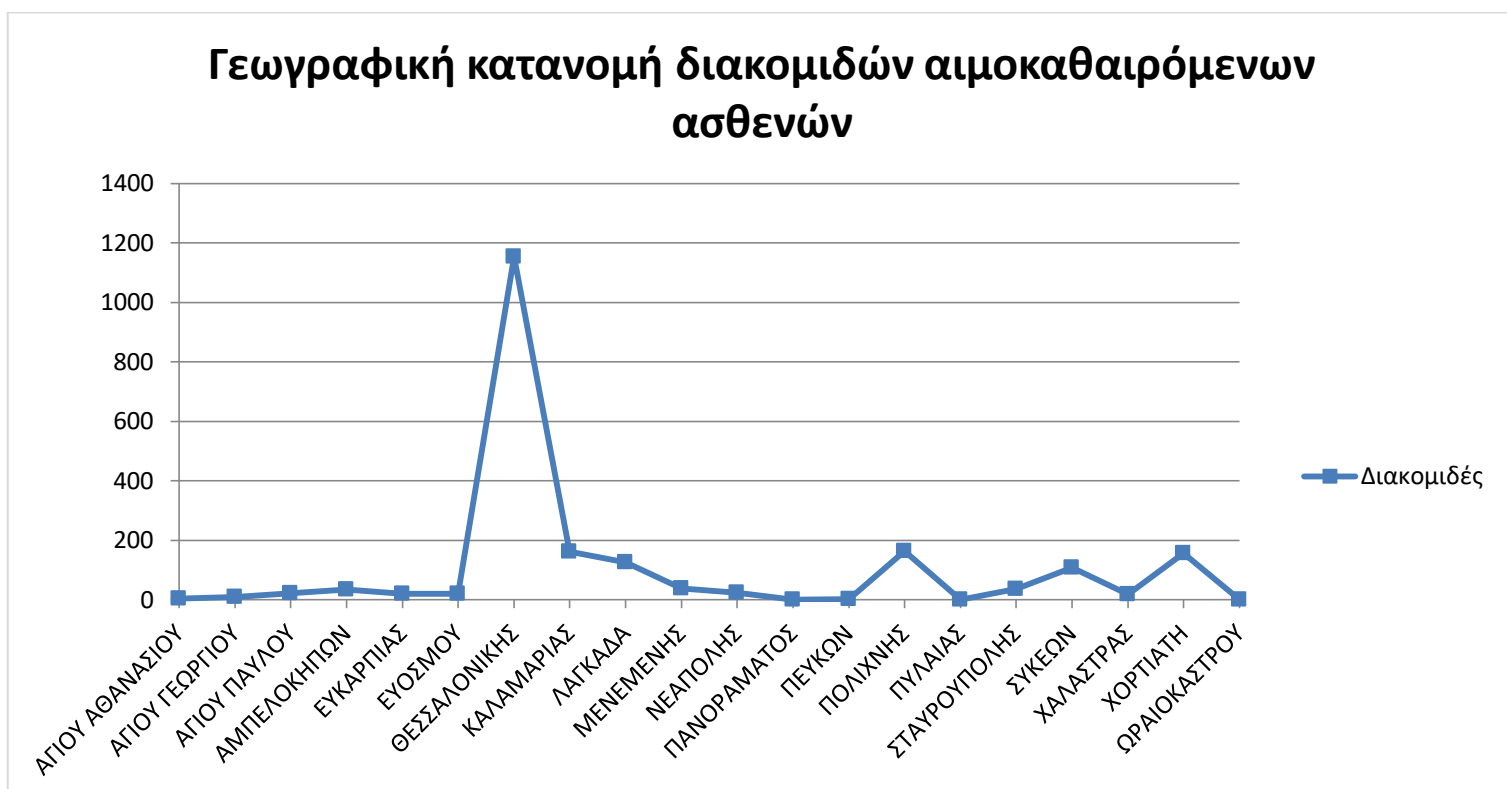
Γράφημα 3. Ποσοστό έγκυρων και άκυρων διακομιδών. Στον παραπάνω πίνακα 97,5% έγκυρες και 2,5% άκυρες διακομιδές.

Από αυτούς τους ασθενείς 1314 ήταν άνδρες ενώ 818 γυναίκες (Γράφημα 4).



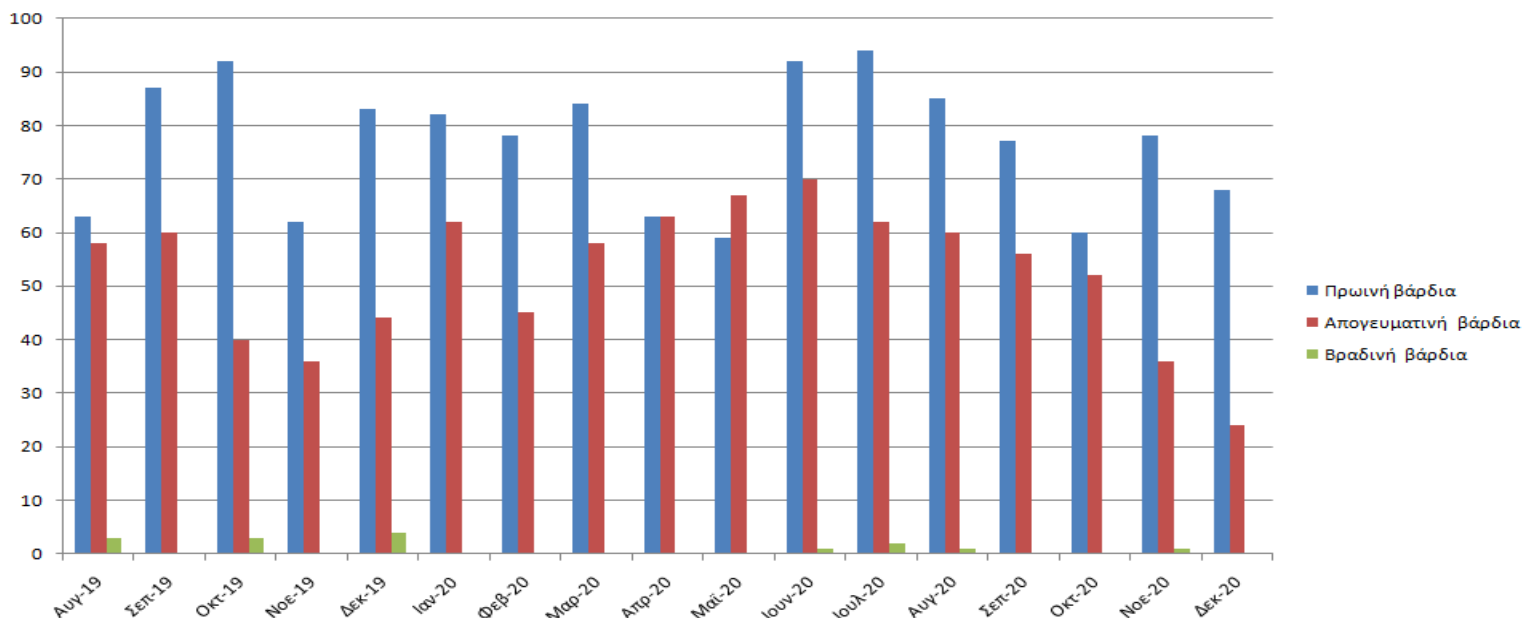
Γράφημα 4. Σύνολο διακομιδών ανάλογα με το φύλο.

Η γεωγραφική κατανομή των διακομιδών αφορά τον νομό Θεσσαλονίκης και αντιστοιχούν στις παρακάτω περιοχές (Γράφημα 5).



Γράφημα 5. Η γεωγραφική κατανομή των ασθενών που χρήζουν μεταφορά σε μονάδες τεχνητού νεφρού στον νομό Θεσσαλονίκης.

