



ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΔΙΟΙΚΗΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ
ΤΜΗΜΑ ΛΟΓΙΣΤΙΚΗΣ ΚΑΙ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΛΟΓΙΣΤΙΚΗΣ ΚΑΙ
ΕΛΕΓΚΤΙΚΗ

Διπλωματική Εργασία

Η ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ ΣΤΗΝ ΥΠΟΒΟΗΘΗΣΗ ΤΗΣ ΕΛΕΓΚΤΙΚΗΣ:
ΜΙΑ ΣΥΓΚΡΙΤΙΚΗ ΜΕΛΕΤΗ ΣΤΟΝ ΕΝΤΟΠΙΣΜΟ ΑΠΑΤΗΣ ΣΤΙΣ ΟΙΚΟΝΟΜΙΚΕΣ ΚΑΤΑΣΤΑΣΕΙΣ ΜΕ
ΤΗΝ ΧΡΗΣΗ ΤΕΧΝΙΚΩΝ ΕΞΟΡΥΞΗΣ ΔΕΔΟΜΕΝΩΝ

ΤΟΥ

ΓΙΑΝΚΟΥΛΙΔΗ ΑΝΔΡΕΑ

Επιβλέπων Καθηγητής: Πρωτόγερος Νικόλαος

Υποβλήθηκε ως απαιτούμενο για την απόκτηση του Μεταπτυχιακού Διπλώματος στην
Εφαρμοσμένη Λογιστική και Ελεγκτική

Οκτώβριος 2020

*Αφιερώνεται στη Μαριάννα
που με ανέχτηκε
και στους γονείς μου...*

Ευχαριστίες

Η παρούσα εργασία αποτελεί διπλωματική εργασία στα πλαίσια του μεταπτυχιακού προγράμματος «Εφαρμοσμένη Λογιστική και Ελεγκτική» του τμήματος Λογιστικής και Χρηματοοικονομικής.

Πριν την παρουσίαση των αποτελεσμάτων της παρούσας διπλωματικής εργασίας, θα ήθελα να ευχαριστήσω αρχικά τον επιβλέποντα καθηγητή της διπλωματικής εργασίας, καθηγητή κύριο Πρωτόγερο Νικόλαο για την πολύτιμη καθοδήγηση του και τη συμβολή του στην επιτυχή ολοκλήρωση της εργασίας και έπειτα όλους τους καθηγητές του Προγράμματος για τις γνώσεις που μου προσέφεραν.

ΠΕΡΙΛΗΨΗ

Στην παρούσα εργασία γίνεται μια σύγκριση προηγούμενων ερευνών στον τομέα του εντοπισμού απάτης στις οικονομικές καταστάσεις με την χρήση τεχνικών εξόρυξης δεδομένων. Τα ερευνητικά ερωτήματα της εργασίας είναι ποιες τεχνικές εξόρυξης δεδομένων υπάρχουν σήμερα για τον έλεγχο των οικονομικών καταστάσεων, πόσο ακριβείς είναι στον εντοπισμό παραποιημένων οικονομικών καταστάσεων και πια η συνεισφορά τους στην ελεγκτική διαδικασία. Γίνεται αναφορά στην ηθική κουλτούρα και συμπεριφορά των εταιρειών για την διάπραξη απάτης. Παρουσιάζονται οι ορισμοί των οικονομικών καταστάσεων, του τριγώνου απάτης και της εξόρυξης δεδομένων. Τα αποτελέσματα των ερευνών αναλύονται με την στατιστική μέθοδο του ελέγχου υποθέσεων όπου φαίνεται η ακρίβεια των τεχνικών εξόρυξης δεδομένων στον εντοπισμό απάτης στις οικονομικές καταστάσεις που εφαρμόστηκαν.

Από την παρούσα συγκριτική μελέτη βλέπουμε ότι η τεχνική εξόρυξης δεδομένων με την υψηλότερη συνολική ακρίβεια είναι το πιθανοτικό νευρωνικό δίκτυο, βέβαια η συγκεκριμένη τεχνική χρησιμοποιείται σε μία μόνο έρευνα. Αντίστοιχα οι τεχνικές που χρησιμοποιούνται στις περισσότερες έρευνες είναι τα τεχνητά νευρωνικά δίκτυα και τα δέντρα απόφασης, δυο πολύ διαδεδομένοι αλγόριθμοι με εφαρμογή σε πολλά επιστημονικά πεδία. Τα αποτελέσματα της παρούσας εργασίας δείχνουν ότι η εξόρυξη δεδομένων, με τις τεχνικές της, μπορεί να βοηθήσει την ελεγκτική διαδικασία στον εντοπισμό παραποιημένων οικονομικών καταστάσεων. Τα αποτελέσματα των ερευνών πάνω σε αυτό το πεδίο, τα τελευταία χρόνια είναι ιδιαίτερος ελπιδοφόρα, ώστε οι ελεγκτές να μπορούν να εμπιστευθούν την εξόρυξη δεδομένων για τον εντοπισμό απάτης στις οικονομικές καταστάσεις εταιρειών.

Abstract

This paper compares previous research in the field of fraud detection in financial statements using data mining techniques. The research questions of the work are what data mining techniques exist today for the audit of financial statements, how accurate they are in detecting falsified financial statements and their contribution to the audit process. Reference is made to the ethical culture and behavior of companies for committing fraud. Are presented the definitions of financial statements, fraud triangle and data mining. The results of the investigations are analyzed with the statistical method of hypothesis control which shows the accuracy of data mining techniques tied to the detection of fraud in the financial statements applied.

From the present comparative study we see that the data mining technique with the highest overall accuracy is the probable neural network, of course this technique is used in a single research. The techniques used in most research are artificial neural networks and decision trees, two very common algorithms applied in many scientific fields. The results of the present work show that data mining, with its techniques, can help the audit process in the detection of falsified financial statements. The results of research in this area in recent years are particularly promising, so that auditors can trust the data mining to detect fraud in the financial statements of companies.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

1.1	ΔΙΑΤΥΠΩΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ	σελ. 1
1.2	ΣΚΟΠΟΣ ΕΡΓΑΣΙΑΣ	σελ. 2
1.3	ΕΡΕΥΝΗΤΙΚΑ ΕΡΩΤΗΜΑΤΑ	σελ. 3
1.4	ΑΝΑΓΚΑΙΟΤΗΤΑ ΕΡΕΥΝΑΣ	σελ. 4
1.5	ΔΙΑΡΘΡΩΣΗ ΕΡΓΑΣΙΑΣ.....	σελ. 5

ΚΕΦΑΛΑΙΟ 2: ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

2.1	ΗΘΙΚΗ ΚΟΥΛΤΟΥΡΑ ΚΑΙ ΗΘΙΚΟ ΚΛΙΜΑ.....	σελ. 6
2.2	ΟΙΚΟΝΟΜΙΚΕΣ ΚΑΤΑΣΤΑΣΕΙΣ ΚΑΙ ΚΙΝΔΥΝΟΙ ΑΠΑΤΗΣ.....	σελ. 8
	2.2.1 Οικονομικές καταστάσεις.....	σελ. 8
	2.2.2 Απάτη Σφάλμα.....	σελ. 8
	2.2.3 Το τρίγωνο της απάτης.....	σελ. 9
2.3	ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ.....	σελ. 10
	2.3.1 Η εξόρυξη δεδομένων στον εντοπισμό απάτης στις οικονομικές καταστάσεις.....	σελ. 10

ΚΕΦΑΛΑΙΟ 3: ΜΕΘΟΔΟΛΟΓΙΑ

3.1	ΜΕΘΟΔΟΛΟΓΙΑ	σελ. 14
3.2	ΣΥΓΚΕΝΤΡΩΣΗ ΔΕΔΟΜΕΝΩΝ	σελ. 20

ΚΕΦΑΛΑΙΟ 4: ΠΑΡΟΥΣΙΑΣΗ ΕΡΓΟΥ

4.1	ΔΗΜΙΟΥΡΓΙΑ ΣΥΓΚΕΝΤΡΩΤΙΚΟΥ ΠΙΝΑΚΑ	σελ. 21
4.2	ΤΑΞΙΝΟΜΗΣΗ ΚΑΤΑ ΑΚΡΙΒΕΙΑ ΤΕΧΝΙΚΗΣ	σελ. 29

ΚΕΦΑΛΑΙΟ 5: ΣΥΜΠΕΡΑΣΜΑΤΑ

σελ. 35

ΕΠΙΛΟΓΟΣ

σελ. 37

ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

σελ. 38

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1	σελ.13
Πίνακας 2	σελ.14
Πίνακας 3	σελ.15
Πίνακας 4	σελ.16
Πίνακας 5	σελ.21
Πίνακας 6	σελ.27
Πίνακας 7.1	σελ.28
Πίνακας 7.2.....	σελ.29
Πίνακας 7.3.....	σελ.29
Πίνακας 7.4.....	σελ.29
Πίνακας 7.5.....	σελ.29
Πίνακας 7.6.....	σελ.30
Πίνακας 7.7.....	σελ.30
Πίνακας 7.8.....	σελ.30
Πίνακας 7.9.....	σελ.30
Πίνακας 7.10.....	σελ.30
Πίνακας 7.11.....	σελ.31
Πίνακας 8.....	σελ.31
Πίνακας 9.....	σελ.32

ΠΙΝΑΚΕΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΕΡΕΥΝΩΝ

1.	σελ. 22
2.	σελ. 23
3.	σελ. 23
4.	σελ. 24
5.	σελ. 24
6.	σελ. 25
7.	σελ. 26
8.	σελ. 26

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

1.1 ΔΙΑΤΥΠΩΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Στις αρχές της προηγούμενης δεκαετίας εντοπίζεται ένα από τα μεγαλύτερα οικονομικά σκάνδαλα στην αμερικανική οικονομία με αλυσιδωτές επιπτώσεις παγκοσμίως (Munawer et al., 2012). Συγκεκριμένα, ο κολοσσός εταιρείας ενέργειας Enron παρουσίαζε έσοδα άνω των 100 δις. Δολαρίων ετησίως. Ωστόσο, ανακαλύφθηκαν σύντομα τα εγχειρήματα λογιστικής απάτης και οι παραποιημένες οικονομικές της καταστάσεις. Από τις αρχές του 2000 και έπειτα, η διεθνής βιβλιογραφία έχει αναδείξει αρκετές περιπτώσεις απάτης σε οικονομικές καταστάσεις εταιρειών εισηγμένων ή μη (Buchanan & Yang, 2005). Πιο αναλυτικά, η ιταλική εταιρεία τροφίμων Parmalat Finanziaria SpA, μία από τις κορυφαίες στον τομέα της παγκοσμίως, το 2003 ανακαλύφθηκε με έλλειμα στα ταμιά της ύψους 14 δις. Ευρώ. Αρχικά, η εταιρεία υπέβαλε αίτηση προστασίας από πτώχευση στα τέλη του έτους χωρίς ανταπόκριση και κατέρρευσε υπό την ονομασία «Europe's Enron».

Ακολούθως, η κατάρρευση της Lehman Brothers που αποτέλεσε μια από τις μεγαλύτερες επενδυτικές τράπεζες κατέρρευσε το 2008 πυροδοτώντας την παγκόσμια χρηματοπιστωτική κρίση (Beccar-Varela et al., 2017). Οι συνέπειες της εν λόγω κατάρρευσης έπληξαν ακόμα και την Ελλάδα σε πολύ σημαντικό βαθμό για την επόμενη δεκαετία. Ο «φυσικός σεισμός» της πτώσης του κολοσσού αυτού προκλήθηκε εξαιτίας απάτης της τάξεως των 691 δις δολαρίων και αναδύθηκε ως η μεγαλύτερη χρεοκοπία των ΗΠΑ στην ιστορία. Οι παραπάνω περιπτώσεις αναφέρονται ενδεικτικά καθώς σε μεταγενέστερο στάδιο εντοπίζονται και άλλα εξίσου σημαντικά οικονομικά σκάνδαλα

Ειδικότερα, σύμφωνα με ισχυρισμούς λογιστικών αρχών, η περίπτωση της Enron, αποτέλεσε μία «τέλεια καταιγίδα» και σύμπτωση απρόβλεπτων και σπάνιων συνθηκών που επέφερε μείωση της εμπιστοσύνης και αξιοπιστίας των επενδυτών στις κεφαλαιαγορές (Reinstein & McMillan, 2004). Η καταστροφή της Enron υπογράμμισε ως επιτακτική την ανάγκη για υιοθέτηση μίας προσέγγισης που αναστοχάζεται πάνω στο επαγγελματικό ήθος του λογιστικού επαγγέλματος. Η κριτική ανάλυση των οικονομικών καταστάσεων καθώς και ο συστηματικός και ενδεδειγμένος έλεγχος προκύπτει ως βασική υποχρέωση των ελεγκτικών εταιρειών. Εντούτοις, προκειμένου

οι διαδικασίες να πραγματοποιούνται χωρίς σφάλματα και βάσει μίας λογικής πρόληψης λαθών και λειτουργικού σχεδιασμού, απαιτείται περαιτέρω επιστημονική διερεύνηση.

Πλέον, λόγω και της υψηλής τεχνολογικής ανάπτυξης των τελευταίων δεκαετιών, υπάρχουν πολλά και διάφορα εργαλεία για την υποστήριξη των διαδικασιών του ελέγχου. Οι ελεγκτικές εταιρείες καλούνται να υποστηρίξουν την δουλειά τους με κάποια από αυτά. Η βιβλιογραφία έχει εμπλουτιστεί με διάφορους μηχανισμούς ελέγχου για τον εντοπισμό παραπονημένων οικονομικών καταστάσεων. Οι μηχανισμοί αυτοί επαφίονται στους κλάδους της στατιστικής, της επιχειρηματικής ευφυΐας ή και των βάσεων δεδομένων.

Η εξόρυξη δεδομένων είναι ένας κλάδος που δίνει τα εργαλεία στους ελεγκτές να εντοπίζουν απάτες στα οικονομικά μεγέθη των εταιρειών. Την τελευταία εικοσαετία έχει δημοσιευθεί πληθώρα εργασιών όπου παρουσιάζουν διάφορες τεχνικές εξόρυξης δεδομένων στον εντοπισμό απάτης στις οικονομικές καταστάσεις με τα αποτελέσματα να είναι αρκετά ελπιδοφόρα.

