

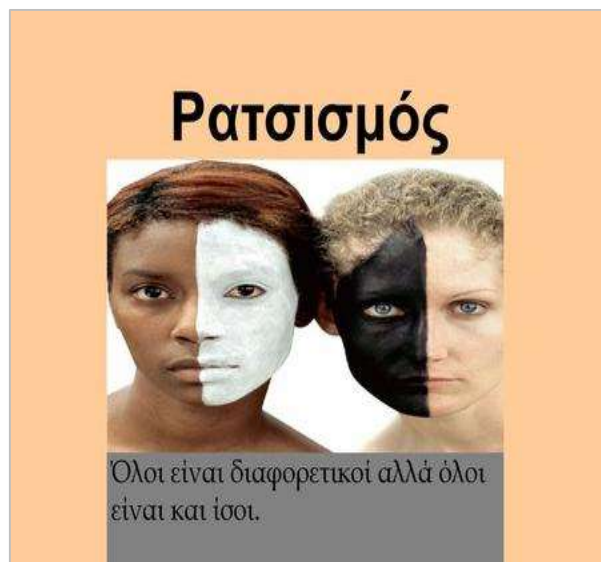
ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ



Πανεπιστήμιο
Ιωαννίνων

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΔΙΕΡΕΥΝΗΣΗ ΤΩΝ ΑΙΤΙΩΝ ΠΟΥ ΕΥΝΟΟΥΝ ΤΗΝ ΑΝΑΠΤΥΞΗ
ΦΑΙΝΟΜΕΝΩΝ ΕΚΦΟΒΙΜΟΥ ΣΕ ΣΧΟΛΙΚΕΣ ΤΑΞΕΙΣ ΜΕ ΧΡΗΣΗ
ΕΥΦΥΩΝ ΜΕΘΟΔΩΝ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ



Καργούδη Σταματία Α.Μ. 11988

Επιβλέπων καθηγητής
Νικόλαος Γιαννακας



Πανεπιστήμιο
Ιωαννίνων

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ Τ.Ε.

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΔΙΕΡΕΥΝΗΣΗ ΤΩΝ ΑΙΤΙΩΝ ΠΟΥ ΕΥΝΟΥΝ ΤΗΝ ΑΝΑΠΤΥΞΗ
ΦΑΙΝΟΜΕΝΩΝ ΕΚΦΟΒΙΜΟΥ ΣΕ ΣΧΟΛΙΚΕΣ ΤΑΞΕΙΣ ΜΕ ΧΡΗΣΗ
ΕΥΦΥΩΝ ΜΕΘΟΔΩΝ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ

Καργούδη Σταματία Α.Μ. 11988

Επιβλέπων Καθηγητής
Νικόλαος Γιαννακέας

- Άρτα 2019 -

INVESTIGATION OF CAUSES THAT LEAD TO THE
DEVELOPMENT OF INTIMIDATING BEHAVIOYR IN SCHOOL
USING MACHINE LEARNIG TECHNIQUES

Εγκρίθηκε από τριμελή εξεταστική επιτροπή

Τόπος, Ημερομηνία

ΕΠΙΤΡΟΠΗ ΑΞΙΟΛΟΓΗΣΗΣ

1. Επιβλέπων καθηγητής

Νικόλαος Γιαννακέας,

Πανεπιστημιακός Υπότροφος

2. Μέλος επιτροπής

Αλέξανδρος Τζάλλας,

Καθηγητής Εφαρμογών

3. Μέλος επιτροπής

Μάρκος Τσίπουρας,

Πανεπιστημιακός Υπότροφος

Ο/Η Προϊστάμενος/η του Τμήματος

Ευριπίδης Γλαβάς,

Καθηγητής

© Επίθετο, Όνομα, έτος.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Δήλωση μη λογοκλοπής

Δηλώνω υπεύθυνα και γνωρίζοντας τις κυρώσεις του Ν. 2121/1993 περί Πνευματικής Ιδιοκτησίας, ότι η παρούσα πτυχιακή εργασία είναι εξ ολοκλήρου αποτέλεσμα δικής μου ερευνητικής εργασίας, δεν αποτελεί προϊόν αντιγραφής ούτε προέρχεται από ανάθεση σε τρίτους. Όλες οι πηγές που χρησιμοποιήθηκαν (κάθε είδους, μορφής και προέλευσης) για τη συγγραφή της περιλαμβάνονται στη βιβλιογραφία.

Καργούδη, Σταματία

Υπογραφή

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω θερμά κ. Γιαννακέα Νικόλαο για την εποικοδομητική βοήθεια του και τη καθοδήγηση του καθ' όλη τη διάρκεια της εκπόνησης της εργασίας.

Περίληψη

Αν και ο σχολικός εκφοβισμός δεν αποτελεί ένα νέο φαινόμενο των τελευταίων χρόνων, απασχολεί έντονα την εκπαιδευτική κοινότητα και παρατηρείται ραγδαία η αύξηση του σε παγκόσμια κλίμακα. Είναι κοινός τόπος ότι οι επιπτώσεις του είναι ιδιαίτερος επικίνδυνες για το θύματα αυτού του φαινομένου. Σκοπός της παρούσας πτυχιακής εργασίας είναι να διερευνήσουμε τα διάφορα είδη του σχολικού εκφοβισμού, τα αίτια και τους παράγοντες που εντείνουν το φαινόμενο αυτό, αλλά και τις επιπτώσεις που έχουν στα θύματα και γενικότερα στο κοινωνικό σύνολο. Επιπλέον στόχος μας είναι να εξετάσουμε ποια είναι η στάση των εκπαιδευτικών στα διαφορά περιστατικά, αλλά και ποιος ο τρόπος αντιμετώπισης που επιλέγει ο κάθε εκπαιδευτικός. Καθώς ο εκφοβισμός λαμβάνει χώρα στο σχολικό περιβάλλον πρώτοι από όλους οι εκπαιδευτικοί είναι αυτοί που είναι υπεύθυνοι για τον εντοπισμό αλλά και την εξάλειψη του φαινομένου. Κάθε περιστατικό θα πρέπει να αντιμετωπίζεται με την ανάλογη σοβαρότητα και προσοχή. Επίσης ο τρόπος αντιμετώπισης των περιστατικών σχολικού εκφοβισμού είναι ιδιαίτερος σημαντικός διότι κάποιες στρατηγικές δεν κρίνονται και τόσο αποτελεσματικές. Στόχος της εργασίας είναι ο εντοπισμός των σχολικών τμημάτων και των τάξεων, που παρουσιάζουν χαρακτηριστικά ευνοϊκά για την ανάπτυξη φαινομένων σχολικού εκφοβισμού, με χρήση τεχνικών Μηχανικής Μάθησης.

Λέξεις-κλειδιά: σχολικός εκφοβισμός, βία, σωματική, λεκτική, ψυχολογική, θύτης, θύμα, σχολικό περιβάλλον, παρατηρητής, οικογένεια, κοινωνικός αποκλεισμός.

Abstract

Even though school bullying isn't a new phenomenon during the last years is a big issue in the schooling community. It is observed a high increase globally while its results to the victims are highly dangerous. The purpose of this diplomatic study is to deal with the different kinds of school bullying the reasons and the factors of this phenomenon but also the results that has in the victims and generally in the whole community. Moreover, our goal is to examine what is the attitude of teachers towards the various incidents and the way each teacher chooses to deal with each incident, also to suggest ways to detect and eliminate the phenomenon. Bullying takes place in school environment that's why teachers are the first to detect and eliminate the phenomenon. Each incident has to be dealt with extreme caution. The aim of the thesis the application of Machine Learning techniques, to identify school classes, where bullying phenomena can be developed.

Keywords: school bullying, violence, physical, verbal, psychological, victim, school environment, observer, family, social exclusion.

Περιεχόμενα

ΚΕΦΑΛΑΙΟ 1:ΣΧΟΛΙΚΟΣ ΕΚΦΟΒΙΣΜΟΣ

- 1.1 Ορισμός- Ιστορική αναδρομή
- 1.2 Μορφές εκφοβισμού
- 1.3 Παράγοντες Ανάπτυξης του φαινομένου
- 1.4 Τρόποι αντιμετώπισης
 - 1.4.1 Στάση εκπαιδευτικού
 - 1.4.2 Διαχείριση σχολικού
 - 1.4.3 Κοινωνικοεθνική ποικιλότητα

ΚΕΦΑΛΑΙΟ 2:ΕΥΦΥΕΙΣ ΜΕΘΟΔΟΙ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ (Machine Learning)

- 2.1 Γενικά μηχανική μάθηση και data mining
- 2.2 Μέθοδοι ταξινόμησης
 - 2.2.1 K-κοντινότερος γείτονας (KNN=K-nearest neighbor)
 - 2.2.2 Μπευζιανός ταξινομητής (Naïve Bayes classifier)
 - 2.2.3 Δένδρα απόφασης (Decision trees)
 - 2.2.4 Μηχανές Διανυσμάτων υποστήριξης (Support Vector Machine-SVM)

ΚΕΦΑΛΑΙΟ 3:ΤΟ ΛΟΓΙΣΜΙΚΟ WEKA

- 3.1 Εισαγωγή
- 3.2 Το περιβάλλον Weka

ΚΕΦΑΛΑΙΟ 4:ΥΛΟΠΟΙΗΣΗ ΠΕΙΡΑΜΑΤΩΝ

- 4.1 Εξαγωγή χαρακτηριστικών
- 4.2 Ηλεκτρονικό ερωτηματολόγιο
- 4.3 Επιλογή χαρακτηριστικών
- 4.4 Ταξινόμηση
- 4.5 Μέτρα αξιολόγησης

ΚΕΦΑΛΑΙΟ 5:ΑΠΟΤΕΛΕΣΜΑΤΑ

- 5.1 Στατιστικά στοιχεία δειγμάτων
- 5.2 Αποτελέσματα ταξινόμησης
 - 5.2.1 K-κοντινότερος γείτονας (KNN=K-nearest neighbor)
 - 5.2.2 Μπευζιανός ταξινομητής (Naïve Bayes classifier)
 - 5.2.3 Δένδρα απόφασης (Decision trees)

5.2.4 Μηχανές Διανυσμάτων υποστήριξης (Support Vector Machine-SVM)

ΚΕΦΑΛΑΙΟ 5:ΣΥΜΠΕΡΑΣΜΑΤΑ

ΒΙΒΛΙΟΓΡΑΦΙΑ

ΚΕΦΑΛΑΙΟ 1

ΣΧΟΛΙΚΟΣ ΕΚΦΟΒΙΣΜΟΣ

1.1 Ορισμός ιστορική αναδρομή

Ο εκφοβισμός (bullying) είναι η επαναλαμβανόμενη άσκηση επιθετικής συμπεριφοράς σωματικής αλλά κυρίως ψυχολογικής κακοποίησης σε ασθενέστερα άτομα μιας ομάδας ή και μεμονωμένα με σκοπό τον κοινωνικό του αποκλεισμό. Πολλοί πιστεύουν πως το bullying είναι ένα καθαρά σχολικό φαινόμενο, στη πραγματικότητα όμως μπορεί να το συναντήσουμε παντού σε κάθε ηλικία σε κάθε κοινωνική τάξη και κάθε κοινωνικό σύνολο. Όπως για παράδειγμα σε αθλητικούς χώρους, στον εργασιακό χώρο, στο στρατό, στη φυλακή ακόμη και στο οικογενειακό περιβάλλον. Εκατοντάδες παραδείγματα επισημαίνουν πως το bullying ή (εκφοβισμός) αυξάνεται με τρομακτική πρόοδο σε σχολεία αλλά και χώρους εργασίας. Το φαινόμενο του σχολικού εκφοβισμού (bullying) μελετήθηκε για πρώτη φορά το 1978 στη Νορβηγία και εννιά χρόνια μετά το 1987 εμφανίστηκε σε επιστημονικά περιοδικά με τον όρο bullying. Το φαινόμενο του σχολικού εκφοβισμού εμφανίζεται πολύ πριν την δεκαετία του 1970 αν και η πρώτη καταγραφεί του φαινομένου είναι τη δεκαετία αυτή. Είναι μια μορφή κακοποίησης και αποτελεί μια έκφραση βίαιης συμπεριφοράς που κάποιος προσπαθεί να επιβάλλει την εξουσία μέσω επιθετική συμπεριφοράς και το θύμα αναφέρεται ως στόχος. Το bullying έχει να κάνει με την περιφρόνηση ότι ένας άνθρωπος είναι κατώτερος και ανάξιος σεβασμού δεν είναι απλά ένας καβγάς που πρέπει να λυθεί. Η Παγκόσμια Οργάνωση Υγείας αναφέρει το bullying ότι είναι μείζον πρόβλημα δημόσιας υγείας στα σχολεία στους εργασιακούς χώρους γενικότερα στη κοινωνία σε άτομα με διαφορετικότητα ευάλωτα και συνεσταλμένα. Κατά τον Olweus σχολικός εκφοβισμός (bullying) θεωρείτε όταν ένα παιδί εκτίθεται κατ'εξακολούθηση σε ψυχοσυναισθηματική κακοποίηση μέσω λεκτικής ή σωματικής επίθεσης. Υπήρχε πάντα σε κάθε σχολική κοινωνία αλλά πηρέ μεγάλες διαστάσεις από την εποχή της παγκοσμιοποίησης και την αλλαγή του εκπαιδευτικού συστήματος σε πολυπολιτισμικό. Το φαινόμενο αυτό λοιπόν ερευνάται

συστηματικά τα τελευταία χρόνια και οι ειδικοί δίνουν ιδιαίτερη έμφαση στη φύση του προβλήματος ώστε να προταθούν αποτελεσματικές παρεμβάσεις [1].