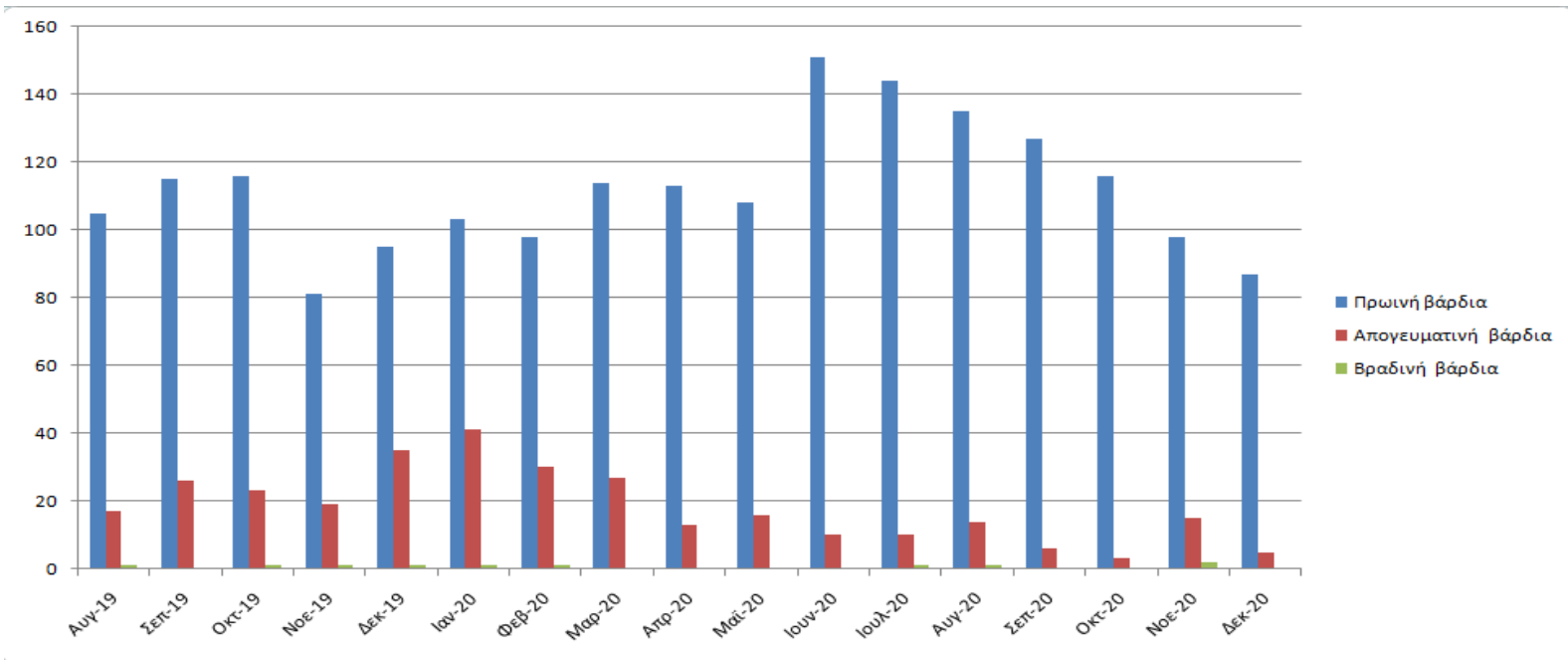
Οι βάρδιες του Ε.Κ.Α.Β. Θεσσαλονίκης πραγματοποιούνται με κυλιόμενο δωρο όσο αναφορά τους τηλεφωνητές και τους ασυρματιστές λόγω της φύσεως της εργασίας που επιτελούν. Το ιατρικό προσωπικό του Ε.Κ.Α.Β. Θεσσαλονίκης αποτελείται από 8 ιατρούς με εξειδίκευση στην προνοσοκομειακή φροντίδα. Το ιατρικό προσωπικό πραγματοποιεί εφημερίες με κυλιόμενο ωράριο εναλλαγής ιατρών ανά 24ωρο ούτως ώστε να υπάρχει κάθε μέρα καθ' όλη την διάρκεια του έτους ειδικά εκπαιδευόμενο προσωπικό άμεσα διαθέσιμο για την διαχείριση επείγουσων καταστάσεων. Το ωράριο εργασίας του υπόλοιπου παραϊατρικού προσωπικού είναι η δωρη βάρδια με ωράρια 07:00 έως 15:00 για την πρωινή βάρδια από της 15:00 έως της 23:00 για την απογευματινή βάρδια και 23:00 έως 07:00 για την βραδινή βάρδια. Ο μεγαλύτερος όγκος κλήσεων πραγματοποιήθηκε κατά την πρωινή και την απογευματινή βάρδια όπως παρατηρείται στο Γράφημα 6 ενώ κατά την βραδινή βάρδια παρατηρείται στον ίδιο πίνακα ένας μικρός αριθμός κλήσεων στο τηλεφωνικό κέντρο. Συνήθως αφορούν ασθενείς με συμπτώματα λόγω των επιπλοκών της ασθένειάς τους, οι οποίοι χρήζουν άμεση μεταφορά στο τμήμα επειγόντων περιστατικών εφημερεύοντος νοσοκομείου.



Γράφημα 6. Σύνολο κλήσεων ανά βάρδια πίνακας για την περίοδο 01/08/2019 έως και 1/1/2021.

Σε αντίθεση με τις εισερχόμενες κλήσεις που δέχεται το τηλεφωνικό κέντρο για διακομιδές ασθενών σε μονάδες τεχνητού νεφρού η ολοκλήρωση της μεταφορά των ασθενών που παρατηρείται στο Γράφημα 7 πραγματοποιείται κατά τις πρωινές ώρες

ως επί το πλείστο. Μια εξήγηση έγκειται στο γεγονός ότι κατά τις πρωινές ώρες ο θεράπων ιατρός πραγματοποιεί τις απαραίτητες εξετάσεις και κρίνει αν ο ασθενής μπορεί να αποχωρήσει από την μονάδα αιμοκάθαρσης για την οικία του. Γεγονός που ενισχύει την υπόθεση ότι στις περισσότερες των περιπτώσεων οι θεραπείες πραγματοποιούνται τις πρωινές ώρες. Στην απογευματινή όπως και στην νυχτερινή βάρδια δύναται να ολοκληρωθούν διακομιδές ασθενών οι οποίες λόγω μεγάλου φόρτου εργασίας δεν ήταν δυνατό να πραγματοποιηθούν κατά τις πρωινές βάρδιες.



Γράφημα 7. Σύνολο ολοκλήρωση διακομιδών ανά βάρδια από 01/08/2019 έως και 1/1/2021.



Στους πίνακες 2 και 3 παρουσιάζονται τα δεδομένα των κλήσεων και των ολοκληρώσεων των διακομιδών σε αριθμούς:

t	Μήνας	Κλήσεις στο Ε.Κ.Α.Β.
1	Αυγ-19	118
2	Σεπ-19	143
3	Οκτ-19	136
4	Νοε-19	101
5	Δεκ-19	131
6	Ιαν-20	145
7	Φεβ-20	129
8	Μαρ-20	144
9	Απρ-20	130
10	Μαΐ-20	123
11	Ιουν-20	161
12	Ιουλ-20	157
13	Αυγ-20	150
14	Σεπ-20	137
15	Οκτ-20	94
16	Νοε-20	133

Πίνακας 2. Δεδομένα εισερχόμενων κλήσεων στο Ε.Κ.Α.Β. Θεσσαλονίκης ξεκινώντας από τον Αύγουστο, Σεπτέμβριο, Οκτώβριο, Νοέμβριο του 2019 και καταλήγουν Αύγουστο, Σεπτέμβριο, Οκτώβριο, Νοέμβριο του 2020.

t	Μήνας	Ολοκλήρωση μεταφοράς από το Ε.Κ.Α.Β.
1	Αυγ-19	129
2	Σεπ-19	142
3	Οκτ-19	137
4	Νοε-19	101
5	Δεκ-19	131
6	Ιαν-20	144
7	Φεβ-20	128
8	Μαρ-20	141
9	Απρ-20	128
10	Μαΐ-20	126
11	Ιουν-20	161
12	Ιουλ-20	155
13	Αυγ-20	150
14	Σεπ-20	133
15	Οκτ-20	118
16	Νοε-20	115

Πίνακας 3. Ομοίως για τα δεδομένα ολοκλήρωσης μεταφοράς διακομιδών από το Ε.Κ.Α.Β. Θεσσαλονίκης.

Στην συνέχεια υπολογίστηκε ο κινητός μέσος όρος για την εξομάλυνση της καμπύλης των τετραμήνων με την χρήση της εντολής του excel average βασιζόμενος στον μαθηματικό τύπο:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n t_i = \frac{1}{n} (t_1 + \dots + t_n)$$

Στην συνάρτηση η παρατήρηση είναι η  $t_i$  και το πλήθος  $n$ .

Το επόμενο βήμα μετά τον υπολογισμό του κινητού μέσου όρου, είναι ο υπολογισμός του κεντραρισμένου κινητού μέσου όρου (Centered Moving Average) με σκοπό την περαιτέρω εξομάλυνση της καμπύλης ούτως ώστε να αναδειχτούν περισσότερα μοτίβα που αφορούν τα δεδομένα τα οποία θα βοηθήσουν στις προβλέψεις. Η μαθηματική πράξη πραγματοποιείται με την χρήση της εντολής του excel average βασιζόμενη στον μαθητικό τύπο:

$$CMA_n = \frac{x_1 + \dots + x_n}{n}$$

Ο κινητός μέσος όρος του πρώτου τετραμήνου προκύπτει από τον υπολογισμό κελιών των κλήσεων για τους μήνες Αυγ-19, Σεπ-19, Οκτ-19, Νοε-19 σε αριθμούς δεδομένων κλήσεων  $118+143+136+101/4=124,50$ . Στην ουσία ο κινητός μέσος όρος 124,5 τοποθετείται ενδιάμεσα στους μήνες, Σεπ-19, Οκτ-19, με αριθμούς κλήσεων 143, 136 όμως στο excel δεν δίνεται αυτή η δυνατότητα. Η συνέχεια είναι πανομοιότυπη αυτή την φορά κατεβαίνοντας ένα κελί ανά τετράδα προς τα κάτω δηλαδή, Σεπ-19, Οκτ-19, Νοε-19, Δεκ-19 όπου η μαθηματική πράξη είναι  $143+136+101+131/4=127,75$  κ.ο.κ. (στο τελευταίο κελί 16 αφαιρείται το αποτέλεσμα του κινητού μέσου όρου για τον λόγο ότι στο τελευταίο τετράμηνο λείπει ένας μήνας). Ο κεντραρισμένος κινητός μέσος όρος υπολογίζεται πάλι με την εντολή average αυτή την φορά η διαίρεση αφορά τον κινητό μέσο όρο ξεκινώντας υπολογισμούς του πρώτου και του δεύτερου ήδη υπολογισμένου κινητού μέσου όρου,  $124,50/127,75=126,13$  κ.ο.κ.. Τα αποτελέσματα των βημάτων σε αριθμούς για την περίοδο από 01/08/2019 έως και 1/1/2021 στους πίνακες 4 και 5:

t	Μήνας	Κλήσεις	Κινητός μέσος όρος	Κεντραρισμένος κινητός μέσος όρος
1	Αυγ-19	118		
2	Σεπ-19	143		
3	Οκτ-19	136	124,50	126,13
4	Νοε-19	101	127,75	128,00
5	Δεκ-19	131	128,25	127,38
6	Ιαν-20	145	126,50	131,88
7	Φεβ-20	129	137,25	137,13
8	Μαρ-20	144	137,00	134,25
9	Απρ-20	130	131,50	135,50
10	Μαϊ-20	123	139,50	141,13
11	Ιουν-20	161	142,75	145,25
12	Ιουλ-20	157	147,75	149,50
13	Αυγ-20	150	151,25	142,88
14	Σεπ-20	137	134,50	131,50
15	Οκτ-20	94	128,50	
16	Νοε-20	133		