1.2 ΣΚΟΠΟΣ ΕΡΓΑΣΙΑΣ

Η παρούσα εργασία έχει ως σκοπό τη διερεύνηση της ακρίβειας εφαρμογής ορισμένων τεχνικών εξόρυξης δεδομένων οι οποίες δύνανται να εντοπίσουν σφάλματα απάτης σε οικονομικές καταστάσεις εταιρειών. Πιο συγκεκριμένα, θα παρατεθούν τα αποτελέσματα ερευνών που διενεργήθηκαν σε διάφορες χώρες παγκοσμίως και θα συγκριθούν τα αποτελέσματά τους με βάση την ακρίβεια, όλων των τεχνικών, στον εντοπισμό απάτης των οικονομικών καταστάσεων. Από την σύγκριση των τεχνικών εξόρυξης δεδομένων, αποσκοπούμε στη δημιουργία ενός πίνακα που θα παρουσιάζει την ακρίβεια της κάθε τεχνικής στον εντοπισμό παραπονημένων οικονομικών καταστάσεων, με σκοπό να μπορούν οι επαγγελματίες ελεγκτές να πληροφορηθούν και να ενημερωθούν ούτως ώστε να μπορούν να επιλέξουν, με ασφάλεια, την τεχνική που θα τους βοηθήσει στην διαδικασία του ελέγχου των οικονομικών καταστάσεων.

1.3 ΕΡΕΥΝΗΤΙΚΑ ΕΡΩΤΗΜΑΤΑ

Λαμβάνοντας υπόψη τη διατύπωση του προβλήματος της εργασίας αλλά και τον σκοπό όπως τον περιγράψαμε παραπάνω, προκύπτουν τα εξής ερευνητικά ερωτήματα:

- ❖ Ποιες τεχνικές εξόρυξης δεδομένων υπάρχουν σήμερα για την υποβοήθηση της ελεγκτικής;
- ❖ Πόσο ακριβείς και αποτελεσματικές είναι οι τεχνικές εξόρυξης δεδομένων που διερευνούμε ως προς τον εντοπισμό της απάτης;

Ακολούθως, οι ερευνητικές υποθέσεις που προκύπτουν βάσει των ερευνητικών ερωτημάτων διατυπώνονται ως εξής:

- ❖ Οι τεχνικές εξόρυξης δεδομένων που χρησιμοποιούνται, σε μεγαλύτερο βαθμό, για την υποβοήθηση της ελεγκτικής είναι, βάσει της βιβλιογραφίας, τα δέντρα απόφασης (DT), τα νευρωνικά δίκτυα (NN), οι μηχανές διανυσμάτων υποστήριξης (SVM), η λογιστική παλινδρόμηση (LR) και τα δίκτυα Bayes (BN).
- ❖ Η ακρίβεια που έχουν παρουσιάσει στο σύνολο των ερευνών είναι ιδιαίτερα υψηλή. Οπότε είναι ικανή η εξόρυξη δεδομένων να δώσει αποτελεσματικές λύσεις στην ελεγκτική.

1.4 ΑΝΑΓΚΑΙΟΤΗΤΑ ΕΡΕΥΝΑΣ

Η ραγδαία τεχνολογική εξέλιξη τις τελευταίες δεκαετίες έχει διευρύνει τους ορίζοντες σε πολλούς κλάδους της επιστήμης, των επιχειρήσεων και όχι μόνο. Αναμφίβολα, η ελεγκτική των επιχειρήσεων συγκεκριμένα δύναται να επωφεληθεί ιδιαίτερα από την παραπάνω τεχνολογική ανάπτυξη καθώς οι ελεγκτές που ελέγχουν τα οικονομικά δεδομένα των εταιρειών χρησιμοποιούν πλέον καινοτόμα ψηφιακά εργαλεία (Tan, Steinbach, & Kumar, 2015). Τα δεδομένα αυτά είναι αποθηκευμένα σε βάσεις δεδομένων και υπόκεινται σε διαχείριση μέσω πληροφοριακών συστημάτων. Ένα πλεονέκτημα των πληροφοριακών συστημάτων είναι η άμεση και έγκυρη παροχή πληροφορίας από την βάση δεδομένων της εταιρείας οπότε και οι ελεγκτές συλλέγουν αξιόπιστη πληροφορία προς έλεγχο.

Παρ' όλα αυτά, τα τελευταία χρόνια, έχουν εμφανιστεί περιπτώσεις εξαπάτησης των ελεγκτών από επιχειρήσεις με απαρχή τις δύο μεγαλύτερες απάτες που προέκυψαν στις

εταιρείες Enron και WorldCom. Ως αποτέλεσμα είναι να έχει ανοίξει η συζήτηση των διαδικασιών ελέγχου που ακολουθούνται και των δικλίδων ασφαλείας στην ελεγκτική διαδικασία (Maka et al., 2020).

Ένας κλάδος που μπορεί να δώσει καινοτόμες λύσεις σε αυτό το πρόβλημα είναι η εξόρυξη δεδομένων. Ο κλάδος αυτός επαφίεται στις βάσεις δεδομένων, της στατιστικής και της επιχειρηματικής ευφυίας μεταξύ άλλων (Santhosh & Mohanapriya, 2020). Έτσι, δίνεται η δυνατότητα στις ελεγκτικές εταιρείες να εκμεταλλευτούν την πρόοδο της τεχνολογίας προκειμένου να ελαχιστοποιήσουν και να περιορίσουν λάθη συστήματος που προκλήθηκαν από λειτουργικές δυσλειτουργίες είτε να ανακαλυφθούν σφάλματα από σκόπιμες παραποιήσεις.

Ωστόσο, το πλήθος τεχνικών εξόρυξης δεδομένων που προέκυψαν δυσκολεύουν τους επαγγελματίες λογιστές να εντοπίσουν την τεχνική που ταιριάζει περισσότερο στην εκάστοτε συνθήκη εργασίας. Υπάρχουν δηλαδή τεχνικές που αξιοποιούν διαφορετικούς αλγόριθμους με διαφορετικό τρόπο διαχείρισης των μεταβλητών, γεγονός που προβληματίζει ως προς την αποτελεσματικότητα και ακρίβεια της κάθε μίας. Επομένως, διαπιστώθηκε η αναγκαιότητα εφαρμογής αξιόπιστων τεχνικών εξόρυξης δεδομένων προς αποφυγή σφαλμάτων που θα μπορούσαν να επιφέρουν ζητήματα ηθικής ακεραιότητας της εταιρείας (Buchanan & Yang, 2005).

Για το λόγο αυτό, στην παρούσα έρευνα επιδιώκεται η ανάλυση της ποιότητας των αποτελεσμάτων βάσει δεικτών ακρίβειας ορισμένων τεχνικών εξόρυξης δεδομένων. Πιο αναλυτικά, η εργασία συγκεντρώνει ένα σημαντικό, αντιπροσωπευτικό δείγμα ερευνών που ασχολούνται με τον εντοπισμό απάτης στις οικονομικές καταστάσεις με συγκεκριμένες τεχνικές εξόρυξης δεδομένων, την τελευταία δεκαετία. Στόχος είναι η πληροφόρηση των ελεγκτών ώστε να μπορούν να αποφασίσουν αν θα χρησιμοποιήσουν κάποια τεχνική εξόρυξης δεδομένων στην υποβοήθηση της εργασίας τους. Τέλος, η εργασία αυτή παρουσιάζει ένα αντιπροσωπευτικό δείγμα της σύγχρονης βιβλιογραφίας ώστε οι ερευνητές να έχουν μια συγκεντρωμένη πληροφόρηση που θα βοηθήσει σε μελλοντικές έρευνες.

1.5 ΔΙΑΡΘΡΩΣΗ ΕΡΓΑΣΙΑΣ

Η παρούσα εργασία αποτελείται από 5 κεφάλαια μέσα από τα οποία παρουσιάζεται η υποβοήθηση της ελεγκτικής διαδικασίας μέσω τεχνικών εξόρυξης δεδομένων.

Στο πρώτο κεφάλαιο γίνεται η εισαγωγή στην εργασία. Πιο αναλυτικά, παρουσιάζεται ο σκοπός της εργασίας, τα ερευνητικά ερωτήματα που τέθηκαν κατά τον σχεδιασμό. Αναλύεται η δομή και καταγράφεται η αναμενόμενη συνεισφορά της.

Στο δεύτερο κεφάλαιο καταγράφεται και παρουσιάζεται η βιβλιογραφική ανασκόπηση. Αναλύονται οι ορισμοί των οικονομικών καταστάσεων, του σφάλματος στις οικονομικές καταστάσεις και η εξόρυξη δεδομένων. Παρουσιάζονται προηγούμενες έρευνες όπου εφαρμόζονται διάφορες τεχνικές εξόρυξης δεδομένων με σκοπό τον εντοπισμό παραποιημένων οικονομικών καταστάσεων.

Στο τρίτο κεφάλαιο παρουσιάζεται η μεθοδολογία των ερευνών που αναλύονται στην παρούσα εργασία. Επίσης αναλύεται η συγκέντρωση των δεδομένων από τα αποτελέσματα των ερευνών που συμμετέχουν στην παρούσα εργασία, σε μορφή πινάκων. Γίνεται αναφορά στον όρο ακρίβεια των τεχνικών αυτών και γίνεται ανάλυση των ερευνών.

Στο τέταρτο κεφάλαιο γίνεται η συγκέντρωση των αποτελεσμάτων σε έναν συγκριτικό πίνακα. Ο πίνακας αυτός είναι ταξινομημένος με βάση την ακρίβεια της κάθε τεχνικής. Στο πέμπτο κεφάλαιο παρουσιάζονται τα αποτελέσματα της συγκριτικής ανάλυσης και του συγκριτικού πίνακα. Γίνεται αναφορά σε προτάσεις για μελλοντική έρευνα, πάνω στο κομμάτι της εξόρυξης δεδομένων και την υποβοήθηση αυτής στον τομέα της ελεγκτικής.

ΚΕΦΑΛΑΙΟ 2: ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

2.1 ΗΘΙΚΗ ΚΟΥΛΤΟΥΡΑ ΚΑΙ ΗΘΙΚΟ ΚΛΙΜΑ

Προκειμένου να αιτιολογηθεί μία επαγγελματική συμπεριφορά, ένας κώδικας κουλτούρας και η συλλογική οργανωτική στάση μίας επιχείρησης χρησιμοποιούνται συχνά οι έννοιες του πολιτισμού και του κλίματος (Suh & Shim, 2020). Πιο αναλυτικά, εστιάζοντας στον πολιτισμό γίνεται λόγος για τις πεποιθήσεις των εργαζομένων αναφορικά με τον τρόπο λειτουργίας του οργανισμού (Agarwal & Malloy, 1999). Από την άλλη πλευρά, με την έννοια «κλίμα» νοείται ως η κοινή αντίληψη περί των συνθηκών λειτουργίας της εταιρείας το οποίο έχει προκύψει ως αποτέλεσμα των ξεχωριστών προσωπικοτήτων που εργάζονται στην επιχείρηση και άρα νοείται ως ένα ενιαίο κράμα διαφορετικών νοοτροπιών. Επομένως, η ηθική κουλτούρα αποσκοπεί πιο συγκεκριμένα στο να επηρεάσει την συμπεριφορά ενός εργαζόμενου ενώ το κλίμα εξετάζει, αναλύει και τελικά συγκεντρώνει τις διαφορετικές συμπεριφορές σε ένα ενιαίο κλίμα που υιοθετείται με επιτυχία ή μη από το σύνολο των εργαζομένων. Από το 1998 κιόλας ο Trevino είχε επιχειρήσει τη διερεύνηση της ηθικής κουλτούρας την οποία είχε περιγράψει ως έννοια – υποκατηγορία της οργανωτικής κουλτούρας. Σαφέστερα, όπως διατύπωσε η οργανωτική κουλτούρα αφορά μία σύνθετη και περίπλοκη αλληλεπίδραση διαφορετικών συστημάτων ελέγχου με επίσημη ή μη επίσημη ταυτότητα που επιδιώκουν την πρόκληση μίας ηθικής ή μίας ανήθικης ενέργειας.

Σε μία περισσότερο αισιόδοξη λογική της ηθικής απάτης ως προς τον συλλογικό δόλο, οι Button & Gee (2013), περιέγραψαν ένα μοντέλο απάτης, ανθεκτικότητας και πολιτισμού το οποίο δημιουργείται από την προσπάθεια μεμονωμένων εργαζομένων που δρουν σε μία δομή οργάνωσης. Σε μία τέτοια περίπτωση, η ίδια η επιχείρηση ως σύστημα εφαρμόζει ένα σύστημα ασφαλείας με σκοπό τον εντοπισμό του ατόμου που προσπαθεί να διαστρεβλώσει την ηθική ακεραιότητα της εταιρείας. Έτσι, αντιλαμβανόμαστε ότι κάθε επιχείρηση πλέον έχει στη διάθεσή της τρόπους για να εντοπίσει άτομα που εμπλέκονται σε απάτη ή όχι. Αναλυτικότερα, κατά τον Schwartz (2013), έχουν αναπτυχθεί τρεις βασικοί παράγοντες η ύπαρξη των οποίων δύναται να επηρεάσει τη δυνατότητα εφαρμογής απάτης από τους εργαζομένους. Οι παράνομες δραστηριότητες θα μπορούσαν να ελαχιστοποιηθούν ή ακόμα καλύτερα να εξαφανιστούν κατά τη διατήρηση μίας ενιαίας και αμετάβλητης ηθικής εταιρικής

κουλτούρας με ενσωματωμένο ένα σύστημα ηθικών αξιών σεβαστό από όλους που να περιλαμβάνει έννοιες όπως η «ακεραιότητα». Επιπλέον, ο δεύτερος παράγοντας θα μπορούσε να αφορά την πιστή εφαρμογή ενός επίσημου προγράμματος δεοντολογίας κατά το οποίο οι εργαζόμενοι θα επιμορφώνονται στον τομέα της «ηθικής» και θα συμμετέχουν σε μία δομημένη εκπαιδευτική διαδικασία για αποφυγή λαθών που θα μπορούσαν να προκύψουν χωρίς δόλο. Τελευταίος αλλά εξίσου σημαντικός παράγοντας είναι ο αναστοχασμός της ηθικής ηγεσίας, του διοικητικού συμβουλίου, των υψηλών στελεχών και των διευθυντών σχετικά με τη διευθέτηση καταστάσεων απάτης και την πρόληψη νέων.