Εικόνα 1.1: Σχολικός εκφοβισμός

1.2 Μορφές Εκφοβισμού

Ρατσισμός-bullying(εκφοβισμός), αρκεί άραγε το χρώμα του δέρματος, η εθνικότητα, η θρησκεία, το φύλο, ο σεξουαλικός προσανατολισμός και πολλά άλλα να διαχωρίζουν τους ανθρώπους μεταξύ τους. Η σωματική και ψυχολογική κακοποίηση σε άτομα ασθενέστερων ομάδων εμφανίζεται σε όλες τις ηλικίες και όλες τις κοινωνικές ομάδες και με διάφορες μορφές. Ο σχολικός εκφοβισμός ορίζεται ως μια επαναλαμβανόμενη έκθεση ενός μαθητή σε αρνητικές καταστάσεις από ένα ή περισσότερα άτομα. Παρακάτω θα αναλύσουμε τις μορφές με τις οποίες εκδηλώνεται ο σχολικός εκφοβισμός αλλά και τους ρόλους του θύτη και του θύματος. Μερικές βασικές μορφές εκφοβισμού είναι η σωματική η λεκτική (ψυχολογική), η σεξουαλική, η ηλεκτρονική και στο σύνολο μιλάμε για ρατσιστικές συμπεριφορές που έχουν ως σκοπό τον εκφοβισμό. Αυτές οι βασικές μορφές εκδηλώνονται με διάφορες επιθετικές συμπεριφορές [2].

- Η σωματική βία είναι μια από τις πιο συχνές μορφές σχολικού εκφοβισμού που παρατηρείτε στις μέρες μας ανάμεσα σε παιδιά όλων των ηλικιών. Τα χτυπήματα, τα σπρωξίματα και η φθορά ατομικής ιδιοκτησίας του θύματος έχουν ως σκοπό το τραυματισμό αλλά και τη δημιουργία της αίσθησης της απειλής ενός ατόμου. Σε πολλά σχολεία ακόμη και ευρωπαϊκών χωρών ο σωματικός εκφοβισμός είναι ένα συχνό φαινόμενο. Οι τρόποι με του οποίους εκδηλώνεται είναι μπουνιές, κλωτσιές, τσιμπήματα, δαγκωνιές ακόμη και βίαιη συμπεριφορά με σκοπό την αφαίρεση προσωπικών αντικειμένων . Με την εκδήλωση σωματικής βίας ο θύτης αποκτά ένα αίσθημα κυριαρχίας απέναντι στο θύμα. Ο σωματικός εκφοβισμός είναι μια από τις μορφές που μπορεί να γίνει εύκολα αντιληπτός από τους εκπαιδευτικούς του σχολείου ή από τους γονείς από τα εμφανή σημάδια κακοποίησης.
- Ένα ακόμη είδος εκφοβισμού είναι η λεκτική βία η οποία εκδηλώνεται με τη χρήση υβριστικών απειλητικών ή απαξιωτικών εκφράσεων από μια ομάδα ατόμων προς ένα άτομο ή και μεμονωμένα. Στα σχολεία αλλά και στη καθημερινή ζωή ένα άτομο μπορεί να δεχτεί λεκτικό εκφοβισμό ποιο συχνά από ότι σωματικό. Αυτή η μορφή εκφοβισμού εκδηλώνεται χρησιμοποιώντας συνήθως φράσεις που έχουν ως σκοπό να ταπεινώσουν και να εξευτελίσουν το θύμα. Συχνά οι θύτες χρησιμοποιούν παρατσούκλια, απειλές, βρισιές, ειρωνικά και κοροϊδευτικά σχόλια για την εμφάνιση του θύματος αλλά και την οικονομική του κατάσταση. Ο λεκτικός εκφοβισμός μπορεί να επηρεάσει ακόμη και άτομα με αυτοπεποίθηση, προκαλεί αρνητική ψυχολογία στο θύμα και του δημιουργεί ένα αίσθημα κατωτερότητας με απώτερο σκοπό την χειραγωγή του και την απομόνωση του από το κοινωνικό σύνολο του σχολείου και κοινωνικές δραστηριότητες.
- Η μορφή του σεξουαλικού (bullying) είναι ένα είδος εκφοβισμού που εμφανίζεται όλο και ποίο συχνά στην εποχή μας σε ένα σχολικό περιβάλλον. Ο σεξουαλικός εκφοβισμός εκδηλώνεται με διάφορους τρόπους απρεπή αγγίγματα, ανήθικες χειρονομίες κυρίως στα γεννητικά όργανα του θύματος, σεξουαλικά υπονοούμενα ακόμη και προσβλητική και υποτιμητική συμπεριφορά σε περίπτωση που κάποιος έχει διαφορετικό σεξουαλικό προσανατολισμό. Αυτή η μορφή εκφοβισμού μπορεί να φτάσει ακόμη και σε σοβαρές σεξουαλικές επιθέσεις. Όλα τα παραπάνω έχουν ως αποτέλεσμα να

οδηγηθεί το θύμα σε μια σεξουαλική πράξη η οποία δεν είναι επιθυμητή. Τα θύματα είναι συνήθως κορίτσια τα οποία υποκύπτουν γιατί δεν έχουν την ευχέρεια να μιλήσουν γι αυτό το οποίο τους συμβαίνει.

- Η ανεξέλεγκτη χρήση κινητών και ηλεκτρονικών υπολογιστών αλλά και η δημιουργία λογαριασμών σε διάφορα social media από μικρή ηλικία έχει ανοίξει τις πύλες της και σε μια ακόμη μορφή εκφοβισμού το λεγόμενο cyber-bullying το οποίο έχει οδηγήσει πολλά θύματα ακόμη και στην αυτοκτονία. Τα θύματα δέχονται μηνύματα ή και βίντεο με απειλητικό ή υβριστικό περιεχόμενο στο κινητό, μηνύματα σεξουαλικού περιεχομένου, υποτιμητικά σχόλια στο facebook και σε προσωπικούς τους λογαριασμούς. Στόχος του θύτη είναι να βλάψει το θύμα, να το ντροπιάσει να το ταπεινώσει ακόμη και να των οδηγήσει σε ακραίες συμπεριφορές.
- Τελειώνοντας η γενικότερη μορφή εκφοβισμού είναι ο ρατσισμός που εστιάζει κυρίως στην διαφορετικότητα ενός ατόμου είτε λόγο καταγωγή του θύματος, κοινωνικού και οικονομικού του επίπεδου, σεξουαλικού του προσανατολισμού αλλά και της σωματικής του διάπλασης (κινητικά προβλήματα ή προβλήματα λόγο ασθένειας).

Μερικά από τα χαρακτηριστικά των παιδιών (θυτών) που εκδηλώνουν εκφοβιστικές συμπεριφορές είναι η αυταρχική προσωπικότητα, η ασέβεια ακόμη και προς μεγαλύτερους ανθρώπους, η προβολή της σωματικής τους δύναμης και η βίαιη συμπεριφορά τους για την επίλυση οποιουδήποτε προβλήματος αντιμετωπίσουν. Χαρακτηρίζονται από υπερβολική αυτοπεποίθηση και αυτοεκτίμηση και αναζητούν ή ακόμη και προκαλούν διαμάχες για να προβάλουν την δύναμη τους και να επιβληθούν. Οι θύτες-bullies αδιαφορούν και δεν νιώθουν καμία ενοχή για τη άσχημη συμπεριφορά προς τα θύματα τους αντιθέτως νιώθουν ευχαρίστηση και ένα αίσθημα υπεροχής γιατί έχουν την ανάγκη να νιώθουν δυνατοί [3].

Αντιθέτως τα θύματα που βιώνουν τον εκφοβισμό-bullying είναι άτομα με χαμηλή αυτοπεποίθηση και αυτοεκτίμηση άτομα με άγχος και πολλές ανασφάλειες σε σχέση με τα υπόλοιπα σύνολο. Είναι άτομα ευαίσθητα, συνεσταλμένα με κλειστό χαρακτήρα και με λίγους φίλους. Τα άτομα αυτά βιώνουν καθημερινά διάφορες μορφές εκφοβισμού στο χώρο του σχολείου και κυρίως για θέματα που αφορούν την σωματική τους διάπλαση και την εμφάνιση τους. Αυτό έχει ως αποτέλεσμα τα θύματα

να νιώθουν ανασφάλεια και ότι υστερούν σε σχέση με τα υπόλοιπα παιδιά. Σπανίως τα άτομα που πέφτουν θύματα εκφοβισμού θα αμυνθούν για να υπερασπιστούν τον εαυτό τους και να προσπαθήσουν να βρουν το δίκιο τους [4].

Βέβαια εκτός από το ρόλο του θύτη και του θύματος υπάρχει και ο ρόλος του παρατηρητή οι μάρτυρες δηλαδή μιας τέτοιας κατάστασης. Οι παρατηρητές είναι μια σιωπηρή δύναμη που μπορεί ακόμη και να αποτρέψει μια πράξη εκφοβισμού. Πολλά παιδιά έχει τύχει να γίνουν θεατές διαφόρων περιστατικών εκφοβισμού και αυτό πολλές φορές έχει ως αποτέλεσμα από τη μια να δημιουργείτε στα παιδιά ένα αίσθημα φόβου μήπως βιώσουν κι αυτά κάτι αντίστοιχο και γίνουν οι ίδιοι στόχοι αλλά από την άλλη είναι πολύ πιθανόν να επηρεαστούν αρνητικά αντικρίζοντας μια τέτοια επιθετική συμπεριφορά ίσως και να την υιοθετήσουν και από απλοί παρατηρητές να γίνουν και οι ίδιοι θύτες. Για το λόγο αυτό και οι γονείς αλλά περισσότερο οι εκπαιδευτικοί μέσα στο σχολείο θα πρέπει να συμβουλεύουν τα παιδιά παρατηρητές να μην μένουν σιωπηλά αντιθέτως να μιλάνε όταν γίνονται θεατές σε κάποιο περιστατικό εκφοβισμού χωρίς να φοβούνται και να σκέφτονται τι επιπτώσεις μπορεί να υπάρξουν καθώς η πράξη τους αυτή μόνο καλό προσφέρει.



Εικόνα 1.2: Εξάλειψη Σχολικού Εκφοβισμού

1.3 Παράγοντες Ανάπτυξης του Φαινομένου

Στη σημερινή εποχή που τα παιδιά δέχονται τόσα πολλά ερεθίσματα οι αιτιολογικοί παράγοντες ανάπτυξης της ενδοσχολικής βίας και της ανάπτυξης των κρουσμάτων σχολικού εκφοβισμού είναι πολλοί. Δεν μπορεί να θεωρηθεί ως ένα μεμονωμένο περιστατικό. Συχνά τέτοιες συμπεριφορές δηλώνουν βαθύτερα προβληματικές καταστάσεις στη ζωή του θύτη. Σημαντικοί παράγοντες που μπορεί να οδηγήσουν στην εκδήλωση φαινομένων εκφοβισμού-bullying από παιδιά είναι η οικογένεια, το σχολείο αλλά και το κοινωνικό του περιβάλλον [5].

Η οικογένεια έχει τον ποιο σημαντικό ρόλο στην ανάπτυξη ενός παιδιού την κοινωνικοποίησης του τη δημιουργία της προσωπικότητας του με τον καλύτερο τρόπο. Έτσι λοιπόν η οικογένεια έχει σαν πρωταρχικό στόχο την ένταξη του παιδιού στο κοινωνικό σύνολο και να το βοηθήσει να αναπτύξει τις δεξιότητες του με τον καλύτερο τρόπο. Η οικογένεια είναι αυτή που πρέπει να του δημιουργήσει το αίσθημα της κοινωνικότητας αν δεν συμβεί κάτι τέτοιο τότε το παιδί είναι πιθανόν να αποκτήσει αντικοινωνική και επιθετική συμπεριφορά. Μέσα στο οικογενειακό περιβάλλον οι αιτίες που οδηγούν ένα παιδί να έχει τάσεις επιθετικότητας είναι αρχικά το κλίμα που επικρατεί στην οικογένεια δηλαδή εάν υπάρχουν συνεχείς και έντονοι καβγάδες μεταξύ των γονιών με υβριστικό λεξιλόγιο ακόμη και χειροδικία. Μια άλλη αιτία είναι το πόσο αυταρχικοί, επιθετικοί είναι η γονείς προς τα παιδιά ακόμη και στη προσπάθειά τους να επιλύσουν ένα πρόβλημα με επιθετική και βίαιη συμπεριφορά χρησιμοποιώντας ακόμη και τιμωρίες. Ένας ακόμη λόγος φυσικά που επηρεάζει και ενισχύει την επιθετικότητα είναι η αδιαφορία η έλλειψη ζεστασίας και στοργής από το οικογενειακό περιβάλλον. Υπάρχουν βέβαια και οι υπέρ-προστατευτικοί γονείς οι οποίοι δημιουργούν τα αντίθετα αποτελέσματα από αυτά που επιθυμούν. Όταν ένα παιδί μεγαλώνει κάτω από ένα προστατευτικό «κάλυμμα» δεν μπορεί να γίνει μια ολοκληρωμένη προσωπικότητα αντιθέτως γίνεται ευάλωτο αδυνατεί να προστατευτεί μόνος του και ενισχύεται η πιθανότητα να πέσει θύμα εκφοβισμού. Σαφώς βέβαια υπάρχουν και γονείς που θεωρούν πως με την υπερβολική ελαστικότητα τους κάνουν καλό στο παιδί, αντιθέτως όμως το παιδί μαθαίνει να είναι ανεξέλεγκτο και να μην έχει όρια να γίνεται επιθετικό και να αντιδράει χωρίς σεβασμό. Όλοι αυτοί οι παράγοντες που αναφέραμε ποιο πάνω προδιαθέτουν ένα παιδί ώστε να εμφανίσει επιθετική συμπεριφορά.