Πίνακας 4. Αποτελέσματα υπολογισμού κινητού μέσου όρου, κεντραρισμένου κινητού μέσου όρου, όσο αναφορά τα δεδομένα κλήσεων στο Ε.Κ.Α.Β. Θεσσαλονίκης.

t	Μήνας	Ολοκλήρωση μεταφοράς	Κινητός μέσος όρος	Κεντραρισμένος κινητός μέσος όρος
1	Αυγ-19	129		
2	Σεπ-19	142		
3	Οκτ-19	137	127,25	127,50
4	Νοε-19	101	127,75	128,00
5	Δεκ-19	131	128,25	127,13
6	Ιαν-20	144	126,00	131,00
7	Φεβ-20	128	136,00	135,63
8	Μαρ-20	141	135,25	133,00
9	Απρ-20	128	130,75	134,88
10	Μαϊ-20	126	139,00	140,75
11	Ιουν-20	161	142,5	145,25
12	Ιουλ-20	155	148,00	148,88
13	Αυγ-20	150	149,75	144,38
14	Σεπ-20	133	139,00	134,00
15	Οκτ-20	118	129,00	
16	Νοε-20	115		

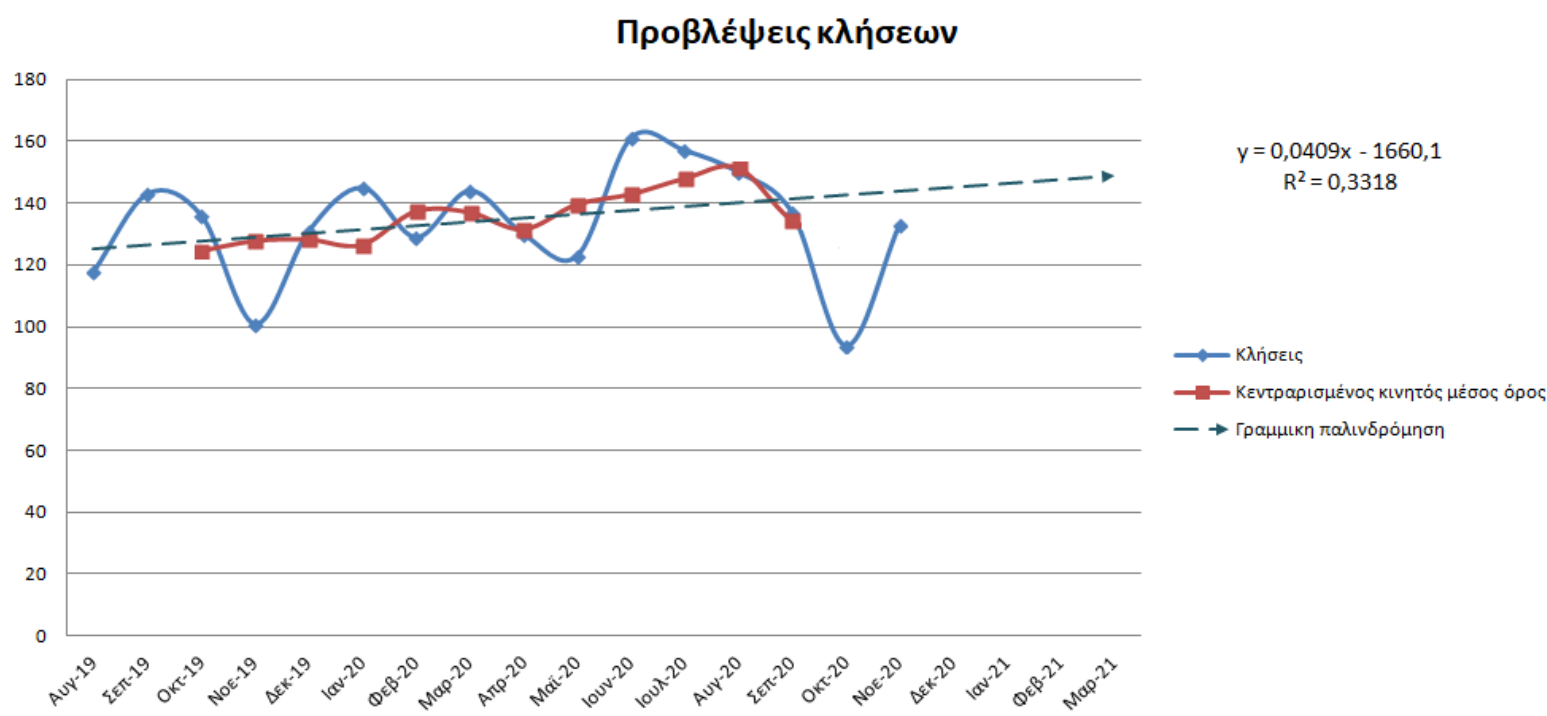
Πίνακας 5. Αποτελέσματα υπολογισμού κινητού μέσου όρου, κεντραρισμένου κινητού μέσου όρου, όσο αναφορά τα δεδομένα ολοκλήρωσης διακομιδών από το Ε.Κ.Α.Β. Θεσσαλονίκης.

Οι προβλέψεις πραγματοποιήθηκαν με την χρήση της γραμμικής παλινδρόμησης μια ευθεία η οποία υπολογίζεται από τον μαθηματικό τύπο:

$$y = mx + b$$

Όπου  $y$  είναι η εξαρτημένη μεταβλητή κλήσεις και οι ολοκληρώσεις διακομιδών,  $x$  οι ανεξάρτητες τιμές του κεντραρισμένου μέσου κινητού μέσου όρου,  $m$  οι συντελεστές που αντιστοιχούν σε κάθε τιμή  $x$  και δείχνουν την τάση της γραμμής και  $b$  η σταθερά τιμή σε αυτήν την ανάλυση το σύνολο του κεντραρισμένου μέσου κινητού μέσου όρου.

Για κλήσεις που ενδέχεται να πραγματοποιηθούν στο Ε.Κ.Α.Β. Θεσσαλονίκης για την διακομιδή ασθενών σε μονάδα αιμοκάθαρσης παρουσιάζονται στο Γράφημα 8:

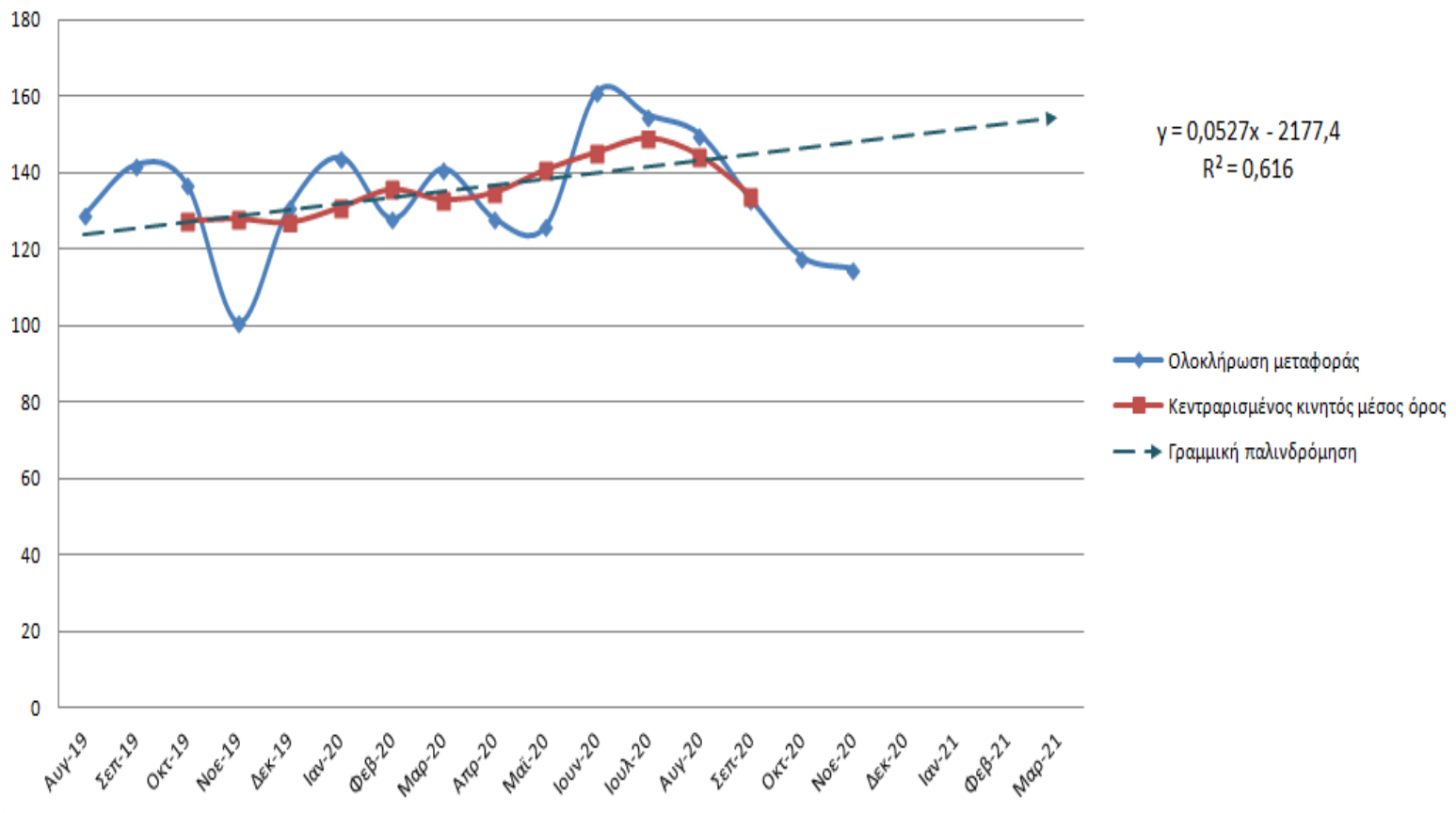


Γράφημα 8. Διάγραμμα χρονοσειράς και προβλέψεων ενδεχομένων κλήσεων στο Ε.Κ.Α.Β. Θεσσαλονίκης για το τετράμηνο 01/12/2020 έως και 31/03/2021.