Στόχος της ηθικής ηγεσίας αποτελεί η λειτουργία της ως πηγή έμπνευσης μέσω λέξεων αλλά και παραδειγμάτων έντιμης διαχείρισης. Με τον παραπάνω τρόπο είναι εφικτή η καλλιέργεια μίας ηθικής ευαισθητοποίησης που θα μπορούσε να ενεργοποιήσει ακόμα και το θάρρος της γνώμης ορισμένων ατόμων για την παραδοχή παρατυπιών. Βεβαίως, δεν θα μπορούσαμε να παραβλέψουμε τις βιβλιογραφικές αναφορές περί ισχυρότερων οικονομικά απατών από διευθυντικά στελέχη οι οποίοι παράκαμψαν προγραμματισμένους ελέγχους και απενεργοποίησαν συστήματα αναφοράς σφαλμάτων με σκοπό να καλύψουν ανήθικες συμπεριφορές προσωπικές τους ή συνεργατών τους (Tirgos, 2002). Επομένως, πρωτίστως θα πρέπει μία εταιρεία να εξασφαλίσει την ακεραιότητα και την ηθική των ανώτερων στελεχών της καθώς και να ελέγχει ανά διαστήματα την αξιοπιστία της και σε δεύτερο βαθμό να εφαρμόζει τεχνικές ελέγχου αξιοπιστίας των εργαζομένων και των συστημάτων της.

Προχωρώντας στην ακεραιότητα στο χώρο εργασίας βρισκόμαστε μπροστά στην ανάγκη διερεύνησης της έννοιας των ηθικών αξιών. Συγκεκριμένα, η ακεραιότητα δύναται να οριστεί ως μία δίκαιη και ειλικρινής προσήλωση σε σταθερά πρότυπα ηθικής που έχουν τεθεί εκ των άνω. Έτσι έχει διατυπωθεί η ύπαρξη μίας θετικής συσχέτισης της ακεραιότητας ενός χώρου εργασίας με την συνολική αξιολόγηση της απάτης του οργανισμού (Suh & Shim, 2020).

Κλείνοντας, είναι σημαντικό να αναφερθεί το στοιχείο που θεωρείται ως βάση και εγείρει αναστοχασμό και αξιολόγηση. Αυτό δεν είναι άλλο από την εκπαιδευτική δεοντολογία, μία έννοια που λειτουργεί ως θεμέλιο πρόληψης και ανίχνευσης παρατυπιών. Ξεκινώντας από πάνω στην πυραμίδα ηγεσίας, τα διευθυντικά στελέχη και έπειτα το προσωπικό και το σύνολο των υπαλλήλων οφείλουν να εκπαιδεύονται σε θέματα που σχετίζονται με την απάτη, τι αυτή περιλαμβάνει, τι μορφές φέρει, σε ποιες συνθήκες μπορεί να προκύψει, από ποιους συχνότερα εφαρμόζεται αλλά και σε

λειτουργικά ζητήματα εντοπισμού, πρόληψης και διαχείρισής της. Είναι αναγκαία δηλαδή η επιμόρφωση σε προγράμματα που προσανατολίζονται στις αξίες και στη συμμόρφωση βάσει του ηθικού κλίματος της εταιρείας.

2.2 ΟΙΚΟΝΟΜΙΚΕΣ ΚΑΤΑΣΤΑΣΕΙΣ ΚΑΙ ΚΙΝΔΥΝΟΙ ΑΠΑΤΗΣ

2.2.1 Οικονομικές καταστάσεις

Με τον όρο οικονομικές καταστάσεις αναφερόμαστε στον ισολογισμό, την κατάσταση αποτελεσμάτων χρήσης, την κατάσταση μεταβολών των ιδίων κεφαλαίων, την κατάσταση ταμειακών ροών και το προσάρτημα (Μπατσινίλας & Πατατούκας, 2012). Κατά την παρ. 2 του άρθρου 42α του Ν. 2190/1920, αυτές οι οικονομικές καταστάσεις μιας επιχείρησης είναι υποχρεωτικό να εμφανίζουν την πραγματική εικόνα της περιουσιακής διάρθρωσης, της χρηματοοικονομικής θέσης και των αποτελεσμάτων χρήσης της εταιρείας. Σε μία επιχείρηση, οι οικονομικές καταστάσεις ενημερώνουν για την κερδοφορία, τη δραστηριότητα, τον κίνδυνο και τη ρευστότητα (Νούλας, 2015). Κατά την ελληνική νομοθεσία (ΕΛΠ 2014), το σύνολο των συναλλαγών και γεγονότων της οικονομικής χρήσης που καταχωρούνται στα λογιστικά βιβλία της εταιρείας εμπεριέχονται στις οικονομικές της καταστάσεις. Οι οικονομικές καταστάσεις παρουσιάζουν τα περιουσιακά στοιχεία, τα στοιχεία εσόδων και εξόδων, κερδών και ζημιών καθώς και τις υποχρεώσεις, την καθαρή θέση και τις χρηματορροές κάθε οικονομικής χρήσης λειτουργίας της οικονομικής μονάδας. Απαρτίζουν ένα ενιαίο σύνολο το οποίο απεικονίζει όλα τα παραπάνω στοιχεία εύλογα από κάθε ουσιώδη άποψη. Χρήστες αυτών των οικονομικών καταστάσεων είναι οι άμεσα και έμμεσα ενδιαφερόμενοι για την επιχείρηση (Μπατσινίλας & Πατατούκας, 2012). Οι άμεσα ενδιαφερόμενοι είναι οι μέτοχοι, εταίροι της επιχείρησης, οι πιστωτές, προμηθευτές της, οι πελάτες, το προσωπικό της, η διοίκηση και η φορολογική αρχή. Τους έμμεσα ενδιαφερόμενους αποτελούν οι οικονομικοί αναλυτές, οι τράπεζες, οι εμπορικές και εργατικές ενώσεις, οι δικηγόροι και το χρηματιστήριο.

2.2.2 Απάτη Σφάλμα

Σύμφωνα με τον διεθνή οργανισμό ΠΑ (Institute of Internal Auditors), απάτη είναι μια πράξη η οποία αποβλέπει στην εξαπάτηση, την απόκρυψη και γενικά την κατάχρηση της εμπιστοσύνης και είναι παράνομη. Δεν οφείλεται σε άσκηση σωματικής βίας ή

εξαναγκασμό. Οι πράξεις αυτές διαπράττονται από οργανισμούς ή άτομα με σκοπό την απόκτηση πλούτου, την αποφυγή πληρωμών ή ζημιών για την εξασφάλιση προσωπικού ή επιχειρηματικού οφέλους. Επίσης, ο οργανισμός ACFE (Association of Certified Fraud Examiners) αναφέρει στον ιστότοπό του ότι απάτη είναι η κατάχρηση, από κάποιον υπάλληλο, της θέσης του μέσα στην επιχείρηση με σκοπό τον προσωπικό πλουτισμό του χρησιμοποιώντας επιχειρησιακούς πόρους ή περιουσιακά στοιχεία της επιχείρησης.

Η απάτη μπορεί να οριστεί ως μία ενέργεια ή πράξη, από ένα ή περισσότερα άτομα της οικονομικής μονάδας που είναι επιφορτισμένοι με τη διακυβέρνηση ή τη διοίκησή της, από εργαζομένους της ή από τρίτα μέρη, η οποία έχει γίνει με πρόθεση και έχει ως σκοπό την απόκτηση μη δίκαιου ή παράνομου πλεονεκτήματος (Λουμιώτης & Τζίφας, 2012). Στις οικονομικές καταστάσεις τα σφάλματα που εντοπίζονται προέρχονται από λάθη ή από απάτη. Η διαφορά του λάθους από την απάτη είναι ότι το πρώτο είναι ακούσιο ενώ το δεύτερο εκούσιο. Ο όρος παραποιημένες οικονομικές καταστάσεις προέρχεται από οικονομικές καταστάσεις όπου τα σφάλματα που εντοπίστηκαν προήλθαν σκόπιμα ούτως ώστε να εξαπατηθούν οι χρήστες τους.

Πιο συγκεκριμένα, ο όρος απάτη σχετίζεται με την αλλοίωση των καταχωρήσεων ή των αποδεικτικών στοιχείων, τη μη σωστή παρουσίαση των λογαριασμών ισολογισμού, την καταχώρηση εικονικών συναλλαγών, την εσφαλμένη εφαρμογή των λογιστικών αρχών κι προτύπων και κυρίως στην εκ προθέσεως απόκρυψη ή παράληψη των επιπτώσεων των παραπάνω καταχωρήσεων (Σπαθής, 2002). Το λάθος ορίζεται ως η ακούσια παρερμηνεία πραγματικών συναλλαγών και αφορά λάθη από αβλεψία ή εκ παραδρομής εσφαλμένη εφαρμογή των λογιστικών προτύπων. Οι δύο αυτές έννοιες δεν έχουν ξεκάθαρα όρια ούτως ώστε να μπορέσει κάποιος να κατηγοριοποιήσει τα σφάλματα στις οικονομικές καταστάσεις.

2.2.3 Το τρίγωνο της απάτης

Η θεωρία του τριγώνου της απάτης αναπτύχθηκε από τον Cressey (1953). Σύμφωνα με την θεωρία αυτή η απάτη είναι πιθανόν να διαπραχθεί από ένα ή περισσότερα άτομα του άμεσου ή έμμεσου ενδιαφέροντος της επιχείρησης είτε λόγω της πίεσης που τους ασκείται, είτε λόγω της εμφάνισης ευκαιρίας να διαπραχθεί η απάτη, είτε λόγω της νοοτροπίας αυτού που διαπράττει την απάτη να εκλογικεύει την πράξη αυτή (Cressey, 1953). Πιο αναλυτικά, πίεση ή κίνητρο για την σύνταξη παραπλανητικών οικονομικών καταστάσεων είναι πιθανό να συμβεί όταν η διοίκηση δέχεται πιέσεις από το εσωτερικό

ή εξωτερικό περιβάλλον της ούτως ώστε να επιτύχει έναν χρηματοοικονομικό στόχο ο οποίος δεν είναι ρεαλιστικός (Λουμιώτης & Τζίφας, 2012). Μια ευκαιρία διάπραξης απάτης είναι, για παράδειγμα όταν ένα πρόσωπο που καθίσταται της εμπιστοσύνης της επιχείρησης και γνωρίζει τις αδυναμίες των εσωτερικών δικλίδων, πιστεύει ότι μπορεί να τις παραβιάσει. Κάποια άτομα λόγω της νοοτροπίας τους και των γενικότερων ηθικών αξιών τους είναι πιθανό να διαπράξουν μια απάτη και να εκλογικεύσουν την πράξη τους αυτή. Η θεωρία αυτή έχει χρησιμοποιηθεί ευρέως από τους επαγγελματίες ως ένα θεωρητικό εργαλείο ούτως ώστε να εξηγήσει τον λόγο που συμβαίνουν οι απάτες (Lin, Chiu, Huang & Yen, 2015).

2.3 ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ

Η εξόρυξη δεδομένων είναι η αυτοματοποιημένη διαδικασία ανακάλυψης χρήσιμων πληροφοριών από μεγάλες βάσεις δεδομένων (Tan, Steinbach, & Kumar, 2015). Μέσα από τις τεχνικές εξόρυξης δεδομένων, παρέχεται η δυνατότητα της πρόβλεψης του αποτελέσματος μιας μελλοντικής παρατήρησης. Ο Κύρκος (2015) αναφέρει ότι η Εξόρυξη Δεδομένων είναι αποτέλεσμα της ανάγκης για επεξεργασία και εξαγωγή χρήσιμων πληροφοριών από αποθηκευμένα δεδομένα. Χρησιμοποιεί μεθοδολογίες από πολλούς επιστημονικούς κλάδους και έχει στόχο να ανακαλύψει πολύτιμες πληροφορίες οι οποίες είναι κρυμμένες σε μεγάλους όγκους δεδομένων.

Οι Tan, Steinbach, & Kumar (2015) αναφέρουν ότι η εξόρυξη δεδομένων έχει συμβολή και σε άλλους επιστημονικούς κλάδους όπως η στατιστική, η τεχνητή νοημοσύνη κ.α. και οι εργασίες της χωρίζονται σε δύο μεγάλες κατηγορίες. Οι προγνωστικές εργασίες και οι περιγραφικές εργασίες. Στόχος της πρώτης κατηγορίας είναι η πρόβλεψη μιας τιμής ενός χαρακτηριστικού, βάση των τιμών των υπόλοιπων χαρακτηριστικών. Η δεύτερη κατηγορία έχει ως στόχο την εξαγωγή συσχετίσεων συνοψίζοντας τις σχέσεις που υπάρχουν μεταξύ των δεδομένων.

2.3.1 Η εξόρυξη δεδομένων στον εντοπισμό απάτης στις οικονομικές καταστάσεις

Ο εντοπισμός απάτης στις οικονομικές καταστάσεις μπορεί να θεωρηθεί ως ένα τυπικό πρόβλημα ταξινόμησης (Kirkos, Spathis & Manolopoulos, 2007). Η εξόρυξη

δεδομένων προτείνει διάφορες μεθόδους ταξινόμησης που προέρχονται από τους τομείς της στατιστικής και της τεχνητής νοημοσύνης.