Ένας ακόμη παράγοντας που βοηθάει στην ανάπτυξη και στην ομαλή κοινωνικοποίηση ενός παιδιού είναι το σχολείο. Ο ρόλος του σχολείου πέρα από τη μόρφωση του παιδιού είναι και η προετοιμασία ένταξης και προσαρμογής του στο εύρη κοινωνικό σύνολο ο εκπαιδευτικός έχει υποχρέωση να συμβάλει στη προσαρμογή του παιδιού στο σύνολο με διάφορους τρόπους. Το σχολικό περιβάλλον όπως όλοι γνωρίζουμε αποτελεί μια μικρογραφία του ευρύτερου κοινωνικού συνόλου πολλές φορές έχουν παρατηρηθεί κρούσματα βίαιης και εκφοβιστικής συμπεριφοράς από και προς τους μαθητές ενός σχολικού περιβάλλοντος. Η επιθετική και εκφοβιστική συμπεριφορά του εκπαιδευτικού προς τους μαθητές είναι μια από τις μεγαλύτερες αιτίες ανάπτυξης του φαινομένου. Στόχος του εκπαιδευτικού αλλά και γενικά όλου του σχολικού περιβάλλοντος είναι να δημιουργήσει στα παιδιά το αίσθημα της ευγενής άμιλλας της σωστής συνεργασίας και της ενθάρρυνσης τους ώστε να ασχοληθούν με σχολικές δραστηριότητες με στόχο να αναδείξουν τις δεξιότητες τους. Η έλλειψη όλων αυτών η αδιάφορη στάση του εκπαιδευτικού τα επιθετικά ξεσπάσματα με σκοπό να προσβάλει ή να μειώσει ένα παιδί, η απουσία ενδιαφέροντος των γονιών για τη σχολική ζωή του παιδιού αλλά και η αδυναμία δημιουργίας προγραμμάτων ψυχοκοινωνικής αγωγής είναι παράγοντες που αυξάνουν τις πιθανότητες να εκδηλώσει το παιδί βίαιες και εκφοβιστικές συμπεριφορές. Μια ακόμη σημαντική αιτία είναι ο διαχωρισμός των παιδιών από τον εκπαιδευτικό με βάση είτε τις επιδόσεις τους στα σχολικά τους καθήκοντα είτε ακόμη και λόγο της καταγωγής τους και του θρησκευματος τους. Αυτό έχει ως αποτέλεσμα από τη μια το παιδί να περιθωριοποιείται να απομονώνεται και να νιώθει μειονεκτικά απέναντι στο υπόλοιπο σύνολο από την άλλη πλευρά όμως υπάρχουν πιθανότητες το παιδί να λειτουργήσει εντελώς διαφορετικά εκδηλώνοντας αντιδραστική και επιθετική συμπεριφορά και προς τον δάσκαλο του αλλά και προς τους συμμαθητές του.

Ένας ακόμη παράγοντας που εντείνει την αύξηση του φαινομένου του σχολικού εκφοβισμού-bullying είναι ότι τα κρούσματα επιθετική συμπεριφοράς και οι διάφορες συγκρούσεις που δημιουργούνται μέσα σε ένα σχολικό περιβάλλον συχνά δεν αντιμετωπίζονται με τον κατάλληλο τρόπο καθώς οι εκπαιδευτικοί είτε τυχαίνει να μην γνωρίζουν πώς να το χειριστούν είτε αδιαφορούν. Οι επιπτώσεις φυσικά που μπορεί να έχει ως αντίκτυπο όλη αυτή η κατάσταση σε ένα παιδί μπορεί να είναι ακόμη και καταστροφικές για την μετέπειτα πορεία του στο ευρύ κοινωνικό σύνολο.

Φυσικά ο σχολικός εκφοβισμός και η βίαιη συμπεριφορά των μαθητών είναι ένα φαινόμενο που δεν αφορά μόνο το εκπαιδευτικό σύστημα και το οικογενειακό περιβάλλον αλλά και ολόκληρο το κοινωνικό σύνολο. Το ελλιπές εκπαιδευτικό σύστημα αλλά και η πιθανή εκδήλωση βίαιης συμπεριφοράς μέσα στο οικογενειακό περιβάλλον δεν είναι οι μοναδικές αιτίες που μπορεί να οδηγήσουν ένα παιδί στο να εκδηλώσει επιθετική συμπεριφορά. Το κοινωνικό περιβάλλον στο οποίο μεγαλώνει ένα παιδί και μέσα σε αυτό διαμορφώνει το χαρακτήρα του είναι ένας σημαντικός παράγοντας που μπορεί σε περίπτωση που δεν είναι ένα υγιές περιβάλλον να έχει ως αποτέλεσμα την εκδήλωση επιθετικής συμπεριφοράς. Μια από τις σημαντικότερες αιτίες που μπορεί να οδηγήσουν σε τέτοιες συμπεριφορές είναι η γειτονία στην οποία μεγαλώνει ένα παιδί εάν είναι υποβαθμισμένη, κακόφημη και υπάρχουν πολλά κρούσματα βίαιης συμπεριφοράς αλλά και οι σχέσεις μεταξύ των κατοίκων της. Η έλλειψη ελεύθερων χώρων αθλητισμού και ψυχαγωγίας στέρει από το παιδί την δυνατότητα εκτόνωσης του από τυχόν άγχος και στρες στρέφοντας το σε άλλες μορφές εκτόνωσης αντιδρώντας με επιθετική συμπεριφορά. Τα υψηλά επίπεδα εγκληματικότητας αλλά και η ανεξέλεγκτη προβολή βίαιων καταστάσεων από τα μέσα μαζικής επικοινωνίας ακόμη και από τα σύγχρονα ηλεκτρονικά παιχνίδια είναι ένας ακόμη παράγοντας που επηρεάζει το παιδί ώστε να υιοθετήσει τέτοιες συμπεριφορές. Παράλληλα σε συνδυασμό με όλα τα παραπάνω μια ακόμη σημαντική αιτία είναι η κοινωνική ανισότητα που πολλά παιδιά πέφτουν θύματα εκφοβισμού-bullying είναι λόγο διαφορετικότητας στην εθνικότητα, στο θρήσκευμα είτε λόγο της σωματικής τους διάπλασης εάν για παράδειγμα είναι παχουλό εάν είναι όμορφο ή ακόμη και εάν έχει κάποιο κινητικό πρόβλημα λόγο ασθενείας. Η παιδική επιθετικότητα ουσιαστικά είναι ένας αντικατοπτρισμός της σύγχρονης σημερινής κοινωνίας μια ένδειξη ότι υπάρχει κάποιο πρόβλημα. Στη σημερινή εποχή που το παιδί δέχεται καθημερινά ένα σορό ερεθίσματα και είναι δύσκολο να το περιορίσεις πολλά παιδιά επηρεάζονται και άθελα τους αναγκάζονται να υιοθετήσουν τέτοιες συμπεριφορές ώστε να επιβιώσουν μέσα το κοινωνικό περιβάλλον για να μην βιώσουν την απομόνωση και το περιθώριο. Έτσι όλο αυτό έχει σαν αποτέλεσμα τα παιδιά ότι βίαιη και εκφοβιστική συμπεριφορά βιώνουν ή είναι απλοί παρατηρητές στη καθημερινή τους ζωή στο κοινωνικό τους περιβάλλον να την αναπαράγουν στο σχολικό περιβάλλον [6].

Οι επιπτώσεις που έχει ως αποτέλεσμα ο σχολικός εκφοβισμός σε ένα παιδί είναι πολύ σοβαρές για τον ψυχισμό του αλλά και την μετέπειτα πορεία του στη κοινωνία δημιουργεί ένα αίσθημα φόβου στο παιδί οδηγώντας το σε αντικοινωνικές συμπεριφορές. η λεκτική και σωματική βία, η ταπείνωση, ο εξευτελισμός που βιώνουν συχνά τα θύματα έχουν αντίκτυπο στην διαμόρφωση της προσωπικότητας τους και στη συμπεριφορά τους. Οι ψυχολογικές επιπτώσεις που επιφέρει ο σχολικός εκφοβισμός σε ένα θύμα είναι μακροχρόνιες η χαμηλή αυτοπεποίθηση και αυτοεκτίμηση, το συναίσθημα κατωτερότητας, η ανασφάλεια, ο φόβος και η αρνητική αυτοκριτική. Πολύ συχνά βέβαια έχει παρατηρηθεί ότι μια τέτοια κατάσταση μπορεί να προκαλέσει στο παιδί ακόμη και ψυχοσωματικές διαταραχές όπως για παράδειγμα απώλεια όρεξης ή το αντίθετο βουλιμία και υπερκατανάλωση τροφής σε μεγάλες ποσότητες. Σε περιπτώσεις βέβαια που ένα παιδί έχει πέσει θύμα σχολικού εκφοβισμού επανειλημμένες φορές παρατηρούμε πως εμφανίζονται σημάδια μελαγχολίας, απομόνωσης καθώς αρχίζει και αποδέχεται τους χαρακτηρισμούς και την όλη κατάσταση. Επιπλέον η εμφάνιση μαθησιακών δυσκολιών όπως είναι η δυσλεξία αλλά και η χαμηλή επίδοση στα μαθητικά του καθήκοντα εντείνει ακόμη περισσότερο το αίσθημα της κατωτερότητας και κάνει το παιδί να αισθάνεται μειονεκτικά απέναντι στους συμμαθητές του. Τελειώνοντας μια από τις πιο σημαντικές επιπτώσεις που έχει ο σχολικός εκφοβισμός σε ένα παιδί είναι η εμφάνιση ακραίων συμπτωμάτων κατάθλιψης που πολλές φορές μπορεί να οδηγήσουν ακόμη και στην αυτοκτονία. Όλες αυτές οι επιπτώσεις της εκφοβιστικής συμπεριφοράς έχουν ως αποτέλεσμα να επηρεάσουν αρνητικά την μετέπειτα συμπεριφορά ενός παιδιού υπάρχουν πολλές πιθανότητες όταν ένα παιδί έχει πέσει θύμα εκφοβισμού ως ενήλικας να είναι ένα άτομο με χαμηλή αυτοπεποίθηση, κλειστός χαρακτήρας και ανεπαρκές στο να διαχειριστεί δύσκολες καταστάσεις.



Εικόνα 1.3: Εκφοβισμό

1.4 Τρόποι Αντιμετώπισης

Τα τελευταία χρόνια η ραγδαία ανάπτυξη του φαινομένου του σχολικού εκφοβισμού-bullying είναι ένα από τα μεγαλύτερα προβλήματα που αντιμετωπίζει η σύγχρονη κοινωνία και το εκπαιδευτικό σύστημα. Κατά κύριο λόγο οι εκπαιδευτικοί αλλά και οι γονείς έχουν βασικό ρόλο στην αντιμετώπιση του φαινομένου. Βέβαια πολλοί γονείς έχουν μια επιθετική στάση και νιώθουν πλήρη απογοήτευση προς το εκπαιδευτικό σύστημα λόγω της ανεπάρκειας του να προστατεύσει τα παιδιά από τέτοιες καταστάσεις. Το γεγονός ότι ο αριθμός των αυτοκτονιών που όσο πάει και αυξάνεται λόγω των σκληρών κρουσμάτων σχολικού εκφοβισμού εντείνει ακόμη περισσότερο την ανάγκη για περισσότερη πρόληψη και αντιμετώπιση τέτοιων καταστάσεων. Έχει παρατηρηθεί ότι σε πολλές χώρες του εξωτερικού όπου αντιμετωπίζουν συχνά κρούσματα σχολικού εκφοβισμού για την εξάλειψη τους έχει ιδρυθεί διάφορα προγράμματα πρόληψης – αντιμετώπισης που δίνουν έμφαση στη σχέση μεταξύ του εκπαιδευτικού με το παιδί, τη σχέση γονέα με το παιδί αλλά και τη συνεργασία γονέα .

1.4.1 Στάση Εκπαιδευτικού

Ένας σημαντικός παράγοντας για την εξάλειψη του φαινομένου του σχολικού εκφοβισμού είναι η στάση του εκπαιδευτικού προς τα παιδιά. Όσο αναφορά την πρόληψη του σχολικού εκφοβισμού και των επιθετικών συμπεριφορών στο σχολικό περιβάλλον σημαντική προϋπόθεση είναι ο εκπαιδευτικός να προάγει το αίσθημα της ομόνοιας και της συνεργασίας ανάμεσα στα παιδιά. Ο εκπαιδευτικός πρέπει να συζητάει με τους μαθητές για του κανόνες συμπεριφοράς μέσα στο σχολικό περιβάλλον που είναι μια μικρογραφία της κοινωνίας αλλά και τους τρόπους αντιμετώπισης του σχολικού εκφοβισμού. Ο ρόλος του εκπαιδευτικού είναι βασικός άξονας στη διαμόρφωση της συμπεριφοράς ενός παιδιού, θα πρέπει να ενισχύει την αλληλεγγύη την συντροφικότητα και την ευγένεια μεταξύ των παιδιών προλαμβάνοντας με αυτό ο τρόπο τα κρούσματα περιθωριοποίησης και απομόνωσης παιδιών από το υπόλοιπο σύνολο. Κάθε μέλος μιας σχολικής κοινότητας έχει ο δικό του ρόλο και την δική του ευθύνη για την πρόληψη και αντιμετώπιση του σχολικού

εκφοβισμού. Βέβαια όλοι μαζί θα πρέπει να συνεργαστούν με πολλούς και διάφορους τρόπους για την αντιμετώπιση τέτοιων περιστατικών.