Στο Γράφημα 8 παρατηρείται μια σχετικά σταθερή αυξητική πορεία κλήσεων όμοια με αυτή των προηγούμενων μηνών χωρίς μεγάλες διακυμάνσεις για το τετράμηνο 01/12/2020 έως και 31/03/2021. Το  $R^2=0,3318$  δείχνει την αναλογία 33,18% των κλήσεων που θα πραγματοποιηθούν σε βάθος χρόνου.

Όσο αναφορά τις προβλέψεις ολοκλήρωσης ενδεχόμενων διακομιδών ασθενών από τις μονάδες αιμοκάθαρσης στην οικία των ασθενών παρουσιάζουν μια παρόμοια εικόνα από αυτή των κλήσεων (Γράφημα 9).

### Προβλέψεις ολοκλήρωσης διακομιδών



Γράφημα 9. Διάγραμμα χρονοσειράς και προβλέψεων ολοκλήρωσης ενδεχόμενων διακομιδών από το Ε.Κ.Α.Β. Θεσσαλονίκης για το τετράμηνο 01/12/2020 έως και 31/03/2021.

Η πρόβλεψη όπως φανερώνει το Γράφημα 9 ακολουθεί για τους μήνες 01/12/2020 έως και 31/03/2021 μια σχετικά σταθερή αυξητική πορεία με αυτή των προηγούμενων μηνών. Το  $R^2=0,616$  δείχνει την αναλογία 61,6% των ολοκληρώσεων των διακομιδών που θα πραγματοποιηθούν σε βάθος χρόνου.

## 8. Συζήτηση

Στην παρούσα διπλωματική εργασία πραγματοποιήθηκε μια προσπάθεια μελλοντικών προβλέψεων από 01/08/2019 έως και 31/03/2021 με πραγματικά δεδομένα τα οποία χορηγήθηκαν από το Ε.Κ.Α.Β. Θεσσαλονίκης για το χρονικό διάστημα 01/08/2019 έως και 01/01/2021. Τα δεδομένα αφορούσαν περιστατικά διακομιδών για ασθενείς οι οποίοι έχουν ανάγκη μεταφοράς σε μονάδες τεχνητού νεφρού. Τα άτομα με νεφρική ανεπάρκεια χρήζουν ειδικής μεταχείρισης κατά την μεταφορά τους στις μονάδες αιμοκάθαρσης. Οι διακομιδές συνήθως είναι προγραμματισμένες και η ιατρική κινητή μονάδα ή το ασθενοφόρο επιστρέφει το γρηγορότερο δυνατό τον ασθενή στην οικία του.

Σε αυτή τη μελέτη, έγινε προσπάθεια εντοπισμού προγνωστικών παραγόντων όσο αναφορά τις κλήσεις που θα πραγματοποιηθούν στο Ε.Κ.Α.Β. Θεσσαλονίκης και την ολοκλήρωση των διακομιδών από τις μονάδες αιμοκάθαρσης προς την οικία των ασθενών για το χρονικό διάστημα 01/12/2020 έως και 31/03/2021.. Χρησιμοποιήθηκαν αυτές οι πληροφορίες για να αναπτυχθεί ένα μοντέλο πρόβλεψης με την μαθηματική τεχνική ανάλυσης των χρονοσειρών, του υπολογισμού του κινητού μέσου όρου, του κεντραρισμένου κινητού μέσου όρου και της γραμμικής παλινδρόμησης όπου εντοπίστηκε όπως παρατηρείται στο Γράφημα 8 και 9 ότι θα υπάρξει μια αυξητική πορεία σε σχέση με τα προηγούμενα τετράμηνα.

Βάσει των αποτελεσμάτων η αξιοπιστία του μοντέλου θεωρείται σημαντική κατά 33,18% όσο αναφορά τις κλήσεις που θα πραγματοποιηθούν στο τηλεφωνικό κέντρο του Ε.Κ.Α.Β. και 61,6% για τις ολοκληρώσεις των διακομιδών. Το χαμηλό ποσοστό της γραμμικής παλινδρόμησης στο 33,18% όσο αναφορά τις κλήσεις που θα πραγματοποιηθούν στο τηλεφωνικό κέντρο του Ε.Κ.Α.Β. οφείλεται κύριος στον περιορισμό των δεδομένων όπως για παράδειγμα για πολύπλοκες αναλύσεις δεδομένων και προβλέψεις που αφορά το Ε.Κ.Α.Β. Θεσσαλονίκης χρειάζεται περαιτέρω έρευνα σε μοντέλα αναλύσεων και προβλέψεων με σκοπό τον εμπλουτισμό της βιβλιογραφίας στον συγκεκριμένο τομέα. Η ανάλυση προβλέψεων με την μέθοδο των χρονοσειρών είναι ένα πρώτο βήμα και βάση για την ανάπτυξη περαιτέρω ερευνών σε νέα μοντέλα πρόβλεψης με την χρήση μηχανικής μάθησης και τεχνητής νοημοσύνης. Όσο αναφορά τις ολοκληρώσεις των διακομιδών η πρόβλεψη είναι

αρκετά ενθαρρυντική κατά 61,6% αποδεδειγμένα ασθενοφόρα που προορίζονται για επείγοντα περιστατικά για το χρονικό διάστημα 01/12/2020 έως και 31/03/2021.

Η κρίση της πανδημίας του κορονοϊού SARS-CoV-2 που έκανε την εμφάνιση της στις 22/02/2020 έπαιξε τεράστιο ρόλο στις συνθήκες κάτω από τις οποίες διακομίζονται τα περιστατικά. Όπως παρατηρείται στο διάγραμμα 8 οι διακομιδές ασθενών σε μονάδες αιμοκάθαρσης τον Σεπτέμβριο και τον Οκτώβριο του 2020 παρουσιάζουν καμπή στο διάγραμμα φανερώοντας απότομη μείωση των διακομιδών. Σύμφωνα με τον ΕΟΔΥ (2020) την συγκεκριμένη περίοδο ξεκίνησε η απότομη αύξηση των κρουσμάτων κορονοϊού πανελλαδικά οδηγώντας στην μείωση κινητικότητας ασθενών σε μονάδες αιμοκάθαρσης.

## 9. Συμπεράσματα

Είναι εμφανές ότι μέσω αναλύσεων των δεδομένων του μοντέλου που προτάθηκε αυτού της ανάλυσης των χρονοσειρών μπορούν να παρατηρηθούν μελλοντικές τάσεις για το σύστημα της προνοσοκομειακής ιατρικής. Σε γενικότερο πλάνο όλες οι υπηρεσίες υγείας μπορούν να χαρακτηριστούν ως ένας γενικότερος αλγόριθμος όπου ο κάθε σύνδεσμος είναι εξαιρετικά χρήσιμος και κρίσιμος όσο αναφορά την εύρυθμη λειτουργία τους. Αυτό που παρατηρείτε σε εθνικό επίπεδο στην Ελλάδα παρόλες τις απίστευτες δυσκολίες που χαρακτηρίζει το υγειονομικό σύστημα, οι φιλότιμες προσπάθειες που γίνονται από ένα μέρος του υγειονομικού προσωπικού όπως για παράδειγμα αυτού της προνοσοκομειακής φροντίδας είναι αξιοσημείωτες.

Στην παρούσα διπλωματική εργασία αυτή η προσπάθεια φανερώνεται μέσω της μελέτης που διενεργήθηκε για το Ε.Κ.Α.Β. Θεσσαλονίκης διατυπώνοντας καταστάσεις προνοσοκομειακών μη επειγόντων περιστατικών. Οι ελλείψεις στα δεδομένα (όπως για παράδειγμα ο όγκος, η ευκρίνεια κ.α.) και οι περιορισμοί αντιπροσώπευσης, όπως για παράδειγμα με τα σημερινά μέσα είναι δύσκολο να διευκρινισθεί στις έρευνες που διατυπώθηκαν λεπτομέρειες για τους ασθενείς (επικείμενα νοσήματα κ.α.) μειώνοντας την ευκρίνεια των αποτελεσμάτων και των εξηγήσεων όπως παρατηρείται στην ανάλυση των εισερχομένων κλήσεων κατά 33,18%, χωρίς αυτό να μειώνει την ευκρίνεια της πρόβλεψης αφού το μεγαλύτερο ποσοστό των θεραπειών είναι προγραμματισμένες και λόγω των επιπτώσεων του κορονοϊού είναι λογικό όλες οι ακυρωμένες συνεδρίες λόγω των lockdown που εφαρμόστηκαν εκείνη την περίοδο, να μετατεθούν σε μέλλοντα χρόνο.