Οι Ravisankar, Raghava & Bose (2011) χρησιμοποίησαν τα δεδομένα 202 εισηγμένων κινέζικων εταιρειών. Οι μισές από αυτές είχαν παραπονημένες οικονομικές καταστάσεις ενώ οι άλλες μισές δεν είχαν. Η ανάλυση των δεδομένων έγινε με τις εξής 6 τεχνικές εξόρυξης δεδομένων: τη μηχανή διανυσμάτων υποστήριξης (SVM), τον εξελικτικό γενετικό προγραμματισμό (GP), το πολυεπίπεδο τεχνητό νευρωνικό δίκτυο (MLFF), τη μέθοδο ομάδων (GMDH), τη λογιστική παλινδρόμηση (LR) και το πιθανοτικό νευρωνικό δίκτυο (PNN). Τα αποτελέσματα της έρευνας έδειξαν ότι όλες οι τεχνικές εξόρυξης εντόπισαν τις παραπονημένες οικονομικές καταστάσεις με ποσοστό ακρίβειας πάνω από 66%. Αυτό σημαίνει ότι όλες οι τεχνικές εντόπισαν τις παραπονημένες οικονομικές καταστάσεις περισσότερες από 2 στις 3 φορές.

Οι Gupta & Gill (2012) συνέλεξαν χρηματοοικονομικά δεδομένα από 114 εισηγμένες εταιρείες σε διάφορα χρηματιστήρια ανά τον κόσμο. Οι 29 από αυτές τις εταιρείες είχαν κατηγορηθεί για παραποίηση οικονομικών καταστάσεων βάσει της επιτροπής κεφαλαιαγοράς των ΗΠΑ (S.E.C.). Τα στοιχεία των εταιρειών αυτών συλλέχθηκαν από το www.wikinvest.com. Οπότε οι 85 από αυτές δεν είχαν παραπονημένες οικονομικές καταστάσεις. Στην έρευνα αυτή χρησιμοποιήθηκαν τρεις τεχνικές εξόρυξης δεδομένων: το δέντρο απόφασης (CART), ο κατηγοριοποιητής του Bayes και ο γενετικός προγραμματισμός. Στα συμπεράσματα η έρευνα αναφέρει ότι οι τρεις μέθοδοι εξόρυξης δεδομένων που χρησιμοποιήθηκαν για τον εντοπισμό απάτης στις οικονομικές καταστάσεις, συγκρίθηκαν με βάση δύο σημαντικά κριτήρια, την ευαισθησία και την ειδικότητα. Ως αποτέλεσμα, αυτές οι τεχνικές είναι σε θέση να εντοπίσουν παραπλανητικές χρηματοοικονομικές αναφορές. Επίσης κατάφεραν να εντοπίσουν εάν η διαχείριση του οργανισμού είναι σε θέση να διαπράξει απάτη στις οικονομικές καταστάσεις.

Η έρευνα των Lin Chi, Chiu, Huang Shaio & Yen David (2015), συγκεντρώνει χρηματοοικονομικά δεδομένα από 129 εταιρείες που είχαν παραπονημένες οικονομικές καταστάσεις κατά την περίοδο 1998 – 2010, σύμφωνα με τις αρχές της Ταϊβάν. Για την ίδια περίοδο συγκέντρωσαν δεδομένα για 447 εταιρείες οι οποίες δεν είχαν παραπονήσει τις οικονομικές τους καταστάσεις. Οι τεχνικές εξόρυξης δεδομένων που χρησιμοποίησαν για την ανάλυση των δεδομένων ήταν η λογιστική παλινδρόμηση (LR), το τεχνητό νευρωνικό δίκτυο (ANN) και το δέντρο απόφασης (CART). Εξαιρώντας όλους τους υπόλοιπους ελέγχους που έγιναν στην έρευνα παρατηρούμε ότι

η συνολική ακρίβεια των τριών τεχνικών στον εντοπισμό απάτης στις οικονομικές καταστάσεις ήταν από 88% έως 92% .

Ο Chen (2016) στην έρευνά του χρησιμοποιεί 4 διαφορετικές τεχνικές εξόρυξης δεδομένων. Το δέντρο απόφασης (DT), το δίκτυο πεποιθήσεως κατά Bayes (BBN), το νευρωνικό δίκτυο (ANN) και τη μηχανή διανυσμάτων υποστήριξης (SVM). Για τα δέντρα απόφασης χρησιμοποιεί δύο διαφορετικούς αλγόριθμους, τον CART C0.5 και τον CHAID. Το δείγμα που χρησιμοποιεί είναι 176 εταιρείες εισηγμένες και μη στο χρηματιστήριο της Ταϊβάν, από τις οποίες οι 44 είχαν παραπονημένες οικονομικές καταστάσεις ενώ οι υπόλοιπες 132 όχι. Για την περίοδο 2002 – 2013. Στα ευρήματα της έρευνας εντοπίζεται ότι το χαμηλότερο ποσοστό ακρίβειας εντοπισμού απάτης στις οικονομικές καταστάσεις είναι το 75% ενώ το υψηλότερο ποσοστό ακρίβειας είναι κοντά στο 88% .

Σε μια μεγάλη έρευνα με σημαντικά αποτελέσματα προχώρησαν οι Dutta I, Dutta S & Raahemi (2017) οι οποίοι χρησιμοποίησαν την βάση δεδομένων Audit Analytics (AA), περιορίζοντας το δείγμα για τις οικονομικές χρήσεις από το 2001 έως το 2014. Συνέλλεξαν δεδομένα για την επαναδιατύπωση οικονομικών καταστάσεων. Επιπλέον, από την βάση δεδομένων COMPUSTAT συνέλλεξαν όλα τα σχετικά δεδομένα για τις περιπτώσεις που αντλήθηκαν από την Audit Analytics. Με διάφορες τεχνικές που χρησιμοποίησαν έφτασαν σε ένα σύνολο περιπτώσεων 64.233 όπου οι 3513 περιπτώσεις ήταν περιπτώσεις παραπονημένων οικονομικών καταστάσεων. Στην συνέχεια χρησιμοποίησαν 4 τεχνικές εξόρυξης δεδομένων για να κάνουν την ανάλυση και τον εντοπισμό των παραπλανητικών οικονομικών καταστάσεων. Οι τεχνικές που χρησιμοποίησαν είναι το δέντρο απόφασης, το νευρωνικό δίκτυο, το δίκτυο πεποιθήσεων κατά Bayes, ο κατηγοριοποιητής κατά Bayes και η μηχανή διανυσμάτων υποστήριξης. Τα αποτελέσματά της έρευνα είναι πολλά υποσχόμενα και παρουσιάζουν αξιόπιστα μοντέλα πρόβλεψης που μπορούν να καθοδηγήσουν τους επενδυτές και τους ελεγκτές να εντοπίσουν μια εταιρεία με παραπονημένες οικονομικές καταστάσεις ή με πιθανότητα ουσιωδών σφαλμάτων στις οικονομικές καταστάσεις.

Οι Cihan Sorkun & Toraman (2017) διερεύνησαν 72 επιχειρήσεις που προήλθαν από μια διαδικασία επιλογής μέσω του προγράμματος e-ledger. Χρησιμοποίησαν 7 διαφορετικές τεχνικές εξόρυξης δεδομένων: τη γραμμική παλινδρόμηση, τα τεχνητά νευρωνικά δίκτυα, τον αλγόριθμο του πλησιέστερου γείτονα, τη μηχανή διανυσμάτων υποστήριξης και τρεις διαφορετικούς αλγόριθμους της τεχνικής δέντρου απόφασης. Στο τελικό αποτέλεσμα αναφέρει ότι το δέντρο απόφασης έχει ποσοστό ακρίβειας 80%

στον εντοπισμό απάτης στις οικονομικές καταστάσεις. Οι υπόλοιπες τεχνικές κυμαίνονται μεταξύ 60% και 77% εκτός από την τεχνική του πλησιέστερου γείτονα που είχε ακρίβεια 37%.

Στην πιο πρόσφατη έρευνα οι Mohammadi, Yazdani, Khanmohammadi, Maham Keyhan (2020) ανέλυσαν τα δεδομένα 165 εταιρειών με παραποιημένες οικονομικές καταστάσεις και 165 εταιρειών χωρίς παραποιημένες τις οικονομικές τους καταστάσεις. Το δείγμα αυτό το άντλησαν από το χρηματιστήριο της Τεχεράνης για τα έτη 2011 έως 2016. Χρησιμοποίησαν τις παρακάτω 5 τεχνικές εξόρυξης δεδομένων: Δίκτυο κατά Bayes, ανάλυση συστάδων, λογιστική παλινδρόμηση, τεχνητά νευρωνικά δίκτυα και μηχανή διανυσμάτων υποστήριξης. Κατά μέσο όρο οι τεχνικές υπέδειξαν ακρίβεια της τάξεως του 68% στον εντοπισμό απάτης στις οικονομικές καταστάσεις.

ΚΕΦΑΛΑΙΟ 3: ΜΕΘΟΔΟΛΟΓΙΑ

3.1 ΜΕΘΟΔΟΛΟΓΙΑ

Οι Özdağoğlu, G., Özdağoğlu, A., Gümüş, Y., & Kurt-Gümüş, G., (2017), αναφέρουν ότι οι τεχνικές εξόρυξης δεδομένων μπορούν να χρησιμοποιηθούν για την ανάπτυξη εργαλείων υποστήριξης στην ελεγκτική διαδικασία. Πιο συγκεκριμένα, αναλύουν στην έρευνά τους την σημασία της γνώμης του ελεγκτή και πως η εξόρυξη δεδομένων μπορεί να βοηθήσει αυτήν την διαδικασία.

Την τελευταία δεκαετία η βιβλιογραφία έχει εμπλουτιστεί με αρκετές έρευνες όπου παρουσιάζονται σημαντικά αποτελέσματα στον εντοπισμό απάτης στις οικονομικές καταστάσεις με την χρήση τεχνικών εξόρυξης δεδομένων. Στο παρόν κεφάλαιο παρουσιάζεται ένα αντιπροσωπευτικό δείγμα από έρευνες, που έγιναν στον τομέα της εξόρυξης δεδομένων με σκοπό τον εντοπισμό απάτης σε οικονομικές καταστάσεις, μέσα στην τελευταία δεκαετία. Στόχος είναι να αναλυθούν και να παρουσιαστούν τα αποτελέσματα των τεχνικών εξόρυξης δεδομένων, ούτως ώστε οι ελεγκτές να έχουν συγκεντρωμένη πληροφορία και να μπορούν να πάρουν την σωστή απόφαση για την τεχνική που θα πρέπει να χρησιμοποιήσουν.

Οι έρευνες που αναλύονται στην παρούσα εργασία παρουσιάζονται στον πίνακα 1:

Authors	Title	Year	Published In	Reference
Ravisankar, P. Ravi, V. Raghava Rao, G. Bose, I.	Detection of financial statement fraud and feature selection using data mining techniques	2011	Decision Support Systems	(Ravisankar et al., 2011)
Gupta Rajan, Singh Gill Nasib	A Data Mining Framework for Prevention and Detection of Financial Statement Fraud	2012	International Journal of Computer Applications	(Gupta & Singh Gill, 2012)
Lin Chi Chen, Chiu An An, Huang Shaio Yan, Yen David C.	Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments	2015	Knowledge-Based Systems	(Lin et al., 2015)
Chen Suduan	Detection of fraudulent financial statements using the hybrid data mining approach	2016	SpringerPlus	(Chen, 2016)
Dutta Ila, Dutta Shantanu, Raahemi Bijan	Detecting financial restatements using data mining techniques	2017	Expert Systems with Applications	(Dutta et al., 2017)
Özdağoğlu G., Özdağoğlu A., Gümüş Y., Kurt- Gümüş G.	The application of data mining techniques in manipulated financial statement classification: The case of turkey	2017	Journal of AI and Data Mining	(Özdağoğlu et al., 2017)
Chen Suduan, Yang Alex	An Effective Financial Statements Fraud Detection Model	2018	DEStech Transactions on Engineering and Technology Research	(CHEN & YANG, 2018)
Mohammadi Mahmood, Yazdani Shohreh, Khanmohammadi Mohammad Hamed, Maham Keyhan	Financial Reporting Fraud Detection: An Analysis of Data Mining Algorithms	2020	International Journal of Finance & Managerial Accounting	(Mohammadi et al., 2020)

Στις εργασίες αυτές παρουσιάζονται αρκετές διαφορετικές τεχνικές εξόρυξης δεδομένων. Κάποιες από αυτές παρουσιάζονται σε παραπάνω από μία έρευνες. Αυτό είναι πολύ φυσιολογικό καθώς οι τεχνικές εξόρυξης δεδομένων είναι αλγόριθμοι ταξινόμησης οι οποίοι μπορούν να εφαρμοστούν σε πολλά επιστημονικά πεδία. Στον παρακάτω πίνακα καταγράφονται οι τεχνικές που παρουσιάζονται στις έρευνες που αναλύει η παρούσα εργασία.

Στον πίνακα 2 βλέπουμε τις τεχνικές εξόρυξης δεδομένων που χρησιμοποιήθηκαν σε αυτές τις έρευνες, για να γίνει η κατηγοριοποίηση των οικονομικών καταστάσεων σε παραποιημένες και μη

Πίνακας 2: Τεχνικές εξόρυξης δεδομένων

1	Artificial Neural Network (ANN)	Τεχνητά Νευρωνικά Δίκτυα
2	Bayesian belief network (BBN)	Δίκτυα Πεποιθήσεων Bayes
3	Discriminant Analysis (DA)	Διακριτική Ανάλυση
4	Decision Tree (DT)	Δέντρα Απόφασης
5	Group method of data handling (GMDH)	Μέθοδος Ομάδων Διαχείρισης Δεδομένων
6	Genetic Programming (GP)	Γενετικός Προγραμματισμός
7	Logistic regression (LR)	Λογιστική Παλινδρόμηση
8	Multi-layer feedforward neural network (MLFF)	Πολυεπίπεδα Νευρωνικά Δίκτυο
9	Naïve Bayes (NB)	Naïve Bayes
10	Probabilistic neural network (PNN)	Πιθανοτικά Νευρωνικά Δίκτυο
11	Support vector machines (SVM)	Μηχανή Διανυσμάτων Υποστήριξης

Όλες αυτές οι τεχνικές εξόρυξης δεδομένων αποτελούν ουσιαστικά αλγόριθμους ταξινόμησης. Η ταξινόμηση θεωρείται κλασικό παράδειγμα εποπτευόμενης μάθησης. Αυτό σημαίνει ότι οι τεχνικές εξόρυξης δεδομένων, που παρουσιάζονται στον πίνακα 2, εκπαιδεύονται, κατ' αρχήν, με ένα σύνολο δεδομένων σωστά προσδιοριζόμενων παρατηρήσεων. Δηλαδή πέρα από τις μεταβλητές εισόδου τους δίνεται και το αποτέλεσμα της ταξινόμησης. Το σύνολο αυτών των δεδομένων είναι το σύνολο δεδομένων εκπαίδευσης.