Η στάση του εκπαιδευτικού θα πρέπει να είναι φιλική προς τα παιδιά να έχει την δυνατότητα το παιδί να πλησιάσει και να μιλήσει για το πρόβλημα που αντιμετωπίζει και όχι να είναι ο τιμωρός εκπαιδευτικός που το παιδί θα φοβάται να του μιλήσει. Ο εκπαιδευτικός πρέπει να δείξει στο παιδί ότι είναι δίπλα του σε ότι χρειαστεί. Όσο εύκολο και αν φαίνεται η στάση που πρέπει να έχει ένας εκπαιδευτικός δεν είναι κάτι απλό πρέπει να αντιμετωπίζει κάθε παιδί ξεχωριστά και να ο πλησιάζει με διαφορετικό τρόπο τα κάθε παιδί δεν είναι ίδιο. Ακόμη περισσότερο βέβαια για τα παιδιά που λόγω κοινωνικής-ταξικής διαφοράς αλλά και λόγω καταγωγής δεν είναι εύκολη η ένταξη τους και η αποδοχή τους.

1.4.2 Διαχείριση Σχολικού Εκφοβισμού

Η διαχείριση του φαινομένου του σχολικού εκφοβισμού δεν θα μπορούσε να χαρακτηριστεί ως κάτι εύκολο. Είναι φανερό ότι για την αντιμετώπιση του σχολικού εκφοβισμού θα πρέπει να τεθεί σε εφαρμογή κάποιο πρόγραμμα διαχείρισης τέτοιων κρουσμάτων το οποίο θα γίνει γνωστό σε μαθητές, εκπαιδευτικούς και γονείς διότι μόνο με την προσπάθεια και την συνεργασία όλων θα επιφέρει κάποιο αποτέλεσμα το εκάστοτε σχέδιο. Φυσικά για την καλύτερη και ποιο αποτελεσματική διαχείριση τέτοιων περιστατικών θα πρέπει η κάθε περίπτωση να αντιμετωπίζεται διαφορετικά ανάλογα με τη φύση και η μορφή του εκφοβισμού. Μια από τις βασικές αρχές διαχείρισης σχολικού εκφοβισμού είναι η διαμόρφωση ομάδας διαχείρισης η οποία θα απαρτίζεται από το διευθυντή και ένα αριθμό εκπαιδευτικών αλλά και η διοργάνωση δράσεων για την ενημέρωση, τη πρόληψη αλλά και την διαχείριση σχολικού εκφοβισμού και βίαιων συμπεριφορών. Αρχικά θα πρέπει να αποφασιστεί ότι πρόκειται για σχολικό εκφοβισμό κάτι τέτοιο φυσικά μπορεί να γίνει και με τη βοήθεια ενός ψυχολόγου. Ο χειρισμός μιας τέτοιας κατάστασης θα πρέπει να γίνεται σε σύντομο χρονικό διάστημα από το συμβάν και σε χώρο όπου ο μαθητής αισθάνεται ασφαλής. Θα πρέπει να γίνεται εκτεταμένη διερεύνηση για περισσότερες πληροφορίες σε ότι αφορά το συμβάν τον εντοπισμό των κινήτρων, το τόπο, το χρόνο αλλά κυρίως για τη μορφή του εκφοβισμού ώστε να αποδοθούν και οι ανάλογες ευθύνες. Απαραίτητη κρίνεται η ενημέρωση των γονέων του θύματος και του θύτη.

Χρειάζεται χρόνος, πολλές συναντήσεις ώστε το παιδί να ανοιχτεί και να μιλήσει αλλά και ψυχολογική στήριξη και από εκπαιδευτικούς και από γονείς διότι τα παιδιά θύματα έχουν την ανάγκη να μιλήσουν χωρίς να φοβούνται τις συνέπειες. Τέλος θα πρέπει να γίνεται καταγραφή των περιστατικών και τήρηση αρχείου σχετικά με τη διαχείριση του περιστατικού με τέτοιο τρόπο ώστε να διαφυλάσσονται τα προσωπικά στοιχεία των θυμάτων [7].

1.4.3 Κοινωνικό-εθνική ποικιλότητα

Έχει παρατηρηθεί ότι πολλές φορές τα θύματα των περιστατικών του σχολικού εκφοβισμού τυχαίνει να είναι παιδιά που διαφέρουν από το υπόλοιπο κοινωνικό σύνολο κυρίως λόγω εθνικότητας, κοινωνικής τάξης και οικογενειακού υπόβαθρου. Ένας ακόμη λόγος που η συγκεκριμένη μερίδα παιδιών πέφτουν θύματα σχολικού εκφοβισμού είναι και η στάση και η συμπεριφορά του εκπαιδευτικού απέναντι σε αυτά τα παιδιά. Στο πλαίσιο του σχολικού εκφοβισμού ένας ακόμη σημαντικός παράγοντας είναι και οι φυλετικές διαφορές που υπάρχουν σε μια σχολική κοινότητα. Κατά τη δεκαετία του '80 η Ελλάδα δέχτηκε ένα τεράστιο μεταναστευτικό κύμα το οποίο εμφανίζεται και στα σχολεία. Όπως έχουμε προ αναφέρει η σχολική κοινωνία είναι μια μικρογραφία της ευρύτερης κοινωνίας μας έτσι όπως υπήρξαν αντιδράσεις και επιφυλακτικότητα στην ένταξη των μεταναστών στην κοινωνία κάτι αντίστοιχο παρατηρήθηκε και στην ένταξη αλλοδαπών μαθητών στα σχολεία. Ακόμη ένας λόγος που εντείνει το πρόβλημα είναι οι διακρίσεις των εκπαιδευτικών ανάμεσα στους μαθητές καθώς δεν υπάρχει ισότητα και σεβασμός. Η εμφάνιση συγκρούσεων και επιθετικής συμπεριφοράς είναι επακόλουθο της ανισότητας που επικρατεί η θυματοποίηση των αλλοδαπών μαθητών αλλά και των μαθητών με χαμηλό οικογενειακό υπόβαθρο είναι ένα συχνό φαινόμενο που πολλές φορές έχει δυσάρεστες συνέπειες στη μαθητικές του επιδόσεις αλλά και στην εξέλιξη του στην κοινωνία. Βέβαια δεν υπάρχει μια ξεκάθαρη εικόνα όσο αναφορά το σχολικό εκφοβισμό σε σχέση με την κοινωνικοοικονομική και φυλετική διαφορετικότητα που υπάρχει μεταξύ των μαθητών. Σημαντικό είναι να γνωρίσουμε τα κίνητρα που κάποιος θα εκφοβίσει λόγο διαφορετικότητας. Ένας τρόπος που θα μπορούσε να βοηθήσει ώστε να μειωθεί ο εκφοβισμός με βάση τη διαφορετικότητα είναι η Διαπολιτισμική Εκπαίδευση που αφορά την επίδραση του ρατσισμού σε μια σχολική

ομάδα. Η σχολική κοινότητα αλλά και οι εκπαιδευτικοί θα έχουν την ευκαιρία να μεταλαμπαδεύσουν στους μαθητές τη « γλύκα » της διαφορετικότητας των ηθών και εθίμων, της γλώσσας, του πολιτισμού αλλά και της κουλτούρας δημιουργώντας ένα κλίμα σεβασμού και αναγνώρισης της διαφορετικότητας είτε είναι εθνική είτε κοινωνική. Η διαπολιτισμική εκπαίδευση επιτυγχάνεται με την οργάνωση δράσεων, ομιλιών μεταξύ των μελών του σχολείου αλλά και τω γονέων και διαφόρων εκδηλώσεων [8-9].



Εικόνα 1.4.3: STOP BULLYING

ΚΕΦΑΛΑΙΟ 2

ΕΥΦΥΕΙΣ ΜΕΘΟΔΟΙ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ (Machine Learning)

2.1 Γενικά για μηχανική μάθηση και data mining

Γενικά η μηχανική μάθηση είναι ένα γνωστικό σύστημα της καθημερινής μας ζωής που μπορεί να συνδεθεί με δυο βασικές ιδιότητες η μια είναι η ικανότητα στην πρόσκτηση γνώσης κατά την αλληλεπίδραση του με το περιβάλλον και η άλλη είναι η ικανότητα βελτίωσης με την επανάληψη τον τρόπο εκτέλεσης μιας ενέργειας. Υπάρχουν διάφοροι ορισμοί για την μάθηση κατά τον Simon ('83) ένας από αυτούς είναι ότι η μάθηση σηματοδοτεί προσαρμοστικές αλλαγές σε ένα σύστημα με την προϋπόθεση ότι αυτές του δίνουν την δυνατότητα να κάνει την ίδια εργασία της ίδιας κατηγορίας πιο αποτελεσματικά από την προηγούμενη φορά. Επίσης κατά τον Minsky ('85) μάθηση είναι να κάνουμε χρήσιμες αλλαγές στο μυαλό μας και τέλος σύμφωνα με τον Michalski είναι η δημιουργία και η αλλαγή αναπαράστασης των εμπειριών [10-11].



Εικόνα 2.1: Data Mining

Ο άνθρωπος καθώς παρατηρεί και δημιουργεί μια απλοποιημένη εκδοχή που ονομάζεται μοντέλο προσπαθεί να κατανοήσει το περιβάλλον του. Η κατασκευή ενός τέτοιου μοντέλου ονομάζεται επαγωγική μάθηση ενώ η διαδικασία ονομάζεται επαγωγή. Είναι σημαντικό να προσθέσουμε ότι ο άνθρωπος μπορεί να οργανώνει και να συσχετίζει όλες τις εμπειρίες και τις παραστάσεις του δημιουργώντας νέες δομές που χαρακτηρίζονται ως πρότυπα.

Η δημιουργία μοντέλων ή προτύπων από ένα σύνολο δεδομένων από
ένα υπολογιστικό σύστημα ονομάζεται μηχανική μάθηση

Αξίζει να σημειωθεί ότι η μηχανική μάθηση έχει πολλές τεχνικές που ανάλογα τη φύση του προβλήματος χρησιμοποιείται κάθε είδος τεχνικής [12].

- Ένα από αυτά είναι η Μάθηση με Επίβλεψη (Supervised Learning) ή αλλιώς μάθηση με παραδείγματα (Learning from observation). Στη μάθηση με επίβλεψη το σύστημα προσπαθεί να μάθει μια έννοια ή μια συνάρτηση από ένα σύνολο δεδομένων που αποτελεί περιγραφή ενός μοντέλου.
- Το δεύτερο είδος τεχνικής είναι η μάθηση χωρίς επίβλεψη (Unsupervised Learning) ή μάθηση από παρατήρηση (Learning from observation). Το σύστημα θα πρέπει από μόνο του να ανακαλύψει τις συσχετίσεις σε ένα σύνολο δεδομένων δημιουργώντας πρότυπα χωρίς όμως να είναι γνωστό αν υπάρχουν πόσα είναι και ποια είναι.

Όσο αναφορά τώρα τον όρο data mining είναι μια έννοια γνωστή μέσα από την εξόρυξη δεδομένων. Με τον όρο αυτό εννοούμε την ανακάλυψη γνώσης μέσα από βάσεις δεδομένων είναι η διερεύνηση μιας ενδιαφέρουσας πληροφορίας από μεγάλες βάσεις δεδομένων με τη χρήση αλγορίθμων ομαδοποίησης ή κατηγοριοποίησης με βάση της αρχές της στατιστικής ή της τεχνητής νοημοσύνης της μηχανικής μάθησης και των συστημάτων βάσεων δεδομένων. Πρέπει να σημειωθεί ότι βασικός σκοπός της εξόρυξης δεδομένων είναι ότι η πληροφορία που θα εξαχθεί αλλά και τα πρότυπα που θα επακολουθήσουν θα έχουν δομή κατανοητή προς τον άνθρωπο με αποτέλεσμα να πάρει τις καλύτερες αποφάσεις. Η εξαγωγή προτύπων από δεδομένα συμβαίνει εδώ και αιώνες οι πρώτες μέθοδοι για τον προσδιορισμό προτύπων ήταν αυτές της Bayes και της ανάλυσης της παλινδρόμησης. Με το πέρασμα του χρόνου βέβαια αλλά και με την εξέλιξη της τεχνολογίας και των υπολογιστών παρουσιάστηκε και αύξηση

όγκου των δεδομένων. Με την αύξηση των δεδομένων αλλά και της πολυπλοκότητας τους ήταν δύσκολη η χειρωνακτική ανάλυση των δεδομένων αυτό είχε σαν αποτέλεσμα την ανάγκη να αντικατασταθεί από μηχανές αυτόματης επεξεργασίας δεδομένων. Όπως προείπαμε με το πέρασμα του χρόνου οι ραγδαίες τεχνολογικές εξελίξεις που ακολούθησαν συνέβαλαν και αυτές σε πολλά επίπεδα για την καλύτερη ανάλυσης δεδομένων. Κάποιες από αυτές ήταν η ανακάλυψη των νευρωνικών δικτύων [13], η συσταδοποίηση, οι γενετικοί αλγόριθμοι [14], τα δένδρα απόφασης και η μηχανική υποστήριξη διανυσμάτων.