Η σοβαρότητα της κατάστασης των ατόμων που πάσχουν από αυτό του είδους την ασθένεια δημιουργεί τεράστιο όγκο και ενδιαφέρον έρευνας που μπορεί να προκύψει όσο αναφορά τις αναλύσεις δεδομένων στον τομέα της προνοσοκομειακής φροντίδας και συγκεκριμένα για το Ε.Κ.Α.Β. Θεσσαλονίκης. Μελλοντικά θέματα έρευνας θα μπορούσαν να προκύψουν από την ανάλυση δεδομένων όχι μόνο των μη επειγόντων περιστατικών όπως παρουσιάστηκαν στην παρούσα διπλωματική εργασία αλλά και των επειγόντων περιστατικών όσο αναφορά τους χρόνους μετακίνησης ασθενοφόρων από και προς τα επείγοντα περιστατικά, την κατανομή των ασθενοφόρων όπως επίσης προβλέψεις επειγόντων περιστατικών και την τάση εμφάνισης αυτών στο νομό Θεσσαλονίκης. Επιπλέον εξαιρετικό θέμα έρευνας είναι οι προβλέψεις σε πραγματικό



χρόνο μετεγκατάστασης ενός μέσου διακομιδής του Ε.Κ.Α.Β. μέσω αναλύσεων των γεωδεδομένων τα οποία θα μπορούσαν να λειτουργήσουν ως εργαλείο για την αντιμετώπιση άσκοπων μετακινήσεων και προβλέψεις επειγόντων περιστατικών με τεράστια ευκρίνεια λόγω της ιδιαιτερότητας των εν λόγω δεδομένων.

Τέλος, όλες οι έρευνες που θα ενσωματώσουν τα Big Data θα συμβάλουν στο να υπάρξει ένα ικανοποιητικό επίπεδο υγείας σε ευρύτερο και βαθύτερο επίπεδο. Επίσης, στην παρούσα διπλωματική εργασία αποτυπώνονται γενικές και σχετικές έννοιες των προνοσοκομειακών Big Data, οι κύριες και βασικές εφαρμογές τους, εισάγοντας βασικές τεχνικές που σχετίζονται με την ανάλυση τους για την διεξαγωγή προβλέψεων. Σε Ελληνικό επίπεδο οι τρέχουσες έρευνες που αφορούν τα προνοσοκομειακά δεδομένα δεν είναι ακόμη ώριμες. Υπάρχουν πολλά προβλήματα που πρέπει να επιλυθούν όπως προαναφέρθηκε η έλλειψη βιβλιογραφίας. Για να επωφεληθούμε πλήρως από τα βαθιά μοτίβα που περιέχονται στα δεδομένα, η αποθήκευση των Big Data, η εξόρυξη και η ανάλυση τους είναι απαραίτητη. Αυτού του είδους οι τεχνολογίες θα υποστηρίξουν την έρευνα των Big Data στην υγειονομική φροντίδα γενικότερα εξυπηρετώντας περαιτέρω ένα ευρύ φάσμα εφαρμογών στους τομείς της δημόσια υγείας και της προνοσοκομειακής ιατρικής.

## Βιβλιογραφία

Abernethy, A. P., Wheeler, J. L., & Bull, J. (2011). Development of a health information technology-based data system in community-based hospice and palliative. DOI: 10.1016/j.amepre.2011.01.012

Aharon, M, M Elad, and A Bruckstein. (2006). *K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation*. Available at: <http://sites.fas.harvard.edu/~cs278/papers/ksvd.pdf>

Aitken, M., & Gauntlett, C. (2013). *Patient apps for improved healthcare: from novelty to mainstream*. IMS Institute for Healthcare Informatics. Available at: [http://ignacoriesgo.es/wp-content/uploads/2014/03/iihi\\_patient\\_apps\\_report\\_editora\\_39\\_2\\_1.pdf](http://ignacoriesgo.es/wp-content/uploads/2014/03/iihi_patient_apps_report_editora_39_2_1.pdf)

Alpaydin, E. (2014). [Introduction to Machine Learning](#), 4<sup>th</sup> edition. London: The MIT Press. ISBN: 9780262043793

Anderson, J. E., & Chang, D. C. (2015). Using electronic health records for surgical quality improvement in the era of big data. *Jama Surgery*, 150(1), 24-29. DOI: 10.1001/jamasurg.2014.947

Artasanchez, A. (2018). *Why your machine learning project will fail*. KD nuggets. Available at: <http://thedata-science.ninja/2018/07/12/why-your-machine-learning-project-will-fail>.

Asante-Korang, A., & Jacobs, J. P. (2016). Big Data and paediatric cardiovascular disease in the era of transparency in healthcare. *Cardiology in the Young*, 26(8), 1597–1602. DOI: 10.1017/S1047951116001736

Ashiwinkumar, U.M., & Anandakumar, K.R. (2010). Ethical and Legal Issues for Medical Data Mining. *International Journal of Computer Applications*, 1(28), 7-11. DOI:10.5120/520-841

Aslanidis Th., Charitidou S., Syrmou E. , Nikolaidou O., Tsioupa A., Karabelidou Ch., Pliadis K., , Matsikoudi Ch.(2016) *Pre-hospital medical emergencies services activity over 10 years in an urban setting, in Greece*. DOI:The Greek E-Journal of Perioperative Medicine 2016; 15(c): 87-100

Backonja, U., Kim, K., Casper, G. R., Patton, T., Ramly, E., & Brennan, P. F. (2012). Observations of daily living: putting the “personal” in personal health records. NI 2012: *11th International Congress on Nursing Informatics*, Montreal, Canada. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3799183>

Bagayoko, C. O., Dufour, J. C., Chaacho, S., Bouhaddou, O., & Fieschi, M. (2010). Open source challenges for hospital information system (HIS) in developing countries: A pilot project in Mali. *BMC Medical Informatics and Decision Making*, 10(22), 1-13. <https://doi.org/10.1186/1472-6947-10-22>

Bengio, Y. (2009). *Learning Deep Architectures for AI*. Now Publishers Inc. DOI:10.1561/22000000006

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation Learning: A Review and New Perspectives. *IEEE Transactions on Pattern Analysis and Machine In* DOI: 10.1109/TPAMI.2013.50.

Bishop, C.M. (2006). *Pattern Recognition and Machine Learning*. Springer. Available at: <http://users.isr.ist.utl.pt/~wurmd/Livros/school/Bishop%20-%20Pattern%20Recognition%20And%20Machine%20Learning%20-%20Springer%20%202006.pdf>

Bostrom, Nick; Yudkowsky, Eliezer (2011). "[THE ETHICS OF ARTIFICIALINTELLIGENCE](http://www.nickbostrom.com/ethics/artificial-intelligence.pdf)" Available at: <https://www.nickbostrom.com/ethics/artificial-intelligence.pdf>

Bozinovski, S. (1981) *Teaching space: A representation concept for adaptive pattern classification*. COINS Technical Report No. 81-28, Computer and Information Science Department, University of Massachusetts at Amherst, MA. Available at: <https://web.cs.umass.edu/publication/docs/1981/UM-CS-1981-028.pdf>

Bozinovski, Stevo (2014) "Modeling mechanisms of cognition-emotion interaction in artificial neural networks, since 1981." *Procedia Computer Science* p. 255-263. <https://doi.org/10.1016/j.procs.2014.11.111>

Breiman, L. (2001) *Statistical Modeling: The Two Cultures*. Available at: [https://projecteuclid.org/download/pdf\\_1/euclid.ss/1009213726](https://projecteuclid.org/download/pdf_1/euclid.ss/1009213726)

Bresnick, J. (2017). Understanding the Many V's of Healthcare Big Data Analytics Volume, velocity, and variety are all vital for healthcare big data analytics, but there are more V-words to think about, too. *Healthy IT Analytics*. Available at: <https://healthitanalytics.com/news/understanding-the-many-vs-of-healthcare-big-data-analytics>

Bzdok, Danilo; Altman, Naomi; Krzywinski, Martin (2018). "*Statistics versus Machine Learning*". 15(4): 233–234. DOI: [10.1038/nmeth.4642](https://doi.org/10.1038/nmeth.4642)

Carney, E.F. (2020). The impact of chronic kidney disease on global health. <https://doi.org/10.1038/s41581-020-0268-7>

Celesti, A., Fazio, M., Romano, A., & Villari, M. (2016). A hospital cloud-based archival information system for the efficient management of HL7 big data. 2016 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). Opatija, Croatia. DOI: 10.1109/MIPRO.2016.7522177

Charles (2007), Predictive Analytics White Paper (PDF), American Institute for Chartered Property Casualty Underwriters/Insurance Institute of America, p. 1. Available at: <https://www.the-digital-insurer.com/wp-content/uploads/2013/12/78-Predictive-Modeling-White-Paper.pdf>

Chow E, Wong H, Hahn-Goldberg S, Chan CT, Morra D. (2014) *Inpatient and emergent resource use of patients on dialysis at an academic medical center* pp:124-127. DOI:10.1159/00036054.