Πιο συγκεκριμένα, στην αρχή των ερευνών οι παραπάνω τεχνικές εφαρμόζονται σε οικονομικές καταστάσεις όπου, εκτός από τις μεταβλητές εισόδου γίνεται γνωστό σε όλες τις τεχνικές αν οι οικονομικές καταστάσεις είναι παραπονημένες ή όχι. Στόχος της συγκεκριμένης διαδικασίας είναι να εκπαιδευτούν οι τεχνικές στην ταξινόμηση των οικονομικών καταστάσεων με βάση τα δεδομένα εισόδου. Στην συνέχεια εφαρμόζονται σε οικονομικές καταστάσεις χωρίς να τους είναι γνωστό αν είναι παραπονημένες ή όχι, οπότε και αναλύονται οι επιδόσεις τους.

Δεδομένα εισόδου είναι οι μεταβλητές βάσει των οποίων γίνεται έλεγχος των παραπονημένων οικονομικών καταστάσεων. Πιο συγκεκριμένα, ως δεδομένα εισόδου των τεχνικών επιλέγονται χρηματοοικονομικοί δείκτες και μη χρηματοοικονομικοί δείκτες που έχει παρατηρηθεί ότι είναι σημαντικοί στην ένδειξη για παραποινή στις οικονομικές καταστάσεις. Αναλυτικότερα, ακολουθούν οι πίνακες 3 και 4 όπου παρουσιάζονται συνολικά οι δείκτες που χρησιμοποιήθηκαν στις έρευνες που αναλύονται στην παρούσα εργασία, ως μεταβλητές εισόδου των τεχνικών εξόρυξης δεδομένων.

Πίνακας 3: Χρηματοοικονομικοί δείκτες

Χρηματοοικονομικοί δείκτες
Ανάπτυξη πωλήσεων
Απαιτήσεις
Απαιτήσεις / Απαιτήσεις προηγούμενης χρήσης
Απαιτήσεις / Σύνολο ενεργητικού
Απαιτήσεις / Σύνολο εσόδων
Απαιτήσεις / Σύνολο πωλήσεων
Αποδοτικότητα επενδυμένων κεφαλαίων
Απόθεμα / Σύνολο ενεργητικού
Αποθέματα / Καθαρές πωλήσεις
Απόθεμα / Σύνολο εσόδων
Αποθέματα / Βραχυπρόθεσμες υποχρεώσεις
Βραχυπρόθεσμες υποχρεώσεις προηγούμενων χρήσεων
Βραχυπρόθεσμες υποχρεώσεις / σύνολο ενεργητικού
Δείκτης άμεσης ρευστότητας
Δείκτης κυκλοφοριακής ρευστότητας
Διατηρούμενα κέρδη / Σύνολο ενεργητικού
Έσοδα / Σύνολο ενεργητικού
Καθαρά έσοδα / Πάγιο ενεργητικό
Καθαρή θέση
Καθαρό κέδρος / Σύνολο ενεργητικού
Καθαρό κέρδος
Καθαρό κέρδος / Μεικτό κέρδος
Καθαρό κέρδος / Σύνολο εσόδων
Κέρδη προ φόρων / Κέρδη προ φόρων προηγούμενης χρήσης
Κέρδη προ φόρων και τόκων (ΚΠΦΤ)
Κέρδος προς διάθεση / Καθαρό κέρδος
Κεφάλαιο και Αποθεματικό / Συνολικό χρέος
Κεφάλαιο κίνησης / Σύνολο ενεργητικού
ΚΠΦΤ / Σύνολο ενεργητικού
Κυκλοφορούν ενεργητικό / Σύνολο ενεργητικού
Κυκλοφοριακή ταχύτητα απαιτήσεων
Κυκλοφοριακή ταχύτητα αποθεμάτων
Κυκλοφοριακή ταχύτητα παγίων
Κυκλοφορούν ενεργητικό
Κυκλοφορούν ενεργητικό / Βραχυπρόθεσμες υποχρεώσεις
Λειτουργικά έξοδα
Μακροπρόθεσμες υποχρεώσεις / Κεφάλαιο και Αποθεματικό
Μακροπρόθεσμες υποχρεώσεις / Σύνολο ενεργητικού
Μέγεθος επιχείρησης στην βάση ενεργητικού
Μεικτό κέρδος
Μεικτό κέρδος / Κέρδη προ φόρων
Μεικτό κέρδος / Σύνολο ενεργητικού

Πάγιο ενεργητικό
Πάγιο ενεργητικό / Σύνολο ενεργητικού
Ρευστοποίηση περιουσιακών στοιχείων
Συνολικές πωλήσεις / Σύνολο ενεργητικού
Συνολικό χρέος / Σύνολο ενεργητικού
Σύνολο ενεργητικού
Σύνολο ενεργητικού / Κεφάλαιο και Αποθεματικό
Σύνολο ενεργητικού / Σύνολο ενεργητικού προηγούμενης χρήσης
Σύνολο εσόδων
Σύνολο εσόδων / Πάγιο ενεργητικό
Σύνολο εσόδων / Σύνολο ενεργητικού
Σύνολο εσόδων / Σύνολο εσόδων προηγούμενης χρήσης
Σύνολο υποχρεώσεων / Σύνολο ενεργητικού
Σύνολο υποχρεώσεων προηγούμενων χρήσεων
Σύνολο χρέους / Σύνολο ιδίων κεφαλαίων
Χρέος
Χρεωστικοί τόκοι
Χρεωστικοί τόκοι / Σύνολο υποχρεώσεων
Χρηματικά διαθέσιμα
Χρηματικά διαθέσιμα / Κυκλοφορούν ενεργητικό
Χρηματικά διαθέσιμα / Σύνολο ενεργητικού
Χρηματοροές
Χρηματοροές ενεργητικού

Πίνακας 4: Μη χρηματοοικονομικοί δείκτες

Μη χρηματοοικονομικοί δείκτες
Ποσοστό ενδοομολικών πωλήσεων
Ποσοστό ενδοομικών αγορών
Δείκτης ενδοομικού κινδύνου
Δείκτης επενδύσεων ιδίων κεφαλαίων
Ποσοστό επανεπένδυσης
Ποσοστό πωλήσεων στο εξωτερικό
Εσωτερική εκμετάλλευση*(πριν δύο χρήσεις) - Εσωτερική εκμετάλλευση (προηγούμενη χρήση)
Μεταβλητή με τιμή 1 εάν η εταιρεία είναι οικογενειακή επιχείρηση, διαφορετικά 0
Απόκλιση κοινών και προνομιούχων μετοχών
Ποσοστό ιδιοκτησίας διοίκησης
Ποσοστό ιδιοκτησίας blockholder
Μεταβλητή με τιμή 1 αν ο εσωτερικός έλεγχος έχει δηλώσει σημαντικές ελλείψεις, διαφορετικά 0
Αριθμός εσωτερικών ελεγκτών που άλλαξαν τις τελευταίες τρεις χρήσεις
Περίοδος από την ίδρυση μέχρι τον εντοπισμό απάτης
Μεταβλητή με τιμή 1 αν ο εξωτερικός έλεγχος γίνεται από εταιρεία Big - 4, διαφορετικά 0
Συμπεριφορά που ταιριάζει με την πράξη
Αριθμός CEO που άλλαξαν τις τελευταίες 3 χρήσεις

Αριθμός CFO που άλλαξαν τις τελευταίες 3 χρήσεις
Αριθμός CPA που άλλαξαν τις τελευταίες 3 χρήσεις
Μεταβλητή με τιμή 1 αν ο πρόεδρος του ΔΣ κατέχει θέση διευθύνοντος συμβούλου στην εταιρεία ή προέδρου ή κάποιας διευθυντικής θέσης, διαφορετικά 0
Αριθμός επαναδημοσίευσης των οικονομικών καταστάσεων που επηρέαζαν τα κέρδη τις δύο προηγούμενες χρήσεις

Οι μεταβλητές αυτές δίνονται στις τεχνικές εξόρυξης δεδομένων ως δεδομένα εισόδου. Έχει παρατηρηθεί ότι όταν οι οικονομικές καταστάσεις είναι παραποιημένες τότε οι μεταβλητές αυτές έχουν ένα συγκεκριμένο εύρος τιμών. Οπότε κατά την εκπαίδευση οι τεχνικές μαθαίνουν πότε οι οικονομικές καταστάσεις περιέχουν σφάλματα. Ως αποτέλεσμα, όταν τους δίνονται δεδομένα χωρίς να γνωρίζουν αν είναι από παραποιημένες οικονομικές καταστάσεις, κατηγοριοποιούν αυτά τα δεδομένα σε παραποιημένες ή μη παραποιημένες οικονομικές καταστάσεις.

Τα αποτελέσματα των τεχνικών αναλύονται με στατιστικές μεθόδους ούτως ώστε να μετρηθεί η ακρίβεια τους στην κατηγοριοποίηση των οικονομικών καταστάσεων σε παραποιημένες ή μη.

Η ακρίβεια τους αναλύεται με την στατιστική μέθοδο των υποθέσεων. Στη μηδενική υπόθεση τίθενται οι περιπτώσεις όπου οι οικονομικές καταστάσεις δεν είναι παραποιημένες. Οπότε στην εναλλακτική υπόθεση είναι οι οικονομικές καταστάσεις που είναι παραποιημένες. Οι τεχνικές εξόρυξης δεδομένων κατηγοριοποιούν τις οικονομικές καταστάσεις σε παραποιημένες ή μη. Άρα, υπολογίζεται η απόδοση της κατηγοριοποίησης των παραποιημένων οικονομικών καταστάσεων ως παραποιημένες. Αντίστοιχα, υπολογίζεται η απόδοση της σωστής κατηγοριοποίησης των μη παραποιημένων οικονομικών καταστάσεων. Έτσι, υπολογίζονται και τα σφάλματα τύπου I και τύπου II. Σφάλμα τύπου I υπάρχει όταν ισχυριζόμαστε ότι έχουμε βρει κάτι το οποίο στην πραγματικότητα δεν υπάρχει (Μενεξές & Οικονόμου, 2002). Ενώ σφάλμα τύπου II υπάρχει όταν αποτυγχάνουμε να βρούμε κάτι το οποίο όμως στην πραγματικότητα υπάρχει. Στην συνολική ακρίβεια των τεχνικών προσμετράτε η σωστή κατηγοριοποίηση και των παραποιημένων και των μη παραποιημένων οικονομικών καταστάσεων.

Ακολουθεί ένα υπόδειγμα ελέγχου υποθέσεων σε γενική μορφή ούτως ώστε να γίνει κατανοητή η μεθοδολογία του ελέγχου των αποτελεσμάτων των ερευνών που παρουσιάζονται στην παρούσα εργασία.

Υπόδειγμα ελέγχου υποθέσεων:

Μη παραπονημένες οικονομικές καταστάσεις = H_0

Παραπονημένες οικονομικές καταστάσεις $\neq H_0$

	H_0	$\neq H_0$
H_0	Σωστό (TP)	Σφάλμα τύπου II (FP)
$\neq H_0$	Σφάλμα τύπου I (FN)	Σωστό (TN)

Στον οριζόντιο άξονα του πίνακα απεικονίζεται η πραγματική κατάσταση των οικονομικών καταστάσεων, ενώ στον κάθετο οι κατηγοριοποιήσεις των τεχνικών εξόρυξης δεδομένων.

3.2 ΣΥΓΚΕΝΤΡΩΣΗ ΔΕΔΟΜΕΝΩΝ

Τα δεδομένα των ερευνών αυτών συγκεντρώνονται από διάφορες βάσεις δεδομένων και καλύπτουν ένα ευρύ φάσμα επιχειρήσεων και εταιρειών παγκοσμίως. Επίσης το χρονικό διάστημα που καλύπτουν είναι από τις αρχές του 2000 έως και τα μέσα τις δεκαετίας του 2010. Οπότε το συνολικό δείγμα είναι ιδιαίτερα αντιπροσωπευτικό για να βγουν ασφαλή συμπεράσματα.