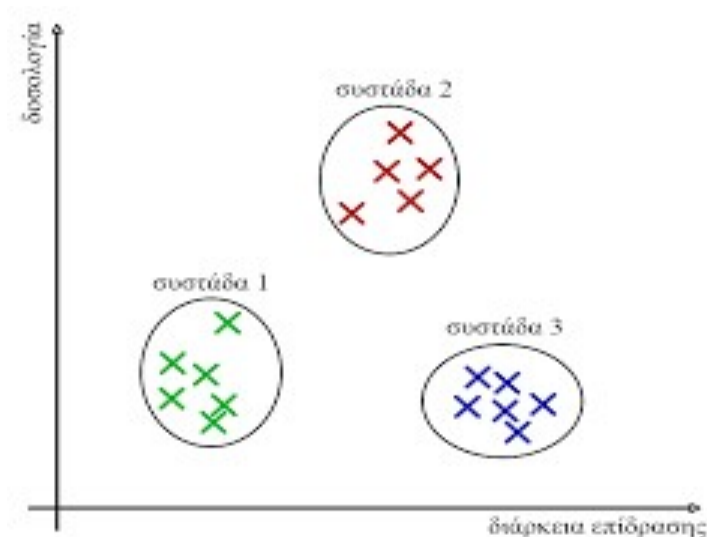
Υπάρχει μια μεγάλη γκάμα τεχνικών εξόρυξης δεδομένων που ανάλογα με το είδος των δεδομένων αλλά και το είδος της γνώσης που εξάγεται διαχωρίζονται σε διαφορετικές κατηγορίες. Κάποιες από τις βασικές μεθόδους εξόρυξης δεδομένων θα της δούμε παρακάτω αναφορικά και στη συνέχεια θα αναλύσουμε κάποιες [15-16].

- Κατηγοριοποίηση /Ταξινόμηση (classification)
- Συσταδοποίηση / Ομαδοποίηση (segmentation)
- Εξαγωγή και Ανάλυση Συσχετίσεων(association)
- Οπτικοποίηση
- Ανίχνευση Ανωμαλιών
- Παλινδρόμηση (regression)

ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ: πρόκειται για ένα σύνολο από εγγραφές και κάθε μια εγγραφή απαρτίζεται από δεδομένα μια από όλες αυτές τις εγγραφές ορίζουμε και ως κατηγορία. Στόχος είναι η δημιουργία ενός μοντέλου με βάση τα δεδομένα που έχουμε που μπορεί να χρησιμοποιηθεί αργότερα για την βέλτιστη κατηγοριοποίηση μελλοντικών δεδομένων. Η κατηγοριοποίηση συχνά τη συνδέουν με το γενικό όρο της πρόβλεψης. Μια κλάση μπορεί να πάρει διακριτές τιμές από ένα πεπερασμένο σύνολο σε αντίθεση με άλλες τεχνικές όπως παλινδρόμηση. Στη κατηγοριοποίηση ανήκουν τα δένδρα αποφάσεων, τα νευρωνικά δίκτυα και η κατηγοριοποίηση κατά Bayes. Η ταξινόμηση είναι μια ευρεία κατηγορία μεθόδων, οι οποίες απαιτούν εκπαίδευση από τα δεδομένα. Το γεγονός αυτό δίνει μεγαλύτερη δυναμική στους αλγόριθμους, οι οποίοι αξιοποιώντας την υπάρχουσα γνώση αποδεικνύονται πιο αποδοτικοί από τους αλγόριθμους χωρίς επίβλεψη/εκπαίδευση (π.χ. αλγόριθμους

ταξινόμησης) [17]. Οι αλγόριθμοι ταξινόμησης έχουν χρησιμοποιηθεί σε διάφορες εφαρμογές με αιχμή του δόρατος μεθόδους στην ιατρική, όπως για παράδειγμα η κατηγοριοποίησης ηλεκτροεγκεφαλογραφήματος [18], και η κατηγοριοποίηση των σταδίων του ύπνου στις αντίστοιχες μελέτες ύπνου [19].

ΣΥΣΤΑΔΟΠΟΙΗΣΗ: είναι μια περιγραφική μέθοδος. Έχουμε ένα σύνολο δεδομένων που συγκροτούνται από κάποια δείγματα τα οποία μεταξύ τους έχουν κάποια ομοιότητα και έτσι δημιουργούνται οι συστάδες/ ομάδες. Ουσιαστικά ψάχνουμε ένα πεπερασμένο σύνολο συστάδων για να περιγράψουμε τα δεδομένα. Μέτρο σύγκρισης αποτελεί η ευκλείδεια απόσταση και πολλά αλλά στοιχεία ανάλογα με την εφαρμογή όπως τον τύπο των δεδομένων που εισάγονται στον αλγόριθμο. Στόχος της συσταδοποίησης είναι η δημιουργία συστάδων/ομάδων που θα περιέχουν όμοια η παρεμφερή δείγματα. Στο παρακάτω παράδειγμα παρατηρούμε το αποτέλεσμα της συσταδοποίησης ενός συνόλου φαρμακευτικών δεδομένων τα οποία έχουν διαχωριστεί σε τρεις συστάδες/ομάδες με βάση τη δοσολογία και την διάρκεια επίδρασης του φαρμάκου. Εφαρμογές της ομαδοποίησης έχουν πραγματοποιηθεί σε επεξεργασία εικόνας και σε προβλήματα κατάτμησης [20-21], στην δημιουργία προφίλ ασθενών για την πρόγνωση του καρκίνου [22] κ.α.



Εικόνα 2.2: Συσταδοποίηση / Ομαδοποίηση

ΕΞΑΓΩΓΗ ΚΑΙ ΑΝΑΛΥΣΗ ΣΥΣΧΕΤΙΣΕΩΝ: είναι μια από τις σημαντικές διεργασίες της εξόρυξης δεδομένων διότι οι μέσα από τους κανόνες συσχέτισης μας

δύνεται η δυνατότητα με συνοπτικό τρόπο να εκφράσουμε χρήσιμες πληροφορίες. Με τους κανόνες συσχέτισης μπορούμε να ανακαλύψουμε κρυφές συσχετίσεις ανάμεσα στα γνωρίσματα ενός συνόλου δεδομένων. Οι συσχετισμοί αυτοί είναι της μορφής $A \rightarrow B$ όπου το A και το B είναι σύνολα χαρακτηριστικών του συνόλου που θα αναλύσουμε. Ένας κανόνας συσχέτισης $A \rightarrow B$ είναι δεδομένο ότι θα εμφανίσει τα χαρακτηριστικά του συνόλου A και φυσικά του συνόλου B . Στόχος είναι η καλύτερη και πιο αποδοτική εξαγωγή συμπερασμάτων μέσω των κανόνων συσχέτισης.

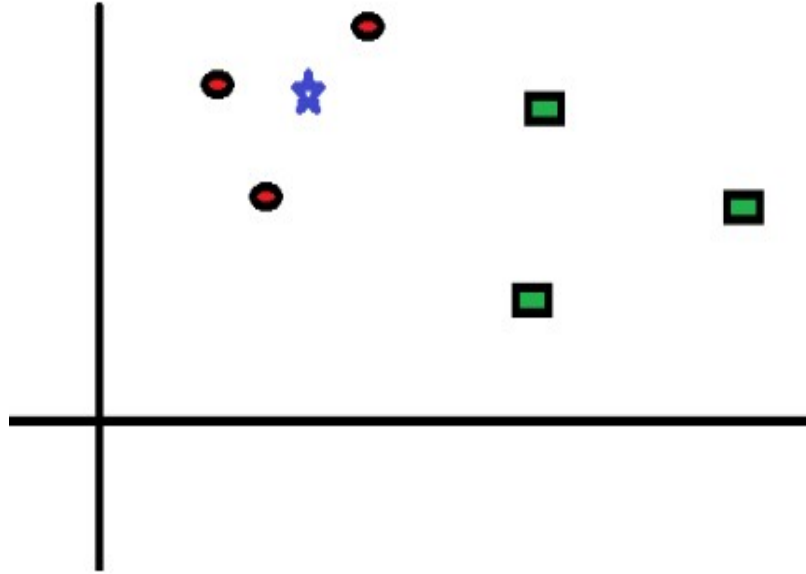
2.2 Μέθοδοι Ταξινόμησης

Σε αυτό το υποκεφάλαιο θα παρουσιάσουμε και θα περιγράψουμε κάποιες από τις πιο σημαντικές τεχνικές ταξινόμησης και πιο συγκεκριμένα:

- Το Κοντινότερο Γείτονα (KNN=K-nearest neighbor)
- Τον Μπευζιανό Ταξινομητή (Naïve Bayes Classifier)
- Τα Δένδρα Απόφασης (Decision Trees)
- Τις Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines-SVM)

2.2.1 K-Κοντινότερος γείτονας (KNN=K-nearest neighbor)

Ο αλγόριθμος ταξινόμησης KNN είναι ένας από τους απλούστερους και πιο σημαντικούς αλγορίθμους στη μηχανική μάθηση και χρησιμοποιείτε κυρίως για ταξινόμηση και παλινδρόμηση. Τρία σημαντικά πλεονεκτήματα αυτού του αλγορίθμου είναι η εύκολη ερμηνεία των αποτελεσμάτων που εξάγει, ο μικρός χρόνος υπολογισμού αλλά και η μεγάλη προγνωστική του δύναμη. Στη παρακάτω εικόνα βλέπουμε ένα παράδειγμα KNN που θα ο εξηγήσουμε στη συνέχεια.



Εικόνα 2.3: KNN(K-nearest neighbor

Όπως παρατηρούμε στην παραπάνω εικόνα υπάρχουν τρεις ομάδες διαφορετικών σχημάτων ένα με κόκκινους κύκλους, ένα με πράσινα τετράγωνα και ένα με μπλε αστέρι. Σύμφωνα με τον αλγόριθμο KNN το μπλε αστέρι θα πρέπει να ενσωματωθεί σε μια από τι δυο κατηγορίες. Με βάση αυτά που γνωρίζουμε για αυτόν τον αλγόριθμο καταλαβαίνουμε ότι το αστέρι θα ενσωματωθεί στη κατηγορία που βρίσκεται πιο κοντά άρα στους κόκκινους κύκλους διότι είναι πιο κοντά του [23].

2.2.2 Μπευζιανός ταξινομητής (Naive Bayes Classifier)

Μια ακόμη τεχνική ταξινόμησης είναι αυτή του Naive Bayes που είναι βασισμένη στο θεώρημα του Bayes. Σύμφωνα με τον Bayes υπάρχει πιθανότητα ένα χαρακτηριστικό να μην έχει καμία σχέση με κάποιο άλλο χαρακτηριστικό σε μια κατηγορία. Η κατασκευή ενός Μπευζιανού ταξινομητή είναι εύκολη και είναι ιδιαίτερα χρήσιμη στην εξαγωγή αποτελεσμάτων για πολύ μεγάλα σύνολα δεδομένων [24-26]. Το θεώρημα αυτό δίνει τη δυνατότητα ενός τρόπου υπολογισμού της οπίσθιας πιθανότητας $\mathcal{P}(c|x)$ από $\mathcal{P}(c)$, $\mathcal{P}(x)$ και $\mathcal{P}(x|c)$.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

Likelihood Class Prior Probability
 ↓ ↓
 Posterior Probability Predictor Prior Probability

$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$$

- $P(c|x)$ είναι η οπίσθια πιθανότητα του στόχου c που δίνει πρόβλεψη x χαρακτηριστικά.
- $P(c)$ είναι η προηγούμενη πιθανότητα κλάσης.
- $P(x|c)$ είναι η πιθανότητα της προβλεπόμενης κατηγορίας προγνωστικού.
- $P(x)$ είναι η προηγούμενη πιθανότητα του προγνωστικού.

Στη συνέχεια θα χρησιμοποιήσουμε ένα απλό παράδειγμα για να μπορέσουμε να κατανοήσουμε καλύτερα τη λειτουργία ενός Naive Bayes Classifier. Με τη χρήση ενός συνόλου δεδομένων καιρικών συνθηκών αλλά και με την αντίστοιχη μεταβλητή προορισμού που μας υποδεικνύει κατά πόσο οι παίκτες θα μπορέσουν να παίξουν ανάλογα με τις καιρικές συνθήκες.

Πρώτο Βήμα: Αυτό που θα πρέπει αρχικά να κάνουμε είναι να μετατρέψουμε τα δεδομένα μας σε ένα πίνακα συχνοτήτων.

Δεύτερο Βήμα: Το επόμενο βήμα που θα πρέπει να ακολουθήσουμε είναι η δημιουργία ενός πίνακα πιθανοτήτων για την εύρεση πιθανοτήτων όπως πιθανότητα συννεφιάς = 0.29 και πιθανότητα παιχνιδιού = 0.64.

Τρίτο Βήμα: Στο τρίτο και τελευταίο βήμα χρησιμοποιούμε τον Μπευζιανό ταξινομητή για να υπολογίσουμε την πιθανότητα για κάθε κατηγορία.