Clifton, Christopher (2010). "[Encyclopædia Britannica: Definition of Data Mining](https://www.britannica.com/technology/data-mining)" Available at: <https://www.britannica.com/technology/data-mining>

Coates, A., Ng, A. Honglak, L. (2011). An Analysis of Single-Layer Networks in Unsupervised Feature Learning. Proceedings of the *Fourteenth International Conference on Artificial Intelligence and Statistics*, PMLR 15:215-223. Available at: <http://proceedings.mlr.press/v15/coates11a/coates11a.pdf>

Cortes, C., Vapnik, V. (1995). Support-Vector networks. *Machine Learning* 20(3), 273–297. Available at: [http://image.diku.dk/imagecanon/material/cortes\\_vapnik95.pdf](http://image.diku.dk/imagecanon/material/cortes_vapnik95.pdf)

Cox, D.R. (2006). *Principles of Statistical Inference*. Cambridge University Press <https://doi.org/10.1017/CBO9780511813559>

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3), 177–188. DOI: 10.1016/0197-2456(86)90046-2

Deserno, T. M., Haak, D., Brandenburg, V., Classen, C., & Specht, P. (2014). Integrated image data and medical record management for rare disease registries. A general framework and its instantiation to the German Calciphylaxis Registry. *Journal of Digital Imaging*, 27(6), 702–713. DOI: 10.1007/s10278-014-9698-8

Dokas, Paul; Ertoz, Levent; Kumar, Vipin; Lazarevic, Aleksandar; Srivastava, Jaideep; Tan, Pang-Ning (2002). "Data mining for network intrusion detection" pp. 21-30 Available at: [https://pdfs.semanticscholar.org/270a/386a914812c79216e8844409f725eb37213c.pdf?\\_ga=2.142162678.517493180.1638218155-1449435257.1638218155](https://pdfs.semanticscholar.org/270a/386a914812c79216e8844409f725eb37213c.pdf?_ga=2.142162678.517493180.1638218155-1449435257.1638218155)

Edionwe, T. (2017). The fight against racist algorithms. *The Outline*. Available at: <https://theoutline.com/post/1571/the-fight-against-racist-algorithms?zd=1&zi=afcgfb44>

Fayyad, Usama, Piatetsky-Shapiro, Gregory, Smyth, Padhraic (1996). "From Data Mining to Knowledge Discovery in Databases". Available at: <https://www.kdnuggets.com/gspubs/aimag-kdd-overview-1996-Fayyad.pdf>

Feldman, B., Martin, E. M., & Skotnes, T. (2012). Big data in healthcare: Hype and hope. Dr. Bonnie360. Available at: [https://www.ghdonline.org/uploads/big-data-in-healthcare\\_B\\_Kaplan\\_2012.pdf](https://www.ghdonline.org/uploads/big-data-in-healthcare_B_Kaplan_2012.pdf)

Finlay, S. (2014). Predictive Analytics, Data Mining and Big Data. Myths, Misconceptions and Methods (1st ed.) ISBN: 978-1-137-37927-6. Available at: <https://link.springer.com/book/10.1057/9781137379283>

Franc, V., Zien, A., & Schölkopf, B. (2011). Support Vector Machines as Probabilistic Models. Proceedings of the 28th International Conference on Machine Learning, ICML 2011. Bellevue, Washington, USA, June 28 - July 2, 2011. Available at: <https://dblp.uni-trier.de/rec/conf/icml/FrancZS11.html?view=bibtex>

Friedman, J. H. (1998). "Data Mining and Statistics: What's the connection? Computing Science and Statistics. 29 (1), 3–9. Available at: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.81.6162&rep=rep1&type=pdf>

Garets, D., & Davis, M. (2006). *Electronic medical records vs Electronic health records: Yes, there is a difference*. HIMSS Analytics. Available at: <https://docplayer.net/661522-Electronic-medical-records-vs-electronic-health-records-yes-there-is-a-difference-a-himss-analytics-tm-white-paper-by-dave-garets-and-mike-davis.html>

Geisser, S. (1993). Predictive Inference: An Introduction. Chapman & Hall. <https://link.springer.com/10.1007/978-1-4899-4467-2>

Google Maps (2021) *Νομός Θεσσαλονίκης*. Available at : <https://www.google.gr/maps/place/%CE%98%CE%B5%CF%83%CF%83%CE%B1%CE%BB%CE%BF%CE%BD%CE%AF%CE%BA%CE%B7/@40.6578371,23.0767734,121765m/data=!3m1!1e3!4m5!3m4!1s0x14a838f41428e0ed:0x52f2c4450c520454!8m2!3d40.7182715!4d23.0965064!5m2!1e2!1e4?hl=el>

Günemann, S., Kremer, H. & Seidl, T. (2011). An extension of the PMML standard to subspace clustering models. *Proceedings of the 2011 workshop on Predictive markup language modeling*. p.48. <https://doi.org/10.1145/2023598.2023605>

Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, 61(4). <https://doi.org/10.1177/0008125619864925>

Hall, T., & Kumar, N. (2017) Why Machine Learning Models Often Fail to Learn: QuickTake Q&A. Available at: <https://www.bloomberg.com/news/articles/2016-11-10/why-machine-learning-models-often-fail-to-learn-quicktake-q-a>

Han, J., & Kamber (2001). *Data mining: concepts and techniques*. Available at: <http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>

Harel Z, Wald R, McArthur E, et al. 2015 Rehospitalizations and emergency department visits after hospital discharge in patients receiving maintenance hemodialysis. pp :3141-3150. DOI:10.1681/ASN.2014060614.

He, C., Jin, X., Zhao, Z., & Xiang, T. (2010). A cloud computing solution for hospital information system. Paper presented at *the 2010 IEEE International Conference on Intelligent Computing and Intelligent Systems*. Xiamen, China. DOI:10.1109/ICICISYS.2010.5658278

Heart, T., Ben-Assuli, O., & Shabtai, I. (2017). A review of PHR, EMR and EHR integration: A more personalized healthcare and public health policy. *Health Policy and Technology*, 6(1), 20–25. DOI:10.1016/J.HLPT.2016.08.002

Hellenic Statistical Authority (2011) *Demographic statistics*, Available at: <https://www.statistics.gr/en/statistics/-/publication/SAM03/>

Henderson, S. G. & Mason, A. J. In Brandeau, M., Sainfort, F., & Pierskalla, W. P. 2004, Ambulance service planning: simulation and data visualization pp77–102. Available at: <https://people.orie.cornell.edu/shane/pubs/HendersonMasonChap.pdf>

Herland, M., Khoshgoftaar, T. M., & Wald, R. (2014). A review of data mining using big data in health informatics. *Journal of Big Data*, 1(2), 1–35. DOI:10.1186/2196-1115-1-2

Hermon, R., & Williams, P. A.H. (2014). Big data in healthcare: What is it used for? *3rd Australian eHealth Informatics and Security Conference*. 1-3 December, 2014 at Edith Cowan University, Joondalup Campus. Perth, Western Australia. DOI:10.4225/75/57982b9431b48

Hernandez D., & Greenwald, T. (2018). IBM Has a Watson Dilemma. *The Wall Street Journal*. Available at: <https://www.wsj.com/articles/ibm-bet-billions-that-watson-could-improve-cancer-treatment-it-hasnt-worked-1533961147>

Hodge, V. J. & Austin, J. (2004). A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review*. 22(2): 85–126. <https://doi.org/10.1007/s10462-004-4304-y>

Hong, C. J., Kaur, M. N., Farrokhyar, F., & Thoma, A. (2015). Accuracy and completeness of electronic medical records obtained from referring physicians in a Hamilton, Ontario, plastic surgery practice: A prospective feasibility study. DOI: 10.4172/plastic-surgery.1000900

Hong, L., Luo, M., Wang, R., Lu, P., Lu, W., & Lu, L. (2019). Big Data in Health Care: Applications and Challenges. *Data and Information Management* 2(3). DOI:10.2478/dim-2018-0014

IBM Cloud Education. (2020). Neural Networks. Available at: <https://www.ibm.com/cloud/learn/neural-networks>

IBM.(2020):*Bayes Net node* Available at: [https://www.ibm.com/support/producthub/icpdata/docs/content/SSQNUZ\\_latest/wsd/nodes/bayesnet.html](https://www.ibm.com/support/producthub/icpdata/docs/content/SSQNUZ_latest/wsd/nodes/bayesnet.html)

Jordan, M. I., & Bishop, C. M. (2004).Neural Networks. In Allen B. Tucker (ed.). *Computer Science Handbook*, Second Edition (Section VII: Intelligent Systems). Boca Raton, Florida: Chapman & Hall/CRC Press LLC. <https://doi.org/10.1201/9780203494455>

Joshi, K., & Yesha, Y. (2012). Workshop on analytics for big data generated by healthcare and personalized medicine domain. Proceedings of the 2012 *Conference of the Center for Advanced Studies on Collaborative Research*, 267-269. DOI:10.5555/2399776.2399818

Khan, W. A., Khattak, A. M., Hussain, M., Amin, M. B., Afzal, M., Nugent, C., & Lee, S. (2014). An adaptive semantic based mediation system for data interoperability among Health Information Systems. *Journal of Medical Systems*, 38(8), 1-18. DOI: 10.1007/s10916-014-0028-y

Kohavi, R. (1995).[A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection](http://web.cs.iastate.edu/~jtian/cs573/Papers/Kohavi-IJCAI-95.pdf)"(PDF).*International Joint Conference on Artificial Intelligence*.Available at: <http://web.cs.iastate.edu/~jtian/cs573/Papers/Kohavi-IJCAI-95.pdf>

Kovalev, V., & Kalinovsky, A. (2015). Big Medical Data: Image Mining, Retrieval and Analytics. Paper presented at *Big Data and Predictive Analytics*. Minsk, Belarus. Available at: <https://www.researchgate.net/publication/280040575>

Koza J.R., Bennett F.H., Andre D., Keane M.A. (1996) Automated Design of Both the Topology and Sizing of Analog Electrical *Circuits* Using Genetic Programming. In: Gero J.S., Sudweeks F. (eds), *Artificial Intelligence in Design '96* (pp. 151-170). Springer, Dordrecht. [https://doi.org/10.1007/978-94-009-0279-4\\_9](https://doi.org/10.1007/978-94-009-0279-4_9)