ΚΕΦΑΛΑΙΟ 4: ΠΑΡΟΥΣΙΑΣΗ ΕΡΓΟΥ

4.1 ΔΗΜΙΟΥΡΓΙΑ ΣΥΓΚΕΝΤΡΩΤΙΚΟΥ ΠΙΝΑΚΑ

Τα αποτελέσματα όλων των ερευνών, που παρουσιάζονται και συγκρίνονται στην παρούσα εργασία, εξάχθηκαν σε μορφή πινάκων. Στους πίνακες αναγράφονται όλα τα στοιχεία που παρουσιάζουν οι ερευνητές στα άρθρα τους στο κεφάλαιο της παρουσίασης των αποτελεσμάτων. Όλες οι εργασίες ακολουθούν την ίδια διαδικασία διερεύνησης του ζητήματος του εντοπισμού παραποιημένων οικονομικών καταστάσεων με τεχνικές εξόρυξης δεδομένων. Κατ' αρχάς παρουσιάζουν τον συνολικό αριθμό των παρατηρήσεων πάνω στο οποίο θα γίνει η εφαρμογή των τεχνικών εξόρυξης δεδομένων. Οι παρατηρήσεις αυτές είναι δημοσιευμένες οικονομικές καταστάσεις διάφορων εταιρειών από διάφορες βάσεις δεδομένων. Επίσης οι οικονομικές καταστάσεις αυτές είναι χαρακτηρισμένες ως παραποιημένες ή όχι από τις εκθέσεις ελεγκτών που τις ακολουθούν. Το κάνουν αυτό οι ερευνητές για να εντοπίσουν την ακρίβεια των τεχνικών που εφαρμόζουν, στην κατηγοριοποίηση των οικονομικών καταστάσεων. Το δεύτερο βήμα της διαδικασίας που ακολουθείτε από τους συγκεκριμένους ερευνητές είναι ο καθορισμός των μεταβλητών, βάσει των οποίων θα γίνει η αναγνώριση από τις τεχνικές εξόρυξης αν μια οικονομική κατάσταση είναι παραποιημένη ή όχι. Πιο συγκεκριμένα, οι μεταβλητές αυτές είναι γνωστοί χρηματοοικονομικοί δείκτες στην επιστήμη της χρηματοοικονομικής και λογιστικής και παρουσιάζονται στο κεφάλαιο 3 της παρούσας εργασίας. Η αξία των δεικτών αυτών εξάγεται από τις οικονομικές καταστάσεις. Επίσης υπάρχουν και μεταβλητές που δεν είναι χρηματοοικονομικής φύσεως. Έχουν να κάνουν με την διοικητική και μετοχική διάρθρωση της εταιρείας, οι οικονομικές καταστάσεις της οποίας συμμετέχουν στο δείγμα της έρευνας. Στο τρίτο βήμα οι ερευνητές παρουσιάζουν τις τεχνικές εξόρυξης δεδομένων που θα χρησιμοποιήσουν και τη μεθοδολογία που θα ακολουθήσουν. Η μεθοδολογία που ακολουθούν είναι ότι στο πρώτο στάδιο των ερευνών τους εκπαιδεύουν τις τεχνικές εξόρυξης δεδομένων. Αυτό σημαίνει ότι εισάγουν τους χρηματοοικονομικούς δείκτες των συγκεκριμένων οικονομικών καταστάσεων, ως δεδομένα εισόδου και τους δίνουν και την κατηγοριοποίηση της συγκεκριμένης παρατήρησης. Δηλαδή τους κάνουν γνωστό το αποτέλεσμα της διαδικασίας που πρέπει να ακολουθήσουν. Αυτό έχει στόχο ώστε οι αλγόριθμοι να κατανοήσουν πότε μια οικονομική κατάσταση είναι παραποιημένη και πότε όχι, με βάση τους

χρηματοοικονομικούς δείκτες της εταιρείας. Αφού ολοκληρωθεί η εκπαίδευση οι ερευνητές εισάγουν δεδομένα στις τεχνικές ούτως ώστε να κάνουν την κατηγοριοποίηση, οι τεχνικές, μα βάσει την εκπαίδευσή τους. Σε αυτό το σημείο μετράνε την ακρίβεια στην κατηγοριοποίηση.

Η παρούσα εργασία εξάγει τα δεδομένα, που παρουσιάζονται στο κεφάλαιο των αποτελεσμάτων αυτών των ερευνών, σε μορφή πίνακα. Στην συνέχεια γίνεται ένα φιλτράρισμα των δεδομένων ώστε να παρουσιαστούν ομοιόμορφα. Τα δεδομένα που παρουσιάζονται είναι:

1. Το ποσοστό σωστής κατηγοριοποίησης των μη παραπονημένων οικονομικών καταστάσεων.
2. Το ποσοστό σωστής κατηγοριοποίησης των παραπονημένων οικονομικών καταστάσεων.
3. Το συνολικό ποσοστό σωστής κατηγοριοποίησης όλων των οικονομικών καταστάσεων παραπονημένων και μη.
4. Τον δείκτη AUC.

Στην συνέχεια παρουσιάζεται ένας συνολικός πίνακας με όλες τις τεχνικές ανά έρευνα συγκεντρωτικά. Στο τέλος γίνεται προσπάθεια υπολογισμού του συνολικού ποσοστού ακρίβειας της κάθε τεχνικής από τις 11 που εφαρμόζονται σε αυτές τις έρευνες ώστε να παρουσιαστούν σε έναν πίνακα, ταξινομημένες με βάση την συνολική τους ακρίβεια σε όλες τις έρευνες.

Η συγκέντρωση των δεδομένων γίνεται από τους πίνακες αποτελεσματικότητας που υπάρχουν στις έρευνες. Κάθε άρθρο έχει δημιουργήσει έναν ή παραπάνω πίνακες με τα αποτελέσματα της έρευνάς του. Τα αποτελέσματα εμφανίζονται σε ποσοστό. Δείχνουν την αποτελεσματικότητα κατηγοριοποίησης των τεχνικών εξόρυξης δεδομένων. Ανάλυση έγινε με την στατιστική μέθοδο της μηδενικής υπόθεσης. Ανάλογα το μέγεθος του δείγματος και των μεταβλητών αλλά και τον αριθμό τεχνικών που χρησιμοποίησε ο κάθε ερευνητής, τα αποτελέσματα εμφανίζονται είτε

συγκεντρωμένα σε έναν πίνακα είτε σε περισσότερους. Τα σφάλματα τύπου I και τύπου II είναι σε αναλογία με την Ευαισθησία και την Ειδικότητα αντίστοιχα.

Το ποσοστό όπου μία τεχνική κατηγοριοποίησε μια παρατήρηση ως παραποιημένη οικονομική κατάσταση σωστά αναφέρεται στις έρευνες ως Ευαισθησία. Αντίστοιχα το ποσοστό σωστής κατηγοριοποίησης των μη παραποιημένων οικονομικών καταστάσεων αναφέρεται στις έρευνες ως Ειδικότητα.

Εκτός από την συνολική ακρίβεια, την ειδικότητα, την ευαισθησία και τα σφάλματα τύπου I και II, στις έρευνες αναφέρεται και ο δείκτης AUC. Ο δείκτης αυτός στα μαθηματικά δείχνει το εμβαδό κάτω από την καμπύλη ROC. Η καμπύλη ROC απεικονίζει το σύνολο των δυνατών ζευγών σωστής κατηγοριοποίησης και σφάλματος τύπου I που μπορεί να ληφθούν. Οπότε ο δείκτης AUC μπορεί να ερμηνευθεί ως α) η μέση τιμή της Ευαισθησίας για όλες τις πιθανές τιμές της ειδικότητας, β) η μέση τιμή της Ειδικότητας για όλες τις πιθανές τιμές της Ευαισθησίας.

Στην συνέχεια παρουσιάζονται οι πίνακες αποτελεσμάτων της κάθε έρευνας με αναφορά στον συνολικό αριθμό του δείγματος που συμμετέχει στην έρευνα. Οι πίνακες αυτοί έχουν υποστεί επεξεργασία στο κομμάτι της εμφάνισης ούτως ώστε να εμφανίζονται ομοιόμορφα στην παρούσα εργασία. Επίσης η εμφάνισή τους γίνεται με χρονολογική σειρά, από την παλαιότερη στην πιο πρόσφατη.

Τέλος για λόγους παρουσίασης οι τεχνικές εξόρυξης δεδομένων αναγράφονται με συντομογραφία. Οι συντομογραφίες παρουσιάζονται στον παρακάτω πίνακα:

Πίνακας5: Συντομογραφίες τεχνικών εξόρυξης δεδομένων

Συν/φία	Ονομασία
ANN	Artificial Neural Network
BBN	Bayesian belief network
DA	Discriminant Analysis
DT	Decision Tree
GMDH	Group method of data handling
GP	Genetic Programming
LR	Logistic regression
MLFF	Multi-layer feedforward neural network
NB	Naïve Bayes
PNN	Probabilistic neural network
SVM	Support vector machines

Στον πίνακα 2 του τρίτου κεφαλαίου της παρούσας εργασίας αναγράφεται η μετάφραση της ονομασίας των τεχνικών αυτών στα Ελληνικά.

ΠΙΝΑΚΕΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΕΡΕΥΝΩΝ:

1.

Ravisankar, P., Ravi, V., Raghava Rao, G., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. <i>Decision Support Systems</i> , 50(2), 491–500.				
Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Εναισθησία	Ειδικότητα	AUC
MLFF	78,77%	76,98%	81,28%	0,7913
SVM	73,41%	72,07%	75,04%	0,7356
GP	92,68%	90,55%	95,27%	0,9291
GMDH	90,68%	93,46%	88,34%	0,9090
LR	70,36%	62,91%	78,88%	0,7090
PNN	95,64%	91,27%	94,16%	0,9272

Το σύνολο των δεδομένων που χρησιμοποιήθηκε σε αυτή την έρευνα είναι 202 κινέζικες εισηγμένες εταιρείες, εκ των οποίων οι 101 είχαν παραποιημένες οικονομικές καταστάσεις και οι υπόλοιπες 101 δεν είχαν. Ως δεδομένα εισόδου των τεχνικών είναι 35 χρηματοοικονομικοί δείκτες, που παρουσιάζονται στον πίνακα 3 του τρίτου κεφαλαίου της παρούσας εργασίας.

2.

Gupta, R., & Singh Gill, N. (2012). A Data Mining Framework for Prevention and Detection of Financial Statement Fraud. <i>International Journal of Computer Applications</i> , 50(8), 7–14.				
Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	AUC
DT	94,74%	86,20%	97,70%	0,9195
NB	87,72%	84,00%	92,90%	0,8845
GP	87,72%	53,00%	99,20%	0,7610

Το σύνολο δεδομένων αυτής της έρευνας είναι 114 εισηγμένες εταιρείες σε διάφορα χρηματιστήρια του κόσμου από τις οποίες οι 85 δεν έχουν παραπονημένες οικονομικές καταστάσεις ενώ οι υπόλοιπες 29 έχουν. Τα δεδομένα συλλέχθηκαν από το www.wikinvest.com.

3.

Lin, C. C., Chiu, A. A., Huang, S. Y., & Yen, D. C. (2015). Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments. <i>Knowledge-Based Systems</i> , 89, 459–470.				
Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	AUC
LR	88,48%	71,79%	93,65%	0,8272
DT	90,32%	81,58%	93,16%	0,8737
ANN	92,90%	82,93%	96,09%	0,8951

Το σύνολο των δεδομένων στην έρευνα αυτή είναι 129 εταιρείες της Ταϊβάν με παραπονημένες οικονομικές καταστάσεις και 447 με μη παραπονημένες οικονομικές καταστάσεις.

4.

Dutta, I., Dutta, S., & Raahemi, B. (2017). Detecting financial restatements using data mining techniques. <i>Expert Systems with Applications</i> , 90, 374–393.				
Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	AUC
DT	77,76%	77,80%	68,80%	0,8180
ANN	77,98%	78,00%	63,70%	0,8300
NB	60,21%	60,20%	68,90%	0,6930
SVM	65,60%	65,60%	66,70%	0,6620
BBN	72,62%	72,60%	68,90%	0,7860

Το σύνολο δεδομένων της έρευνας έχει συλλεχθεί από αρκετές βάσεις δεδομένων παγκοσμίως που κρατούν στοιχεία οικονομικών καταστάσεων διαφόρων εταιρειών ανά τον κόσμο. Καλύπτει ένα διάστημα από το 2001 έως το 2014 και έχει συνολικά 64.233 παρατηρήσεις από τις οποίες οι 3513 είναι περιπτώσεις παραποιημένων οικονομικών καταστάσεων.

5.

Chen, S. (2016). Detection of fraudulent financial statements using the hybrid data mining approach. <i>SpringerPlus</i> , 5(1), 1–16.				
Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	AUC
CART–DT(CART)	83,19	88,59	77,78	0,8319
CART–DT(CHRID)	80,70	81,88	79,51	0,8070
CART–BBN	75,20	77,18	73,22	0,7520
CART–SVM	74,68	75,17	74,19	0,7468
CART–ANN	75,00	75,84	74,16	0,7500
CHRID–CART	87,97	92,69	83,24	0,8797
CHRID–CHRID	75,28	79,19	71,37	0,7528
CHRID–BBN	81,01	81,88	80,13	0,8101
CHRID–SVM	79,05	79,87	78,23	0,7905
CHRID–ANN	82,40	83,20	81,59	0,8240

Ένα σύνολο δεδομένων 176 εταιρειών της Ταιβάν από τις οποίες οι 44 έχουν παραποιημένες οικονομικές καταστάσεις και οι 132 δεν έχουν, είναι το σύνολο των δεδομένων της έρευνας αυτής. Η ιδιαιτερότητα της εν λόγω έρευνας είναι ότι χρησιμοποιεί και δύο αλγόριθμους από την τεχνική του δέντρου απόφασης για να επιλέξει τις μεταβλητές που θα χρησιμοποιηθούν ως δεδομένα εισόδου. Δεν θα μπορούμε σε περαιτέρω λεπτομέρειες γι' αυτήν την διαδικασία καθώς το ζήτημα γίνεται πολύ τεχνικό και δεν έχει καμία σημασία για την εν λόγω εργασία.

6.

<p>Özdağoğlu, G., Özdağoğlu, A., Gümüş, Y., & Kurt Gümüş, G. (2017). The application of data mining techniques in manipulated financial statement classification: The case of turkey. <i>Journal of AI and Data Mining</i>, 5(1), 67–77.</p>				
Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	AUC
DT	82,24%	92,73%	71,15%	0,8194
LR	82,24%	92,73%	71,15%	0,8194
ANN	83,18%	90,00%	75,96%	0,8298

Αυτή η έρευνα αναλύει ένα σύνολο δεδομένων 214 τούρκικων εταιρειών από τις οποίες οι 110 έχουν είτε αρνητική γνώμη είτε γνώμη με επιφύλαξη από τους ελεγκτές. Οι υπόλοιπες 114 έχουν θετική γνώμη από τους ελεγκτές. Η χρονική περίοδος που καλύπτεται είναι από το 2009 έως το 2013 και τα δεδομένα συλλέχθηκαν από την πλατφόρμα Finnet.

7.

CHEN, S., & YANG, A. (2018). An Effective Financial Statements Fraud Detection Model. <i>DEStech Transactions on Engineering and Technology Research</i> , pmsms, 8–11.				
Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	AUC
ANN-CHAID	93,41%	95,45%	87,27%	0,9136
CART-SVM	91,82%	94,55%	83,64%	0,8909
CART-CHAID	91,59%	95,45%	80,00%	0,8773
ANN-ANN	90,91%	92,73%	85,45%	0,8909
ANN-SVM	90,68%	93,64%	81,82%	0,8773
CART-ANN	88,07%	91,36%	78,18%	0,8477

Το σύνολο δεδομένων αυτής της έρευνας είναι 220 εταιρείες της Ταϊβάν όπου οι 55 έχουν παραποιημένες οικονομικές καταστάσεις και οι 165 δεν έχουν. Επίσης, πέρα από τις τεχνικές που χρησιμοποιεί για να ταξινομήσει τα οικονομικές καταστάσεις, χρησιμοποιεί και δύο αλγόριθμους για επιλογή μεταβλητών ως δεδομένα εισόδου. Οι αλγόριθμοι αυτοί είναι το τεχνητό νευρωνικό δίκτυο και ο αλγόριθμος CART που είναι δέντρο απόφασης.