Weather	Play
Sunny	No
Overcast	Yes
Rainy	Yes
Sunny	Yes
Sunny	Yes
Overcast	Yes
Rainy	No
Rainy	No
Sunny	Yes
Rainy	Yes
Sunny	No
Overcast	Yes
Overcast	Yes
Rainy	No

Frequency Table		
Weather	No	Yes
Overcast		4
Rainy	3	2
Sunny	2	3
Grand Total	5	9

Likelihood table				
Weather	No	Yes		
Overcast		4	=4/14	0.29
Rainy	3	2	=5/14	0.36
Sunny	2	3	=5/14	0.36
All	5	9		
	=5/14	=9/14		
	0.36	0.64		

Εικόνα 2.4: Παράδειγμα Αλγορίθμου Naïve Bayes

Στη συνέχεια μας δίνεται ένα πρόβλημα με την εξής ερώτηση «οι παίκτες θα παίξουν εάν ο καιρός είναι ηλιόλουστος;» και θα πρέπει να απαντήσουμε εάν είναι σωστή ή λανθασμένη.

$$P(Nai|Sunny) = P(Sunny|Nai) * P(Nai)/P(Sunny)$$

$$P(Sunny|Nai) = 3/9 = 0.33, \quad P(Sunny) = 5/14 = 0.36$$

$$P(Nai) = 9/14 = 0.64$$

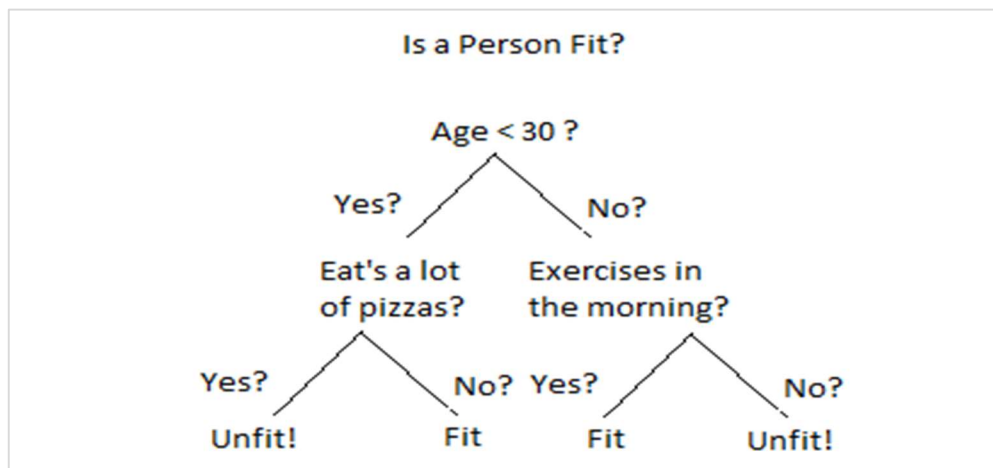
Άρα: $P(Nai|Sunny) = 0.33 * 0.64/0.36 = 0.60$ υπάρχει μεγάλη πιθανότητα.

Επομένως καταλήγουμε ότι ο Μπευζιανός ταξινομητής κάνει χρήση ενός παρόμοιου τρόπου για να μπορέσει να προβλέψει την πιθανότητα διαφορετικής κατηγορίας με διάφορα χαρακτηριστικά. Τέλος ο αλγόριθμος αυτός χρησιμοποιείται κατά κύριο λόγο για την ταξινόμηση κειμένου αλλά και για προβλήματα που έχουν πολλαπλές κλάσεις.

2.3 Δένδρα Αποφάσεων (Decision Trees)

Ακόμη ένας αρκετά σημαντικός αλγόριθμος ταξινόμησης για την σωστή λήψη αποφάσεων είναι τα δένδρα αποφάσεων όπου τα δεδομένα του κάθε προβλήματος μπορούν να αναπαρασταθούν γραφικά [27-29]. Ένα σύστημα ταξινόμησης decision trees είναι ένα εργαλείο λήψης αποφάσεων που η απεικόνιση του είναι ίδια με αυτή ενός δέντρου που περιλαμβάνει όλες τις πιθανές αποφάσεις, όλες τις επιρροές που μπορεί να δεχτεί αλλά και όλα τα πιθανά αποτελέσματα. Συχνά τα δένδρα αποφάσεων

μιμούνται την ανθρώπινη σκέψη με αποτέλεσμα η κατανόηση του είναι πολύ απλή. Τα δέντρα αυτά λειτουργούν με τα εξής δύο βήματα. Μια πορεία προς τα εμπρός και μια πορεία προς τα πίσω. Όσο αναφορά την προς τα εμπρός πορεία έχει να κάνει με την ανάμειξη αυτού που λαμβάνει μια απόφαση στην ταυτοποίηση της απόφασης που πρέπει να ληφθεί. Η προς τα εμπρός πορεία αποκαλύπτει τη δομή του προβλήματος στο στάδιο που απαιτείται ο υπολογισμός των απολαβών και των πιθανοτήτων που αφορούν μελλοντικά γεγονότα. Ενώ η πορεία προς τα πίσω συνδέεται με την ανάλυση του προβλήματος και σκοπός της είναι ο υπολογισμός της προστιθέμενης αξίας ξεκινώντας από πίσω προς τα εμπρός μέχρι να έχουμε την τελική απόφαση. Παρακάτω θα εξετάσουμε ένα παράδειγμα δέντρου αποφάσεων. Στο παρακάτω παράδειγμα ο χρήστης θα πρέπει να εξετάσει αν ένα άτομο έχει καλή φυσική κατάσταση ανάλογα με τις διατροφικές επιλογές και την άσκηση του.



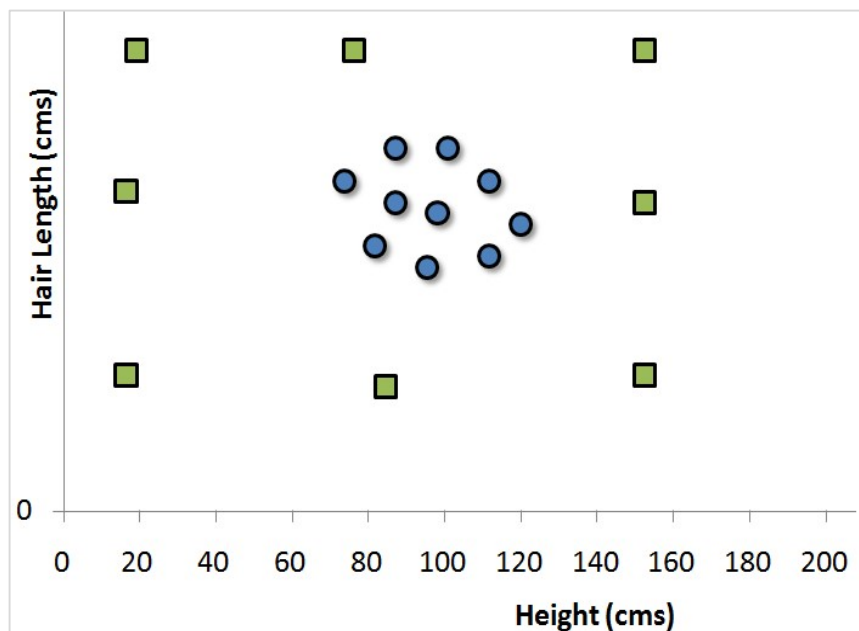
Εικόνα 2.3: Δέντρα Αποφάσεων

Ξεκινώντας να σχεδιάσουμε και να αναλύσουμε το παραπάνω δέντρο αποφάσεων στο πρώτος κόμβος που συναντάμε αναζητάμε να βρούμε την ηλικία του ατόμου δηλαδή εάν είναι πάνω από 30 ή κάτω από 30. Σε αυτό το κόμβο υπάρχουν δυο απαντήσεις Yes και No. Εάν αυτό το άτομο είναι πάνω από 30 ο πρώτος παράγοντας που επηρεάζει τη καλή ή κακή του φυσική κατάσταση είναι η μεγάλη κατανάλωση πίτσας. Ένας δεύτερος επίσης σημαντικός παράγοντας που επηρεάζει είναι η καθημερινή σωματική του άσκηση. Άρα στη περίπτωση που είναι πάνω από 30 και καταναλώνει πολλές πίτσες τότε καταλαβαίνουμε ότι δεν θα έχει καλή φυσική κατάσταση. Επίσης στη περίπτωση που είναι κάτω των 30 αλλά ασκείτε καθημερινά η φυσική του κατάσταση θα είναι καλή Όπως καταλαβαίνουμε πολλές φορές οι αποφάσεις μας μπορεί να επηρεαστούν από διάφορους εξωγενείς παράγοντες.

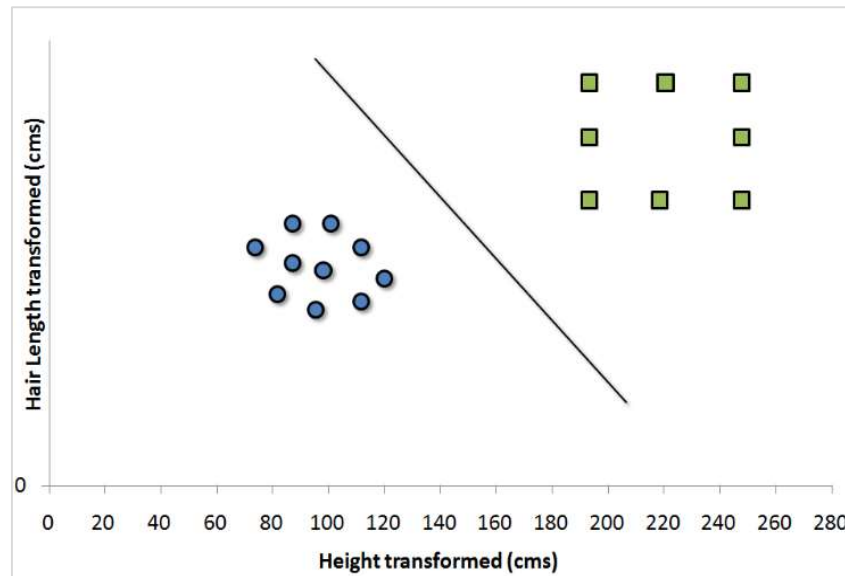
2.4 Μηχανές Διανυσμάτων υποστήριξης (Support Vector Machines-SVM)

Στη μηχανική μάθηση μία ακόμη σημαντική τεχνική ταξινόμησης είναι Μηχανές Διανυσμάτων υποστήριξης. Οι μηχανές αυτές είναι πρότυπα μάθησης αλγορίθμων για την ανάλυση και ταξινόμηση δεδομένων. Με τον αλγόριθμο αυτό το κάθε δεδομένο που σχεδιάζουμε είναι σε n -διάστατο χώρο όπου n είναι ο αριθμός των χαρακτηριστικών που έχουμε. Στο SVM είναι εύκολο να έχουμε ένα γραμμικό υπέρ-επίπεδο μεταξύ των δυο κατηγοριών. Πραγματοποιώντας την ταξινόμηση καθορίζουμε το υπέρ-επίπεδο που διαχωρίζει τις δυο κατηγορίες. Το SVM έχει μια ακόμη λειτουργία που ονομάζεται κόλπο πυρήνα συνήθως όμως είναι χρήσιμη σε προβλήματα μη γραμμικού διαχωρισμού. Στο παρακάτω παράδειγμα θα δούμε τι θα συμβεί σε περίπτωση που η διανομή των στοιχείων ήταν τυχαία [30-32].

Στη περίπτωση αυτή όπως φαίνεται στη φωτογραφία δεν υπάρχει κάποια διαχωριστική γραμμή που να ξεχωρίζει τις δυο κατηγορίες. Αυτό έχει σαν αποτέλεσμα να πρέπει να χαρτογραφήσουμε αυτά τα διανύσματα σε ένα υψηλότερο επίπεδο διαστάσεων ώστε να διαχωριστούν.



Εικόνα 2.5: Παράδειγμα SVM (τυχαία διανομή στοιχείων)



Εικόνα 2.6: Παράδειγμα SVM (ταξινομημένο)

Τελειώνοντας οι Μηχανές Διανυσμάτων Υποστήριξης – SVM είναι ένας πολύ ισχυρός αλγόριθμός που όταν χρησιμοποιείται με τυχαία διανομή και σε συνδυασμό και άλλα εργαλεία μάθησης μηχανών έχει σαν αποτέλεσμα να δίνει μια διαφορετική προσέγγιση στα μοντέλα συνόλων. Η χρησιμότητα αυτού του αλγορίθμου είναι μεγάλη κυρίως σε περιπτώσεις που απαιτείτε ανάγκη για μεγάλη προγνωστική δύναμη.

ΚΕΦΑΛΑΙΟ 3

Το Λογισμικό Weka

3.1 Εισαγωγή

Το λογισμικό Weka(Waikato Environment for Knowledge Analysis) είναι μια μηχανή μηχανικής μάθησης εξόρυξης δεδομένων. Δημιουργήθηκε στο πανεπιστήμιο του Waikato στη Ν. Ζηλανδία. Το όνομα του προήλθε από ένα πουλί που είναι είδος υπό εξαφάνιση στη Ν. Ζηλανδία. Το Weka είναι ευρέως διαδεδομένο στο κοινό και διατίθεται δημοσίως ελεύθερα ώστε να το χρησιμοποιούν οι χρήστες σύμφωνα με την άδεια του GNU(General Public License) [33].