Krumholz, H. M. (2014). Big data and new knowledge in medicine: The thinking, training, and tools needed for a learning health system. *Health Affairs*, 33(7), 1163–1170. DOI: 10.1377/hlthaff.2014.0053

Kruse, S. C., Karem, P., Shifflett, K., Vegi, L., Ravi, K., & Brooks, M. (2018). Evaluating barriers to adopting telemedicine worldwide: A systematic review. *Journal of telemedicine and telecare*, 24(1), 4–12. DOI: 10.1177/1357633X16674087

Kshetri, Nir (2014). [Big data's impact on privacy, security and consumer welfare](#). *Telecommunications Policy*.38(11), 1134-1145.  
<http://dx.doi.org/10.1016/j.telpol.2014.10.002>

Kumar, B. (2015). An encyclopedic overview of ‘big data’ analytics. *International Journal of Applied Engineering Research*, 10, (3), 5681-5705. Available at: [https://www.academia.edu/14718806/An\\_Encyclopedic\\_overview\\_of\\_Big\\_Data\\_Analytics](https://www.academia.edu/14718806/An_Encyclopedic_overview_of_Big_Data_Analytics)

Kumar, V., & Grag, M.L. (2018). [International Journal of Computer Applications](#) 182(1), 31-37. Available at: <https://www.researchgate.net/journal/International-Journal-of-Computer-Applications-0975-8887>

Kyoungyoung, J., & Gang-Hoon, K. (2013). Potentiality of Big Data in the Medical Sector: Focus on How to Reshape the Healthcare System. *Healthcare Information Research*, 19(2), 79-85. DOI:10.4258/hir.2013.19.2.79

Lauritzen, S. L., & Spiegelhalter, D. J. (1988). *Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems*.pp 157–224. Available at: <http://www.jstor.org/stable/2345762>

Le Roux, N., Bengio, J., & Fitzgibbon, A. (2012). Improving First and Second-Order Methods by Modeling Uncertainty . In Sra, S., Nowozin, S., Wright, Stephen J. *Optimization for Machine Learning*. MIT Press. Available at: <https://www.cse.iitb.ac.in/~cs709/2015a/notes/readingAssignment/ImprovingOptByModelingUncertainty.pdf>

Link Technologies (2008). *E-DISPATCH*. Available at: <https://www.link-tech.gr/project/e-dispatch/>

Lodish, H. (2008). *Molecular cell biology*. San Francisco, CA: W.H.Freeman and Company. DOI 10.1002/bmb.20373

López-Vallverdú, J.A., Riaño, D., & Bohada, J.A. (2012). Improving medical decision trees by combining relevant health-care criteria. *Expert Systems with Applications*, 39(14), 11782-11791. <https://doi.org/10.1016/j.eswa.2012.04.073>



Loran MJ, McErlean M, Eisele G, et al (2002). The emergency department care of hemodialysis patients. *Clin Nephrol*. pp 57:439-44. DOI: 10.1016/j.ajem.2020.12.006

Lu, H. Plataniotis, K.N., Venetsanopoulos, A.N. (2011). A survey of multilinear subspace learning for tensor data. *Pattern Recognition*, 44(7), 1540-1551. DOI:10.1016/j.patcog.2011.01.004

McGraw-Hill. (2002). McGraw-Hill Concise Dictionary of Modern Medicine. Available at: <https://www.healthcatalyst.com/healthcare-analytics-adoption-model/>

Michie, D., Spiegelhalter, D. J., & Taylor, C. C. (1994). Machine Learning, Neural and Statistical Classification. Ellis Horwood Series in Artificial Intelligence. Available at: <https://ui.adsabs.harvard.edu/abs/1994mlns.book.....M/abstract>

Mitchell, T. (1997). *Machine Learning*. New York: McGraw Hill. ISBN:978-0-07-042807-2

Mohr, D. C., Burns, M. N., Schueller, S. M., Clarke, G., & Klinkman, M. (2013). Behavioral intervention technologies: Evidence review and recommendations for future research in mental health. *General Hospital Psychiatry*, 35(4), 332–338. DOI: 10.1016/j.genhosppsy.2013.03.008

Mohri ,Mehyyar(2012) Rostamizadeh, Afshin; Talwalkar, Ameet (2012).*Foundations of Machine Learning*. USA, Massachusetts: MIT Press. ISBN 9782062018258

NASCIO Research Brief, September (2004) :*Think Before You Dig: Privacy Implications of Data Mining & Aggregation*. Available at: <https://web.archive.org/web/20081217063043/http://www.nascio.org/publications/documents/NASCIO-dataMining.pdf>

Naito, M. (2014). Utilization and application of public health data in descriptive epidemiology. *Journal of Epidemiology*, 24(6), 435–436. DOI: 10.2188/jea.JE20140182

Obenshain, M. K. (2004). Application of data mining techniques to healthcare data. *Infection Control and Hospital Epidemiology*, 25(8), 690–695. DOI: 10.1086/502460

Ohm, P. (2012). Don't Build a Database of Ruin Can be found. *Harvard Business Review* Available at: <https://hbr.org/2012/08/dont-build-a-database-of-ruin>

Page, C.; Sbat, M.; Vazquez, K.; Yalcin, Z.D. (2013) *Analysis of Emergency Medical System across the world*, Worcester Polytechnic Institute. Available at: <http://pacdaoman.gov.om/images/pdf/research/MQFIQP2809.pdf>

Pai, F. Y., & Huang, K. I. (2011). Applying the technology acceptance model to the introduction of healthcare information systems. *Technological Forecasting and Social Change*, 78(4), 650–660. <https://doi.org/10.1016/j.techfore.2010.11.007>

Pavlyshenko B. (2016). "Machine learning, linear and Bayesian models for logistic regression in failure detection problems," 2016 IEEE International Conference on Big Data (Big Data), pp. 2046-2050. DOI: 10.1109/BigData.2016.7840828.

Papaspyrou, E., Setzis, D., Grosomanidis, V., Manikis, D., Boutlis, D., & Ressos, C. (2004). *International EMS systems: Greece. Resuscitation*, 63(3), 255–259. <https://doi.org/10.1016/j.resuscitation.2004.06.009>

Pentland, A., Reid, T., & Heibeck, T. (2013). *Big data and health: Revolutionizing medicine and public health*. Report of the Big Data and Health Working Group 2013. Available at: <http://www.wish-qatar.org/summits/wish-2013/forums-research-chairs/big-data-healthcare/>.

Podgorelec, V., Kokol, P., Stiglic, B., & Rozman, I. (2002) Decision trees: an overview and their use in medicine. *Journal of Medical Systems*, 26, 445-463 <https://doi.org/10.1023/A:1016409317640>

Pontius, Si (2014) *The total operating characteristic to measure diagnostic ability for multiple thresholds* :pp:570-583. DOI:[10.1080/13658816.2013.862623](https://doi.org/10.1080/13658816.2013.862623)

Poole, D., & Macworthh, A. (2010). *Artificial Intelligence. Foundations of Computational Agents*. Cambridge University Press. Available at: <https://www.cambridge.org/gr/academic/subjects/computer-science/artificial-intelligence-and-natural-language-processing/artificial-intelligence-foundations-computational-agents-2nd-edition?format=HB&isbn=9781107195394>

Poulymenopoulou, M., Malamateniou, F., Prentza, A., & Vassilacopoulos, G. (2015). Challenges of evolving PINCLOUD PHR into a PHR-based health analytics system. Paper presented at the Proceedings of the *European, Mediterranean & Middle Eastern Conference on Information Systems EMCIS*. Available at: [https://www.researchgate.net/publication/275208193\\_PINCLOUD\\_II\\_An\\_enhanced\\_cloud-based\\_PHR\\_service\\_with\\_healthcare\\_analytics](https://www.researchgate.net/publication/275208193_PINCLOUD_II_An_enhanced_cloud-based_PHR_service_with_healthcare_analytics)

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*, 2(3). DOI: 10.1186/2047-2501-2-3

Resig, J. (2004). *A Framework for Mining Instant Messaging Services*. Available at: <https://johnresig.com/files/research/SIAMPaper.pdf>

Rome, J. (2005) Data Mining, Dashboards and Data Quality. Available at: [https://web.archive.org/web/20140201170452/http://www.okairp.org/documents/2005%20Fall/F05\\_ROMEDataQualityETC.pdf](https://web.archive.org/web/20140201170452/http://www.okairp.org/documents/2005%20Fall/F05_ROMEDataQualityETC.pdf)

Ronksley PE, Tonelli M, Manns BJ, et al. (2017) Emergency department use among patients with CKD: a population-based analysis. pp :304-314. DOI: 10.2215/CJN.06280616

Russell, S. J. & Norvig, P. (2010) *Artificial Intelligence: A modern Approach* (Third ed.). Prentice Hall. Available at: <https://cs.calvin.edu/courses/cs/344/kvlinden/resources/AIMA-3rd-edition.pdf>

Safavi, S., & Shukur, Z. (2014). Conceptual privacy framework for health information on wearable device. *PLoS One*, 9(12), e114306. <https://doi.org/10.1371/journal.pone.0114306>

Seltzer, W. (2005). *The Promise and Pitfalls of Data Mining: Ethical Issues*. ASA Section on Government Statistics. American Statistical Association. Available at: <https://ww2.amstat.org/committees/ethics/linksdire/Jsm2005Seltzer.pdf>