8.

Mohammadi, M., Yazdani, S., Khanmohammadi, M. H., & Maham, K. (2020). Financial Reporting Fraud Detection: An Analysis of Data Mining Algorithms. <i>International Journal of Finance & Managerial Accounting</i> , 4(16), 1–12.				
Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	AUC
BN	65,80%	69,70%	63,00%	0,6635
DA	62,00%	67,00%	58,00%	0,625
LR	67,90%	69,10%	66,10%	0,676
ANN	75,00%	69,80%	80,20%	0,75
SVM	67,00%	67,50%	66,70%	0,671

Αυτή η έρευνα αναλύει τα δεδομένα 330 εισηγμένων Ιρανικών εταιρειών από τις οποίες οι μισές έχουν παραπονημένες οικονομικές καταστάσεις και οι άλλες μισές δεν έχουν. Το χρονικό διάστημα που καλύπτει είναι από το 2001 έως το 2016.

4.2 ΤΑΞΙΝΟΜΗΣΗ ΚΑΤΑ ΑΚΡΙΒΕΙΑ ΤΕΧΝΙΚΗΣ

Στην συνέχεια ακολουθεί ένας πίνακας που παρουσιάζει συνολικά όλες τις τεχνικές των ερευνών που αναφέρθηκαν προηγουμένως.

Πίνακας 6:

Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	AUC
ANN	92,90%	82,93%	96,09%	0,8951
ANN	77,98%	78,00%	63,70%	0,8300
ANN	83,18%	90,00%	75,96%	0,8298
ANN	75,00%	69,80%	80,20%	0,75
ANN-ANN	90,91%	92,73%	85,45%	0,8909
ANN-CHAID	93,41%	95,45%	87,27%	0,9136
ANN-SVM	90,68%	93,64%	81,82%	0,8773
BBN	72,62%	72,60%	68,90%	0,7860
BN	65,80%	69,70%	63,00%	0,6635
CART-ANN	88,07%	91,36%	78,18%	0,8477
CART-ANN	75,00%	75,84%	74,16%	0,7500
CART-BBN	75,20%	77,18%	73,22%	0,7520
CART-CHAID	91,59%	95,45%	80,00%	0,8773
CART-DT(CART)	83,19%	88,59%	77,78%	0,8319
CART-DT(CHAID)	80,70%	81,88%	79,51%	0,8070
CART-SVM	91,82%	94,55%	83,64%	0,8909
CART-SVM	74,68%	75,17%	74,19%	0,7468
CHAID-ANN	82,40%	83,20%	81,59%	0,8240
CHAID-BBN	81,01%	81,88%	80,13%	0,8101
CHAID-CART	87,97%	92,69%	83,24%	0,8797
CHAID-CHAID	75,28%	79,19%	71,37%	0,7528
CHAID-SVM	79,05%	79,87%	78,23%	0,7905
DA	62,00%	67,00%	58,00%	0,625
DT	94,74%	86,20%	97,70%	0,9195

DT	90,32%	81,58%	93,16%	0,8737
DT	77,76%	77,80%	68,80%	0,8180
DT	82,24%	92,73%	71,15%	0,8194
GMDH	90,68%	93,46%	88,34%	0,9090
GP	92,68%	90,55%	95,27%	0,9291
GP	87,72%	53,00%	99,20%	0,7610
LR	70,36%	62,91%	78,88%	0,7090
LR	88,48%	71,79%	93,65%	0,8272
LR	82,24%	92,73%	71,15%	0,8194
LR	67,90%	69,10%	66,10%	0,676
MLFF	78,77%	76,98%	81,28%	0,7913
NB	87,72%	84,00%	92,90%	0,8845
NB	60,21%	60,20%	68,90%	0,6930
PNN	95,64%	91,27%	94,16%	0,9272
SVM	73,41%	72,07%	75,04%	0,7356
SVM	65,60%	65,60%	66,70%	0,6620
SVM	67,00%	67,50%	66,70%	0,671

Οι πίνακες που ακολουθούν παρουσιάζουν τις τεχνικές εξόρυξης δεδομένων, των ερευνών που αναλύει η παρούσα εργασία, ανά τεχνική εξόρυξης. Δηλαδή κάθε πίνακας εμφανίζει τα αποτελέσματα των ερευνών ανά τεχνική. Επίσης έχει γίνει ο υπολογισμός του μέσου όρου της ακρίβειας που παρουσιάζουν οι τεχνικές αυτές.

Πίνακας 7.1:

Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	Μ. Ο. Ακρίβειας
ANN	92,90%	82,93%	96,09%	83,18%
ANN	77,98%	78,00%	63,70%	
ANN	83,18%	90,00%	75,96%	
ANN	75,00%	69,80%	80,20%	
ANN-ANN	90,91%	92,73%	85,45%	
CART-ANN	88,07%	91,36%	78,18%	
CART-ANN	75,00%	75,84	74,16	
CHAID-ANN	82,40%	83,20	81,59	

Πίνακας 7.2:

Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	Μ. Ο. Ακρίβειας
BBN	72,62%	72,60%	68,90%	76,28%
CART-BBN	75,20%	77,18%	73,22%	
CHAID-BBN	81,01%	81,88%	80,13%	

Πίνακας 7.3:

Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	Μ. Ο. Ακρίβειας
BN	65,80%	69,70%	63,00%	71,24%
BN	87,72%	84,00%	92,90%	
BN	60,21%	60,20%	68,90%	

Πίνακας 7.4:

Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	Μ. Ο. Ακρίβειας
CART-CHAID	91,59%	95,45%	80,00%	85,72%
CART-DT(CART)	83,19%	88,59%	77,78%	
CART-DT(CHAID)	80,70%	81,88%	79,51%	
CHAID-CART	87,97%	92,69%	83,24%	
CHAID-CHAID	75,28%	79,19%	71,37%	
ANN-CHAID	93,41%	95,45%	87,27%	
DT	94,74%	86,20%	97,70%	
DT	90,32%	81,58%	93,16%	
DT	77,76%	77,80%	68,80%	
DT	82,24%	92,73%	71,15%	

Πίνακας 7.5:

Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	Μ. Ο. Ακρίβειας
DA	62,00%	67,00%	58,00%	62,00%

Πίνακας 7.6:

Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	Μ. Ο. Ακρίβειας
GMDH	90,68%	93,46%	88,34%	90,68%

Πίνακας 7.7:

Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	Μ. Ο. Ακρίβειας
GP	92,68%	90,55%	95,27%	90,20%
GP	87,72%	53,00%	99,20%	

Πίνακας 7.8:

Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	Μ. Ο. Ακρίβειας
LR	70,36%	62,91%	78,88%	77,25%
LR	88,48%	71,79%	93,65%	
LR	82,24%	92,73%	71,15%	
LR	67,90%	69,10%	66,10%	

Πίνακας 7.9:

Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	Μ. Ο. Ακρίβειας
MLFF	78,77%	76,98%	81,28%	78,77%

Πίνακας 7.10:

Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	Μ. Ο. Ακρίβειας
PNN	95,64%	91,27%	94,16%	95,64%

Πίνακας 7.11:

Τεχνική εξόρυξης δεδομένων	Συνολική ακρίβεια	Ευαισθησία	Ειδικότητα	Μ. Ο. Ακρίβειας
SVM	73,41%	72,07%	75,04%	77,46%
SVM	65,60%	65,60%	66,70%	
SVM	67,00%	67,50%	66,70%	
ANN-SVM	90,68%	93,64%	81,82%	
CART-SVM	91,82%	94,55%	83,64%	
CART-SVM	74,68%	75,17%	74,19%	
CHAID-SVM	79,05%	79,87%	78,23%	

Ο μέσος όρος δεν είναι η σταθμισμένη μέση τιμή αλλά ο μέσος όρος της συνολικής ακρίβειας των τεχνικών προς τον αριθμό των ερευνών που εμφανίζονται. Με αυτή την προσέγγιση ακολουθεί ένας πίνακας που παρουσιάζει τις τεχνικές ταξινομημένες με βάση τον μέσο όρο όπως αναλύθηκε.

Επίσης στον πίνακα που ακολουθεί εμφανίζεται και ο αριθμός των ερευνών που χρησιμοποίησαν την κάθε τεχνική για να ταξινομήσουν τις οικονομικές καταστάσεις σε παραποιημένες ή μη.

Πίνακας 8:

Τεχνική εξόρυξης δεδομένων	Μ. Ο. ακρίβειας	Αριθμός εργασιών
Πιθανοτικά Νευρωνικά Δίκτυο	95,64%	1
Μέθοδος Ομάδων Διαχείρισης Δεδομένων	90,68%	1
Γενετικός Προγραμματισμός	90,20%	2
Δέντρα Απόφασης	85,72%	6
Τεχνητά Νευρωνικά Δίκτυα	83,18%	6
Πολυεπίπεδα Νευρωνικά Δίκτυο	78,77%	1
Μηχανή Διανυσμάτων Υποστήριξης	77,46%	5
Λογιστική Παλινδρόμηση	77,25%	4
Δίκτυα Πεποιθήσεων Bayes	76,28%	2
Naïve Bayes	71,24%	3
Διακριτική Ανάλυση	62,00%	1

Παρατηρούμε ότι 4 από τις 11 τεχνικές χρησιμοποιούνται μόνο σε μία έρευνα. Δύο τεχνικές χρησιμοποιούνται σε δύο εργασίες. Ενώ σε τρεις έρευνες χρησιμοποιείται η τεχνική Naïve Bayes και σε 4 η λογιστική παλινδρόμηση. Πιο συχνά χρησιμοποιούμενη τεχνική εξόρυξης δεδομένων είναι η Μηχανή Διανυσμάτων Υποστήριξης καθώς χρησιμοποιείται σε 5 έρευνες. Αντίστοιχα οι περισσότερο χρησιμοποιούμενες από όλες είναι τα Δέντρα Απόφασης και τα Τεχνητά Νευρωνικά Δίκτυα που εμφανίζονται να χρησιμοποιούνται από τους ερευνητές σε 6 διαφορετικές έρευνες.

Ο ακόλουθος πίνακας παρουσιάζει τις τεχνικές εξόρυξης δεδομένων ταξινομημένες κατά τον σταθμισμένο μέσο όρο της συνολικής ακρίβειας. Ο μαθηματικός τύπος του σταθμισμένου μέσου όρου είναι:

$$\bar{x} = \frac{\sum_{i=1}^n x_i p_i}{\sum_{i=1}^n p_i} = \frac{x_1 p_1 + x_2 p_2 + \dots + x_n p_n}{p_1 + p_2 + \dots + p_n}$$

Για τον υπολογισμό χρησιμοποιήθηκε η εφαρμογή από την ιστοσελίδα

<https://www.calculat.org/gr>.

Ως τιμή X_i δόθηκε η συνολική ακρίβεια της τεχνικής στην κάθε εργασία που χρησιμοποιήθηκε, ενώ ως τιμή P_i δόθηκε ο συνολικός αριθμός του δείγματος δεδομένων της κάθε έρευνας.

Πίνακας 9:

Τεχνική εξόρυξης δεδομένων	Σταθμικός Μ. Ο.	Αριθμός εργασιών
Πιθανοτικά Νευρωνικά Δίκτυο	95,64%	1
Γενετικός Προγραμματισμός	90,89%	2
Μέθοδος Ομάδων Διαχείρισης Δεδομένων	90,68%	1
Πολυεπίπεδα Νευρωνικά Δίκτυο	78,77%	1
Τεχνητά Νευρωνικά Δίκτυα	78,09%	6
Δέντρα Απόφασης	77,97%	6
Λογιστική Παλινδρόμηση	75,01%	4
Δίκτυα Πεποιθήσεων Bayes	72,65%	2
Μηχανή Διανυσμάτων Υποστήριξης	65,86%	5
Διακριτική Ανάλυση	62,00%	1
Naïve Bayes	60,29%	3

ΚΕΦΑΛΑΙΟ 5: ΣΥΜΠΕΡΑΣΜΑΤΑ

Στην παρούσα εργασία έγινε μια συγκριτική ανάλυση συγκεκριμένων ερευνών που ασχολούνται με τον εντοπισμό παραπονημένων οικονομικών καταστάσεων με τεχνικές εξόρυξης δεδομένων. Οι έρευνες αυτές δημοσιεύτηκαν την τελευταία δεκαετία. Το σύνολο δεδομένων που χρησιμοποιήσαν λαμβάνει χώρα από τις αρχές τις προηγούμενης δεκαετίας μέχρι και τα μέσα της τρέχουσας δεκαετίας. Λόγω και της παγκόσμιας οικονομικής κρίσης από το 2008 έως κι το 2015, το δείγμα αυτό των ερευνών αναλύει δεδομένα της εν λόγω περιόδου κι έτσι, εύστοχα, μπορεί να παρατηρηθεί η συμπεριφορά των εταιρειών σε μια ιδιαίτερα σημαντική και δύσκολη περίοδο για την οικονομία παγκοσμίως.