Εικόνα 3.1: Το Weka

Είναι ένα πολύ γνωστό λογισμικό εξόρυξης δεδομένων που έχει χρησιμοποιηθεί σε πολλές εργασίες. Πολλά επιστημονικά βιβλία εξόρυξης δεδομένων πολύ συχνά κάνουν αναφορές σε αυτό λόγω των χαρακτηριστικών του αλλά και των πολλών δυνατοτήτων που προσφέρει. Κάποιες από αυτές θα τις δούμε παρακάτω:

- Το μεγάλο σύνολο μεθόδων κατηγοριοποίησης, παλινδρόμησης, ανάλυσης συστάδων, και κανόνων συσχέτισης είναι ένα δυνατό χαρακτηριστικό του Weka. Δίνει στο χρήστη τη δυνατότητα προεπεξεργασίας των δεδομένων.

- Επίσης δίνει τη δυνατότητα στους χρήστες με γνώσεις προγραμματισμού να το τροποποιήσουν ή ακόμη και να το εξελίξουν μιας και ο κώδικάς του είναι ανοιχτός.
- Είναι γραμμένο στη γλώσσα προγραμματισμού Java που μας δίνει την δυνατότητα εγκατάστασης σε διάφορες πλατφόρμες υλικού και λογισμικού.
- Διαθέτει γραφικό περιβάλλον εργασίας. Η μεγάλη γκάμα βιβλιοθηκών στο διαδίκτυο καθιστά ακόμη πιο εύκολη την εκμάθηση της μηχανικής μάθησης και της εξόρυξης δεδομένων.

Είναι επίσης σημαντικό να τονίσουμε ότι το Weka διαθέτει και δύο διαφορετικές εκδόσεις:

- Η πρώτη έκδοση είναι η λεγόμενη «σταθερή» (stable) που είναι καθαρά για χρήστες που δεν έχουν γνώσεις προγραμματισμού και αφορά την τελευταία έκδοση του βιβλίου των Witten, Frank and Hall(2011)
- Ενώ η δεύτερη έκδοση είναι αυτή που αφορά χρήστες που έχουν γνώσεις προγραμματισμού. Η δεύτερη έκδοση πρέπει να αναφέρουμε ότι χρησιμοποιείτε κυρίως από την κοινότητα των προγραμματιστών του Weka για την καλύτερη λειτουργία του και την εξέλιξη του.

Οι χρήστες έχουν την δυνατότητα να προμηθευτούν και να εγκαταστήσουν το Weka μέσα από τη σελίδα <http://www.cs.waikato.ac.nz/ml/weka/downloading.html> που είναι διαθέσιμες επιλογές για διάφορα λειτουργικά συστήματα όπως για παράδειγμα (Windows, Mac, OS X και Linux).

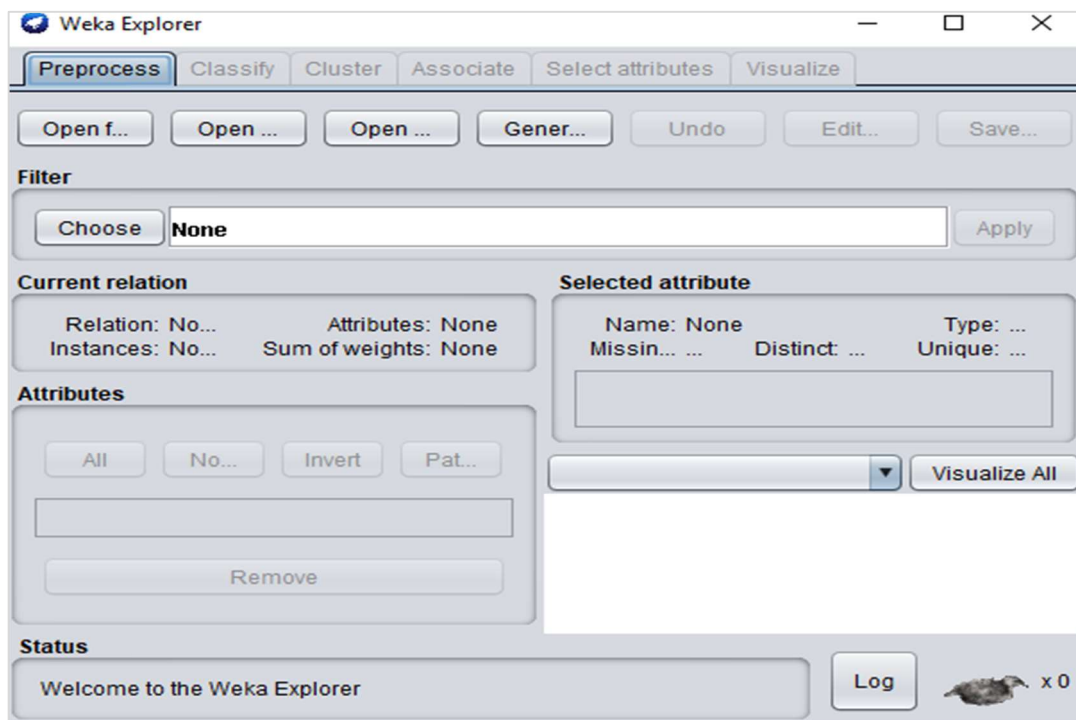
3.2 Το περιβάλλον Weka

Κατά το άνοιγμα του προγράμματος Weka όπως εμφανίζεται στην Εικόνα10 ο χρήστης έχει τη δυνατότητα να ξεκινήσει με μια από τις κύριες εφαρμογές του που είναι ο Explorer, Experimenter, Knowledge-Flow, Workbench, Simple CLI. Εμείς θα ασχοληθούμε με την πιο γνωστή εφαρμογή του Weka τον Explorer. Όπως προαναφέραμε είναι μια από τις κύριες εφαρμογές του, ο χρήστης έχει τη δυνατότητα να πραγματοποιήσει όλες τις διεργασίες της Εξόρυξης Δεδομένων.



Εικόνα 3.2: Το περιβάλλον Weka

Έτσι λοιπόν κατά την εκκίνηση του προγράμματος και με την επιλογή του Explorer μας εμφανίζεται ένα παράθυρο με έξι επιλογές για κατηγοριοποίηση, παλινδρόμηση, ανάλυση συστάδων, ανακάλυψη κανόνων συσχέτισης, προ επεξεργασία δεδομένων και οπτικοποίηση.



Εικόνα 3.3: Weka-Explorer

Καθώς επιλέξουμε την λειτουργία Weka-Explorer το πρώτο που πρέπει να κάνουμε είναι να εισάγουμε τα δεδομένα μας. Ένας από τους πιο συνηθισμένους τρόπους

εισαγωγής αρχείων στο Weka είναι με τη δημιουργία ενός αρχείου ARFF. Να συμπληρώσουμε ότι ο χρήστης μπορεί επίσης να εισάγει δεδομένα μέσα από ένα url ή μια βάση δεδομένων sql. Τα αρχεία ARFF είναι αρχεία κειμένου που απλά όλα τα στοιχεία είναι διαχωρισμένα με κόμμα. Κάθε τέτοιο αρχείο πρέπει να περιέχει και μια επικεφαλίδα ορίζοντας το όνομα τις σχέσης.

```
@relation heart-disease-simplified

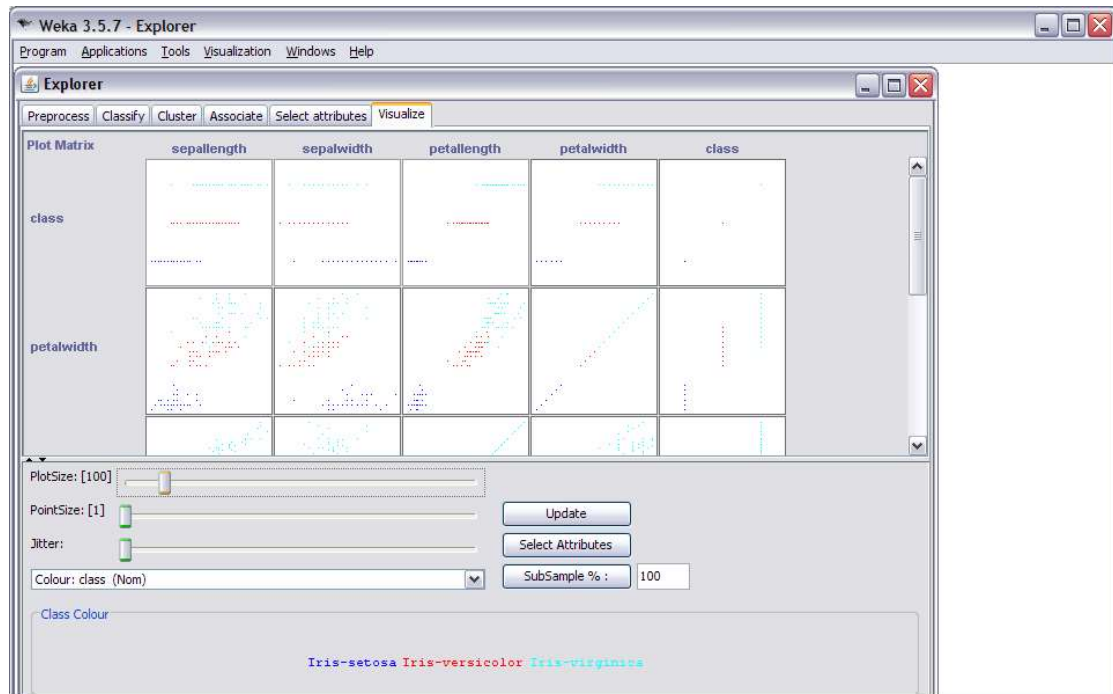
@attribute age:numeric
@attribute sex: { female, male}
@attribute chest_pain_type: { typ_angina, asympt, non_anginal, atyp_angina}
@attribute cholesterol:numeric
@attribute exercise_induced_angina: { no, yes}
@attribute class: { present, not_present}

@data
63,male,typ_angina,233,no,not_present
67,male,asympt,286,yes,present
67,male,asympt,229,yes,present
38,female,non_anginal,?,no,not_present
```

Εικόνα 3.4: Παράδειγμα Αρχείου Arff

Όπως παρατηρούμε στην αρχή του αρχείου υπάρχει η λέξη @relation και στη συνέχεια ακολουθεί το όνομα του αρχείου. Παρακάτω θα πρέπει να δηλώσουμε τα πεδία για να γίνει αυτό χρειάζεται να υπάρχει η λέξη @attribute και στη συνέχεια το όνομα του κάθε πεδίου και ο τύπος του πεδίου ανάλογα αν είναι αριθμητικό πρέπει να βάλουμε τη λέξη numeric ενώ αν είναι ονομαστικό μέσα σε αγκύλες θα πρέπει να δηλώσουμε τις ανάλογες τιμές. Στο παραπάνω παράδειγμα που φαίνεται στην Εικόνα 11 πριν από το όνομα του πίνακα (heart-diseases-simplified) υπάρχει η λέξη @relation ενώ στη συνέχεια ακολουθεί η δήλωση των πεδίων με @attribute και δίπλα το όνομα και ο τύπος του πεδίου. Τα πεδία που έχουμε στο παραπάνω παράδειγμα είναι τα εξής, @attribute age, @attribute sex, @attribute chest pain, @attribute cholesterol, @attribute exercise induced angina, @attribute class. Αφού λοιπόν δηλώσουμε το όνομα του πίνακα τα πεδία το όνομα τους και το τύπο τους ακολουθούν τα δεδομένα μας τα οποία διαχωρίζονται με κόμμα αφού πρώτα βάλουμε

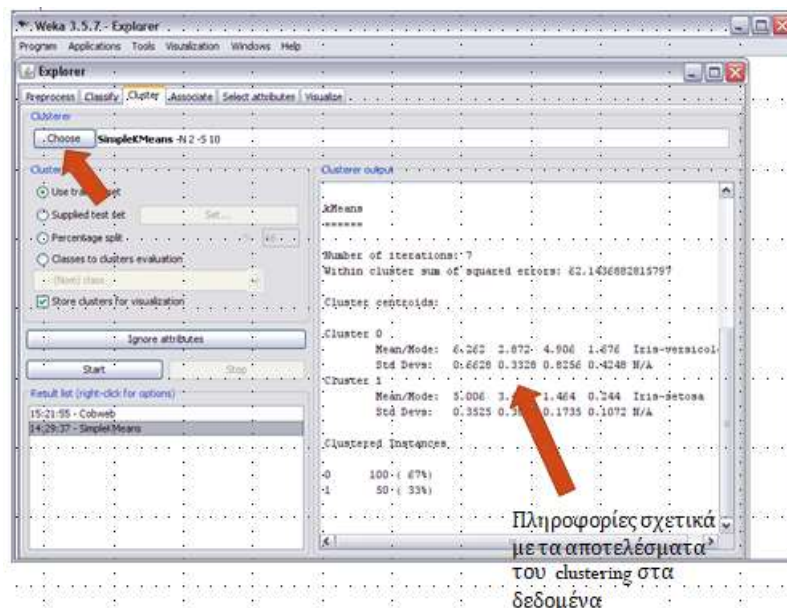
τη λέξη @data. Μέσα από την επιλογή Visualize μας δίνεται η δυνατότητα για γραφική αναπαράσταση κάθε γνωρίσματος σε σχέση με κάθε άλλο γνώρισμα. Αυτή είναι και η λεγόμενη Οπτικοποίηση των Δεδομένων.