Sheskin, D. J. (2011). *Handbook of Parametric and Nonparametric Statistical Procedures*. Available at: [https://fmipa.umri.ac.id/wp-content/uploads/2016/03/David\\_J.\\_Sheskin\\_David\\_Sheskin\\_Handbook\\_of\\_ParaBook\\_Fi.org\\_.pdf](https://fmipa.umri.ac.id/wp-content/uploads/2016/03/David_J._Sheskin_David_Sheskin_Handbook_of_ParaBook_Fi.org_.pdf)

Sheta, O. E., & Eldeen, A. N. (2013). The technology of using a data warehouse to support decision-making in health care. *International Journal of Database Management Systems*, 5(3), 75-86. <https://doi.org/10.5121/ijdms.2013.5305>

Sindhu, V, Nivedha, S., & Prakash M, (2020) AN EMPIRICAL SCIENCE RESEARCH ON BIOINFORMATICS IN MACHINE LEARNING. *Journal of Mechanics of Continua and Mathematical Sciences*, 7, 86-94. <https://doi.org/10.26782/jmcms.spl.7/2020.02.00006>

Spaite, D. W., Criss, E. A., Valenzuela, T. D., & Guisto, J. (1995). *Emergency medical service systems research: problems of the past, challenges of the future*. *Annals of emergency medicine*, 26(2), 146–152. [https://doi.org/10.1016/s0196-0644\(95\)70144-3](https://doi.org/10.1016/s0196-0644(95)70144-3)

Srinivasan, U., & Arunasalam, B. (2013). Leveraging Big Data Analytics to Reduce Healthcare Costs. IEEE Computer Society. *IT Professional*, 15(6). DOI:10.1109/MITP.2013.55

Steinbrook, R. (2008). Personally controlled online health data—The next big thing in medical care? *The New England Journal of Medicine*, 358(16), 1653–1656. DOI: 10.1056/NEJMp0801736

Stevenson, C. (2016). Tutorial: Polynomial Regression in Excel. Available at: <https://facultystaff.richmond.edu/~cstevens/301/Excel4.html> Retrieved 24 November 2020.

Stewart, G. B., Higgins, J. P., Schünemann, H., & Meader, N. (2015). The use of Bayesian networks to assess the quality of evidence from research synthesis: 1. *PloS one*, 10(3), e0114497. DOI: 10.1371/journal.pone.0114497

Sun, Z. (2018). *10 Bigs: Big Data and Its Ten Big Characteristics*. BAIS No. 17010, PNG University of Technology. <https://doi.org/10.1145/3291801.3291822>

Swan, M. (2013). The quantified self: Fundamental disruption in big data science and biological discovery. *Big Data*, 1(2), 85–99. DOI: 10.1089/big.2012.0002

Tan, S. S., Gao, G., & Koch, S. (2015). Big Data and Analytics in Healthcare. *Methods of Information in Medicine*, 54(6), 546–547. DOI: 10.3414/ME15-06-1001

Taverner, T., Karpievitch, Y. V., Polpitiya, A. D., Brown, J. N., Dabney, A. R., Anderson, G. A., & Smith, R. D. (2012). DanteR: An extensible R-based tool for quantitative analysis of -omics data. *Bioinformatics* (Oxford, England), 28(18), 2404–2406. DOI: 10.1093/bioinformatics/bts449

Tillmann, A. M. (2015). On the Computational Intractability of Exact and Approximate Dictionary Learning. *IEEE Signal Processing Letters*, 22 (1), 45–49. DOI: 10.1109/LSP.2014.2345761

Tola, K., Abebe, H., Gebremariam, Y., & Jikamo, B. (2017). Improving Completeness of Inpatient Medical Records in Menelik II Referral Hospital, Addis Ababa, Ethiopia. *Advances in Public Health*, 2017, 1–5. <https://doi.org/10.1155/2017/8389414>

Tsumoto, S., Hirano, S., & Iwata, H. (2013). Mining nursing care plan from data extracted from hospital information system. Paper presented at the *2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. Niagara Falls, ON, Canada. <https://doi.org/10.1145/2492517.2500332>

Tzourakis, M. (1996). The Healthcare Industry and Data Quality. Available at: <http://mitiq.mit.edu/iciq/Documents/IQ%20Conference%201996/Papers/TheHealthCareIndustryandDataQuality.pdf>

U.S. Department of Health & Human Services (2014) . Hospital Uses Data Analytics and Predictive Modeling To Identify and Allocate Scarce Resources to High-Risk Patients, Leading to Fewer Readmissions. Available at: <https://innovations.ahrq.gov/profiles/hospital-uses-data-analytics-and-predictive-modeling-identify-and-allocate-scarce-resources>

Usami, Y., Cho, H. C., Okazaki, N., & Tsujii, J. I. (2011). Automatic acquisition of huge training data for bio-medical named entity recognition. *Proceedings of BioNLP 2011 Workshop* 5, 65-73. DOI:10.5555/2002902.2002912

van Otterlo, M. & Wiering, M. (2012). Reinforcement learning and markov decision processes. Reinforcement Learning. *Adaptation, Learning, and Optimization*, 12. pp. 3–42. DOI: 10.1007/978-3-642-27645-3\_1

Vesna, V. (2000). The Visible Human Project: Informatic bodies and posthuman medicine. *AI & Society*, 14(2), 262–263. DOI:10.1007/BF01205456

Wang, L., & Alexander, C. A. (2013). Applications of automated identification technology in EHR/EMR. *International Journal of Public Health Science*, 2(3), 109–122. DOI:10.11591/IJPHS.V2I3.3300

Windridge, D., & Bober, M. (2014). A kernel-based framework for medical big-data analytics. In A. Holzinger & I. Jursica (Eds.), *Interactive knowledge discovery and data mining in biomedical informatics* (pp. 197-208). Berlin, Germany: Springer-Verlag. DOI: 10.1007/978-3-662-43968-5\_11

World Health Organization (WHO) (2008) *Emergency Medical Services Systems in the European Union pp 22* Available at:

[https://www.euro.who.int/\\_data/assets/pdf\\_file/0016/114406/E92038.pdf](https://www.euro.who.int/_data/assets/pdf_file/0016/114406/E92038.pdf)

Wu, Luo, J., Wu, M., Gopukumar, D., & Zhao, Y. (2016). Big data application in biomedical research and health care: A literature review. *Biomedical Informatics Insights*, 8, 1–10. DOI: 10.4137/BII.S31559

Yuen-Reed, G., & Mojsilović, A. (2016). The role of big data and analytics in health payer transformation to consumer-centricity. In C. Weaver, M. Ball, G. Kim & J. Kiel (Eds.), *Healthcare information management systems* (pp. 399–420). Switzerland: Springer. DOI:10.1007/978-3-319-20765-0\_23

Zimek, A. & Schubert, E. (2017). Outlier Detection. In *Encyclopedia of Database Systems*. Springer New York. DOI: [https://doi.org/10.1007/978-1-4899-7993-3\\_80719-1](https://doi.org/10.1007/978-1-4899-7993-3_80719-1)

Γιαννόπουλος Α, Ασλανίδης Θ , Τσικρικώνης Γ , Βαρελά Α , Παυλίδης Ε , Γιατσιάτσος Σ , Στολτίδης Π , Αραμπατζή Α , Αλεξανδρίδης Κ , Τσιαουσίδης Κ , Προδρόμου Β , Τανίδης Π , Βιάρου Μ , Βορνίτης Η , Τατατζίδης Χ, Σταμάτης Ε, Τσιούπα Α (2018): *Διαχείριση προνοσοκομειακών περιστατικών από διασώστες. Πρόδρομη ανάλυση 1010 περιστατικών από τη Βόρεια Ελλάδα*. DOI:The Greek E-Journal of Perioperative Medicine2018;17(d):50-66

ΕΟΔΥ (2020) *Ημερήσια έκθεση επιδημιολογικής επιτήρησης λοίμωξης από το νέο κορωνοϊό (COVID-19) Δεδομένα έως 28 Οκτωβρίου 2020*. Available at: <https://eody.gov.gr/wp-content/uploads/2020/10/covid-gr-daily-report-20201028.pdf>

Εθνικό Κέντρο Άμεσης Βοήθειας (2021). *Χωροταξική Κατανομή* Ε.Κ.Α.Β.. Available at: <https://www.ekab.gr/chorotaxiki-katanomi/>

Εθνικό Κέντρο Άμεσης Βοήθειας (2021). *Μέσα Διακομιδής* Ε.Κ.Α.Β.. Available at: <https://www.ekab.gr/mesa-diakomidis/>

Εθνικό Κέντρο Άμεσης Βοήθειας (2021) *ΟΡΓΑΝΙΣΜΟΣ Ε.Κ.Α.Β. (ΑΠΟΣΤΟΛΗ & ΣΚΟΠΟΙ)*. Available at: <https://www.ekab.gr/files/2019/04/%CE%A3%CE%A7%CE%95%CE%94%CE%99%CE%9F-%CE%A1%CE%93%CE%91%CE%9D%CE%99%CE%A3%CE%9C%CE%9F%CE%A5-%CE%95%CE%9A%CE%91%CE%92.pdf>

Νόμος 1579/1985 (1985) *Εφαρμογή και ανάπτυξη του Εθνικού Κέντρου Άμεσης Βοήθειας (ΦΕΚ 217/τ. Α'/23-12- 1985)*. Available at: <https://www.e-nomothesia.gr/kat-ygeia/n-1579-1985.html>

Νόμος 2071, (1992), *Εκσυγχρονισμός και Οργάνωση του συστήματος υγείας*, ΦΕΚ 123, Τεύχος Α' κεφαλαίο 3. Available at: <https://www.e-nomothesia.gr/kat-ygeia/n-2071-1992.html>