Από την ανάλυση που έγινε στα προηγούμενα κεφάλαια της εν λόγω εργασίας, παρατηρούμε ότι οι τεχνικές εξόρυξης δεδομένων Τεχνητά Νευρωνικά Δίκτυα, Δέντρα Απόφασης, Μηχανή Διανυσμάτων Υποστήριξης και Λογιστική Παλινδρόμηση είναι οι πιο ευρέως εφαρμοσμένες σε έρευνες για τον εντοπισμό απάτης στις οικονομικές καταστάσεις. Αυτό φαντάζει φυσιολογικό καθώς και οι τέσσερις αυτές τεχνικές είναι πολύ διαδεδομένες σε όλα τα επιστημονικά πεδία και έχουν μεγάλο φάσμα εφαρμογών. Η ακρίβεια που παρουσιάζουν στην σωστή κατηγοριοποίηση των παραπονημένων οικονομικών καταστάσεων και των μη παραπονημένων είναι ιδιαίτερα υψηλή. Κυμαίνεται από το χαμηλό 65,86% που παρουσιάζει η Μηχανή Διανυσμάτων Υποστήριξης και φτάνει έως το υψηλό 78,09% που παρουσιάζουν τα Τεχνητά Νευρωνικά Δίκτυα. Βέβαια οι τρεις από τις τέσσερις αυτές τεχνικές έχουν ποσοστό ακρίβειας μεταξύ 75% και 78,09%. Αυτό σημαίνει ότι σε ένα σημαντικό αντιπροσωπευτικό δείγμα οικονομικών καταστάσεων, οι τεχνικές εξόρυξης δεδομένων που επιτυγχάνουν να κατηγοριοποιήσουν σωστά περισσότερες από 3 στις 4 οικονομικές καταστάσεις είναι τα Τεχνητά Νευρωνικά Δίκτυα, το Δέντρα Απόφασης και Λογιστική Παλινδρόμηση.

Δεν πρέπει να παραλείψουμε να αναφέρουμε ότι ιδιαίτερα υψηλό ποσοστό ακρίβειας πέτυχε το Πιθανοτικό Νευρωνικό Δίκτυο. Σε μια έρευνα που χρησιμοποιήθηκε κατάφερε να έχει ακρίβεια της τάξεως του 95,64%. Αντίστοιχα σημαντικά αποτελέσματα πέτυχαν και οι τεχνικές εξόρυξης δεδομένων Γενετικός Προγραμματισμός και Μέθοδος Ομάδων Διαχείρισης Δεδομένων, με 90,89% και 90,68% ακρίβεια στον εντοπισμό παραπονημένων οικονομικών καταστάσεων

αντίστοιχα. Αυτό σημαίνει ότι οι τρεις τεχνικές εξόρυξης δεδομένων που κατάφεραν να κατηγοριοποιήσουν σωστά περισσότερες από τις 9 στις 10 οικονομικές καταστάσεις είναι τα Πιθανοτικά Νευρωνικά Δίκτυα, ο Γενετικός Προγραμματισμός και η Μέθοδος Ομάδων Διαχείρισης Δεδομένων.

Η τεχνική Πολυεπίπεδα Νευρωνικά Δίκτυα χρησιμοποιήθηκε σε μία μόνο έρευνα και πέτυχε να κατηγοριοποιήσει σωστά τις οικονομικές καταστάσεις σε ποσοστό 78,77%. Η τεχνική της Διακριτική Ανάλυσης χρησιμοποιήθηκε από μία μόνο έρευνα και δεν μπόρεσε να ξεπεράσει το 62%. Τέλος τα Δίκτυα Πεποιθήσεων κατά Bayes και τα Μπεϋνσιανά δίκτυα, ενώ χρησιμοποιήθηκαν σε δύο και τρεις έρευνες αντίστοιχα πέτυχαν συνολική ακρίβεια 72,65% και 60,29%. Τα Μπεϋνσιανά δίκτυα πέτυχαν το χαμηλότερο ποσοστό ακρίβειας από όλες τις τεχνικές στο σύνολο των ερευνών που αναλύει η παρούσα εργασία.

Συνολικά μπορούμε να πούμε ότι αυτές οι τεχνικές εξόρυξης δεδομένων μπορούν να κατηγοριοποιήσουν σωστά τις οικονομικές καταστάσεις τουλάχιστον 6 στις 10 περιπτώσεις.

Συμπερασματικά μπορούμε να πούμε ότι οι τεχνικές εξόρυξης δεδομένων που μπορούν να δώσουν μια πραγματική βοήθεια στην εργασία των ελεγκτών για τον εντοπισμό απάτης στις οικονομικές καταστάσεις είναι τα δέντρα απόφασης και οι υποκατηγορίες των νευρωνικών δικτύων. Αυτό παρατηρείται καθώς βλέπουμε ότι τα Δέντρα απόφασης είναι τα περισσότερο χρησιμοποιούμενα στην διαδικασία του εντοπισμού απάτης και μπορούν να εντοπίσουν σωστά σχεδόν 8 στις 10 οικονομικές καταστάσεις. Αντίστοιχα οι υποκατηγορίες των νευρωνικών δικτύων, που είναι τα Τεχνητά Νευρωνικά Δίκτυα, τα Πολυεπίπεδα Νευρωνικά Δίκτυα και τα Πιθανοτικά Νευρωνικά Δίκτυα, έχουν εξίσου σημαντική παρουσία στην διερεύνηση του εντοπισμού απάτης στις οικονομικές καταστάσεις και συνολικά μπορούν να κατηγοριοποιήσουν πάνω από 8 στις 10 οικονομικές καταστάσεις σωστά.

Αυτό που μπορεί να καταγραφεί με ακρίβεια είναι ότι οι τεχνικές εξόρυξης δεδομένων μπορούν να δώσουν μια πολύ σημαντική βοήθεια στην εργασία των ελεγκτών και κυρίως τα δέντρα απόφασης και τα νευρωνικά δίκτυα. Σαφώς οι τεχνικές αυτές δεν μπορούν σε καμία περίπτωση να υποκαταστήσουν την εργασία του ελεγκτή, αλλά σίγουρα μπορούν να του παρέχουν ένα εργαλείο όπου θα μπορεί να κάνει τον έλεγχο πιο δομημένο και να περιορίσει τον χρόνο που δαπανάται για τον έλεγχο των

οικονομικών καταστάσεων. Χωρίς να μειωθεί η ποιότητα των αποτελεσμάτων του ελέγχου.

ΕΠΙΛΟΓΟΣ

ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ

Η εξόρυξη δεδομένων μπορεί να συνδράμει στο έργο των ελεγκτών και σε άλλα στάδια της ελεγκτικής διαδικασίας πέρα από τον έλεγχο των οικονομικών καταστάσεων. Η βιβλιογραφία πλέον είναι αρκετά εμπλουτισμένη με έρευνες για τον έλεγχο των οικονομικών καταστάσεων και φυσικά η έρευνα αυτή συνεχίζεται. Θα μπορούσε στο μέλλον να γίνει διερεύνηση στο κατά πόσο οι τεχνικές εξόρυξης δεδομένων θα μπορούσαν να βοηθήσουν στην συλλογή των ελεγκτικών τεκμηρίων, ούτως ώστε να μειωθεί στο ελάχιστο η πιθανότητα ο έλεγχος να βγάλει εσφαλμένα συμπεράσματα. Επίσης η εξόρυξη δεδομένων θα μπορούσε να έχει λόγο και για την τελική γνώμη του ελεγκτή ανάλογα με τα ελεγκτικά τεκμήρια, τον έλεγχο των οικονομικών καταστάσεων και γενικά την διαδικασία του ελέγχου. Έτσι ώστε ο ελεγκτής να έχει ένα εργαλείο καθ' όλη την διάρκεια του ελέγχου που θα μπορεί να συμβουλευέται και να παίρνει αποφάσεις γρηγορότερα.

ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

1. Μπατσινίλας, Ε., & Πατατούκας, Κ. (2012). Σύγχρονη Λογιστική (Σύμφωνα με τα Ελληνικά Λογιστικά Πρότυπα και με αναφορά κατά θέμα στα Διεθνή Λογιστικά Πρότυπα) Τόμος Α'. Αθήνα : Σταμούλη Α.Ε.
2. Νούλας, Α. Γ. (2015). Χρηματοοικονομική Διοίκηση. Επενδυτικές και Χρηματοδοτικές Αποφάσεις. Θεσσαλονίκη.
3. Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data Mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications*, 32(4), 995–1003. <https://doi.org/10.1016/j.eswa.2006.02.016>
4. Λουμιώτης, Β. Ι., & Τζίφας, Β. Ν. (2012). Βασικές οδηγίες εφαρμογής Διεθνών Προτύπων Ελέγχου. Εκδόσεις Σταμούλη.
5. Μενεξές, Γ., & Οικονόμου, Α. (2002). Σφάλματα και Παρανοήσεις στους Στατιστικούς Ελέγχους Υποθέσεων. Υπέρβαση μέσω της Ανάλυσης Δεδομένων. *Τετράδια Ανάλυσης Δεδομένων-Data Analysis Bulletin*, 2, 52-64.
6. Σπαθή, Χ. (2002). Εμπειρική διερεύνηση της παραποίησης οικονομικών καταστάσεων (Vol. 570).
7. Κύρκος Ε. (2015). Επιχειρηματική ευφυΐα και εξόρυξη δεδομένων. Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών.
8. [ΕΛΠ 2014]: Ελληνικά Λογιστικά Πρότυπα, συναφείς ρυθμίσεις και άλλες διατάξεις, 2014, Νόμος Υπ' αριθμό. 4308
9. Tan, P.-n., Steinbach, M., & Kumar, V. (2015). Εισαγωγή στην Εξόρυξη Δεδομένων. (Β. Σ. Βερούκιος , Επιμ., & Σ. Σουραβλάς, Μεταφρ.) Θεσσαλονίκη: Εκδόσεις Τζιόλα.
10. Agarwal, J., & Malloy, D. C. (1999). Ethical work climate dimensions in a not-for-profit organization: An empirical study. *Journal of Business Ethics*, 20(1), 1–14. <https://doi.org/10.1023/A:1005974718602>
11. Beccar-Varela, M. P., Mariani, M. C., Tweneboah, O. K., & Florescu, I. (2017). Analysis of the Lehman Brothers collapse and the Flash Crash event by applying wavelets methodologies. *Physica A: Statistical Mechanics and Its Applications*, 474, 162–171. <https://doi.org/10.1016/j.physa.2017.01.064>
12. Buchanan, B., & Yang, T. (2005). The benefits and costs of controlling shareholders: The rise and fall of Parmalat. *Research in International Business and Finance*,

- 19(1), 27–52. <https://doi.org/10.1016/j.ribaf.2004.10.002>
13. Chen, S. (2016). Detection of fraudulent financial statements using the hybrid data mining approach. *SpringerPlus*, 5(1), 1–16. <https://doi.org/10.1186/s40064-016-1707-6>
 14. CHEN, S., & YANG, A. (2018). An Effective Financial Statements Fraud Detection Model. *DEStech Transactions on Engineering and Technology Research*, pmsms, 8–11. <https://doi.org/10.12783/dtetr/pmsms2018/24902>
 15. Cihan Sorkun, M., & Toraman, T. (2017). International Journal of Intelligent Systems and Applications in Engineering Fraud Detection on Financial Statements Using Data Mining Techniques. 5(3). <https://doi.org/10.1039/b000000x>
 16. Cressey, D. R. (1953). Other people's money; a study of the social psychology of embezzlement. In *Other people's money; a study of the social psychology of embezzlement*. Free Press.
 17. Dutta, I., Dutta, S., & Raahemi, B. (2017). Detecting financial restatements using data mining techniques. *Expert Systems with Applications*, 90, 374–393. <https://doi.org/10.1016/j.eswa.2017.08.030>
 18. Gupta, R., & Singh Gill, N. (2012). A Data Mining Framework for Prevention and Detection of Financial Statement Fraud. *International Journal of Computer Applications*, 50(8), 7–14. <https://doi.org/10.5120/7789-0889>
 19. Lin, C. C., Chiu, A. A., Huang, S. Y., & Yen, D. C. (2015). Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments. *Knowledge-Based Systems*, 89, 459–470. <https://doi.org/10.1016/j.knosys.2015.08.011>
 20. Maka, K., Pazhanirajan, S., & Mallapur, S. (2020). Selection of most significant variables to detect fraud in financial statements. *Materials Today: Proceedings*. <https://doi.org/10.1016/j.matpr.2020.09.613>
 21. Mohammadi, M., Yazdani, S., Khanmohammadi, M. H., & Maham, K. (2020). Financial Reporting Fraud Detection: An Analysis of Data Mining Algorithms. *International Journal of Finance & Managerial Accounting*, 4(16), 1–12.
 22. Munawer, Z., Yahya, S., & Siti-Nabiha, A. K. (2012). Sell-side Security Analysts: Re-reporting of Enron Corporation Fraudulent Financial Data. *Procedia - Social and Behavioral Sciences*, 62, 749–760. <https://doi.org/10.1016/j.sbspro.2012.09.127>

23. Özdağoğlu, G., Özdağoğlu, A., Gümüş, Y., & Kurt-Gümüş, G. (2017). The application of data mining techniques in manipulated financial statement classification: The case of turkey. *Journal of AI and Data Mining*, 5(1), 67–77. <https://doi.org/10.22044/jadm.2016.664>
24. Ravisankar, P., Ravi, V., Raghava Rao, G., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support Systems*, 50(2), 491–500. <https://doi.org/10.1016/j.dss.2010.11.006>
25. Reinstein, A., & McMillan, J. J. (2004). The Enron debacle: More than a perfect storm. *Critical Perspectives on Accounting*, 15(6–7), 955–970. <https://doi.org/10.1016/j.cpa.2003.08.006>
26. Santhosh, R., & Mohanapriya, M. (2020). Generalized fuzzy logic based performance prediction in data mining. *Materials Today: Proceedings*. <https://doi.org/10.1016/j.matpr.2020.08.626>
27. Suh, J. B., & Shim, H. S. (2020). The effect of ethical corporate culture on anti-fraud strategies in South Korean financial companies: Mediation of whistleblowing and a sectoral comparison approach in depository institutions. *International Journal of Law, Crime and Justice*, 60, 100361. <https://doi.org/10.1016/j.ijlcrj.2019.100361>
28. [ACFE]: Association of Certified Fraud Examiners, Retrieved data October 3rd, 2020, <http://www.acfe.com/fraud-101.aspx>
29. [IIA 2016]: International Standards for the Professional Practice of Internal Auditing (Standards), Retrieved October 3rd, 2020 from [https://na.theiia.org/standards-guidance/Public%20Documents/IPPF Standards-2017.pdf](https://na.theiia.org/standards-guidance/Public%20Documents/IPPF%20Standards-2017.pdf)