Εικόνα 3.5: Οπτικοποίηση δεδομένων

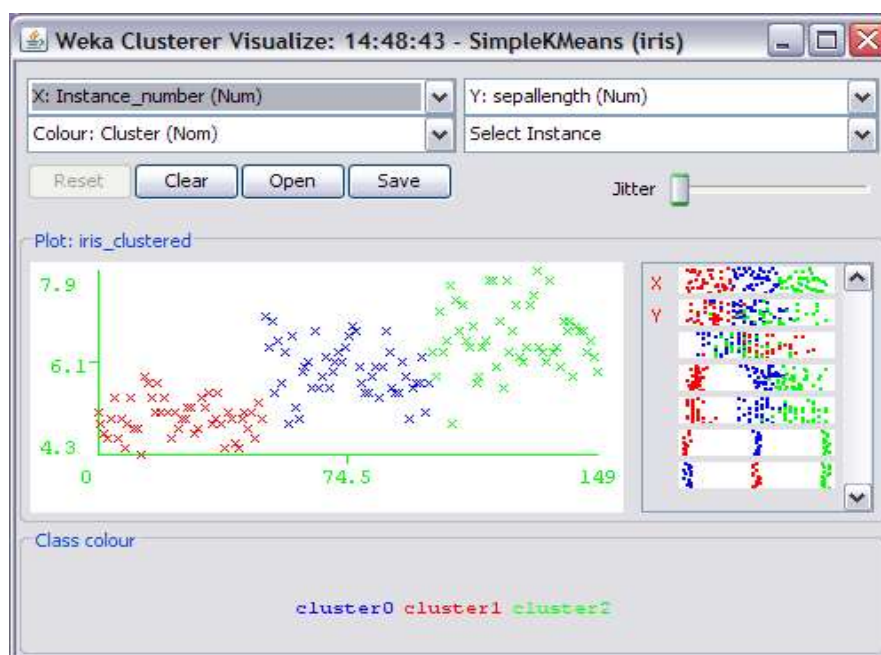
Εφόσον έχουμε επιλέξει ένα σύνολο δεδομένων μπορεί πλέον να γίνει η ομαδοποίηση-συσταδοποίηση δηλαδή η εύρεση ίδιων ομαδοποιημένων δεδομένων. Πατώντας το tab cluster επιλέγεται ένας αλγόριθμος που με την βοήθεια του θα επιτευχθεί η συσταδοποίηση και επιλέγοντας το κουμπί start θα αρχίσει η εκτέλεση του αλγορίθμου. Παρακάτω θα αναφέρουμε κάποιους αλγορίθμους συσταδοποίησης:

- Cobweb (ιεραρχική συσταδοποίηση)
- DBScan
- EM
- Farthest First
- OPTICS
- SimpleKmeans
- Xmeans



Εικόνα 3.6: Classified-Weka-Explorer

Οι τιμές των παραμέτρων κάθε αλγορίθμου συσταδοποίησης μπορούν να τροποποιηθούν κάνοντας διπλό κλικ πάνω στο όνομα του αλγορίθμου. Για την καλύτερη οπτικοποίηση των αποτελεσμάτων κάνοντας διπλό κλικ πάνω στο όνομα του αλγορίθμου που έχει είδη εκτελεστεί στα δεδομένα από το (result list) εμφανίζεται ένα βοηθητικό μενού με το οποίο μπορεί να εμφανιστεί η γραφική παράσταση των δεδομένων ανάλογα με τις ομάδες που έχουν προκύψει από το clustering [21].



Εικόνα 3.7:Οπτικοποίηση δεδομένων(2)

ΚΕΦΑΛΑΙΟ 4

Υλοποίηση Πειραμάτων

4.1 Εξαγωγή χαρακτηριστικών

Για την καλύτερη εξαγωγή χαρακτηριστικών ενός προβλήματος μηχανικής μάθησης δημιουργήσαμε το παρακάτω ηλεκτρονικό ερωτηματολόγιο. Το ερωτηματολόγιο είναι διαχωρισμένο σε τέσσερις κατηγορίες ερωτήσεων στο οικογενειακό περιβάλλον το οποίο μεγάλωσε το παιδί, στο σχολικό περιβάλλον, στη περιοχή την οποία μεγάλωσε και στα χαρακτηριστικά του υπεύθυνου καθηγητή. Το ερωτηματολόγιο διαμορφώθηκε έτσι ώστε οι απαντήσεις να είναι ομαδοποιημένες για την καλύτερη εξαγωγή χαρακτηριστικών μέσω της μηχανικής μάθησης.

ΕΡΩΤΗΜΑΤΟΛΟΓΙΟ

1. ΦΥΛΟ

- α) Άνδρας
- β) Γυναίκα

2. ΗΛΙΚΙΑ

- α) 15-18
- β) 18-22
- γ) 22-25
- δ) >25

3. ΕΘΝΙΚΟΤΗΤΑ

- α) Έλληνας
- β) Άλλο

ΟΙΚΟΓΕΝΕΙΑ

1. ΕΧΕΤΕ ΜΕΓΑΛΩΣΕΙ ΣΕ ΜΟΝΟΓΟΝΕΙΚΗ ΟΙΚΟΓΕΝΕΙΑ;

- α) Ναι
- β) Όχι

2. ΠΩΣ ΘΑ ΧΑΡΑΚΤΗΡΙΖΑΤΕ ΤΗ ΣΧΕΣΗ ΜΕΤΑΞΥ ΤΟΥ ΠΑΤΕΡΑ ΚΑΙ ΤΗΣ

ΜΗΤΕΡΑΣ ΣΑΣ ;

- α) Πολύ καλή
- β) Καλή
- γ) Μέτρια
- δ) Άσχημη

3. ΠΟΣΑ ΑΔΕΡΦΙΑ ΕΧΕΤΕ;

- α) 1
- β) 2
- γ) 3
- δ) +4

4. ΠΩΣ ΘΑ ΣΚΙΑΓΡΑΦΟΥΣΑΤΕ ΤΙΣ ΣΧΕΣΕΙΣ ΣΑΣ ΜΕ ΤΑ ΥΠΟΛΟΙΠΑ ΜΕΛΗ ΤΗΣ ΟΙΚΟΓΕΝΕΙΑΣ ΣΑΣ;

- α) Πολύ καλή
- β) Καλή
- γ) Μέτρια
- δ) Άσχημη

5. ΠΟΣΟ ΑΥΣΤΗΡΟΙ ΕΙΝΑΙ Ή ΗΤΑΝ ΟΙ ΓΟΝΕΙΣ ΣΑΣ ΜΕ ΕΣΑΣ ΑΛΛΑ ΚΑΙ ΜΕ ΤΑ ΥΠΟΛΟΙΠΑ ΜΕΛΗ ΤΗΣ ΟΙΚΟΓΕΝΕΙΑΣ;

- α) Ήταν αυστηροί αλλά δεν ασκούσαν βία
- β) Ασκούσαν σωματική βία
- γ) Δεν ήταν αυστηροί

6. ΠΩΣ ΘΑ ΧΑΡΑΚΤΗΡΙΖΑΤΕ ΤΗΝ ΟΙΚΟΜΙΚΗ ΚΑΤΑΣΤΑΣΗ ΤΟΥ ΟΙΚΟΓΕΝΕΙΑΚΟΥ ΣΑΣ ΠΕΡΙΒΑΛΛΟΝΤΟΣ;

- α) Ευκατάστατη
- β) Μεσαία
- γ) Δυσχερής

7. ΠΩΣ ΘΑ ΧΑΡΑΚΤΗΡΙΖΑΤΕ ΤΗΝ ΚΟΙΝΩΝΙΚΗ ΤΑΞΗ ΤΗΣ ΟΙΚΟΓΕΝΕΙΑΣ ΣΑΣ;

- α) Αστική
- β) Εργατική

ΣΧΟΛΕΙΟ-ΕΚΠΑΙΔΕΥΣΗ

1. ΥΠΗΡΞΕ ΠΟΤΕ ΚΑΠΟΙΟ ΚΡΟΥΣΜΑ ΕΚΦΟΒΙΣΜΟΥ ΣΤΗ ΤΑΞΗ ΣΑΣ;

- α) Ναι

β) Όχι

2.ΠΩΣ ΑΣΚΕΙΤΑΙ Ο ΕΚΦΟΒΙΣΜΟΣ ΣΕ ΕΝΑ ΔΙΑΦΟΡΕΤΙΚΟ ΑΤΟΜΟ ΣΤΗ ΤΑΞΗ ΣΑΣ;

- α) Μεμονωμένα από ένα μαθητή σε έναν άλλο
- β) Από μια ομάδα μαθητών σε έναν μαθητή
- γ) Από μια παρέα μαθητών σε μια άλλη παρέα

3.ΕΙΣΤΕ ΑΠΟΦΟΙΤΟΣ

- α) Υποχρεωτικής εκπαίδευσης
- β) Δευτεροβάθμιας εκπαίδευσης
- γ) Τριτοβάθμια εκπαίδευσης
- δ) Μεταπτυχιακής εκπαίδευσης

4.ΣΕ ΤΙ ΣΧΟΛΕΙΟ ΦΟΙΤΗΣΑΤΕ;

- α) Δημόσιο
- β) Ιδιωτικό

5.ΘΑ ΧΑΡΑΚΤΗΡΙΖΑΤΕ ΤΗΝ ΤΑΞΗ ΤΟΥ ΣΧΟΛΕΙΟΥ ΣΑΣ ΑΝΟΜΟΙΟΓΕΝΗ;

- α) Ναι
- β) Όχι

6.Η ΔΙΑΦΟΡΕΤΙΚΟΤΗΤΑ ΣΤΑ ΠΑΙΔΙΑ ΤΗΣ ΤΑΞΗΣ ΣΑΣ ΒΑΣΙΖΕΤΑΙ ΣΤΗΝ ΕΘΝΙΚΟΤΗΤΑ, ΣΕ ΙΑΤΡΙΚΑ ΠΡΟΒΛΗΜΑΤΑ Η ΣΤΗΝ ΣΩΜΑΤΙΚΗ ΔΙΑΠΛΑΣΗ;

- α) Υπάρχουν παιδιά μεταναστών;
- β) Υπάρχουν παιδιά με ιατρικά προβλήματα;
- γ) Υπάρχουν παιδιά με πρόβλημα στην εμφάνιση;

ΠΕΡΙΟΧΗ

1.ΕΧΕΤΕ ΠΑΡΑΤΗΡΗΣΕΙ ΦΑΙΝΟΜΕΝΑ ΡΑΤΣΙΣΜΟΥ-BULLYING ΣΤΗΝ ΠΕΡΙΟΧΗ ΣΑΣ;

- α) Ναι
- β) Όχι
- γ) Άλλο

2.ΠΩΣ ΘΑ ΧΑΡΑΚΤΗΡΙΖΑΤΕ ΤΗΝ ΠΕΡΙΟΧΗ ΠΟΥ ΒΡΙΣΚΕΤΑΙ Τ ΣΧΟΛΕΙΟ ΣΑΣ;

- α) Εύπορη
- β) Υποβαθμισμένη

γ) Αστική

3.Η ΠΕΡΙΟΧΗ ΠΟΥ ΒΡΙΣΚΕΤΑΙ ΤΟ ΣΧΟΛΕΙΟ ΣΑΣ ΕΙΝΑΙ ΑΝΟΜΟΙΟΓΕΝΗΣ;

α) Ναι

β) Όχι

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΠΕΥΘΥΝΟΥ ΚΑΘΗΓΗΤΗ

1.ΗΛΙΚΙΑ ΚΑΘΗΓΗΤΗ:

α) 25-30

β) 30-35

γ) 35-40

δ) >40

2.ΦΥΛΟ ΚΑΘΗΓΗΤΗ

α) Άνδρας

β) Γυναίκα

3.ΕΘΝΙΚΟΤΗΤΑ ΚΑΘΗΓΗΤΗ:

α) Έλληνας

β) Άλλο

4.ΜΟΡΦΩΤΙΚΟ ΕΠΙΠΕΔΟ:

α) Πανεπιστημιακή εκπαίδευση

β) Μεταπτυχιακή εκπαίδευση

γ) Άλλο

5.ΠΩΣ ΘΑ ΧΑΡΑΚΤΗΡΙΖΑΤΕ ΤΟΝ ΥΠΕΥΘΥΝΟ ΚΑΘΗΓΗΤΗ ΣΑΣ;

α) Αυστηρός

β) Προσιτός

γ) Αδιάφορος

6.ΠΙΣΤΕΥΕΤΑΙ ΟΤΙ Ο ΚΑΘΗΓΗΤΗΣ ΣΑΣ ΣΥΜΒΑΛΕΙ ΣΤΗΝ ΕΞΟΜΑΛΥΝΣΗ ΤΩΝ ΣΧΕΣΕΩΝ ΤΩΝ ΜΑΘΗΤΩΝ Κ ΟΤΙ ΠΡΩΩΘΕΙ ΤΟΝ ΑΛΛΗΛΟΣΕΒΑΣΜΟ;

α) Ναι

β) Όχι

ΥΠΗΡΧΑΝ ΦΑΙΝΟΜΕΝΑ BULLYING ΣΤΗ ΤΑΞΗ ΣΑΣ;

α) Ναι ένα άτομο ασκούσε bullying σε άλλα άτομα της τάξης

β) Ναι μια ομάδα ατόμων ασκούσαν bullying σε άλλα άτομα της τάξης

γ) Η πλειοψηφία της τάξης ασκούσε bullying σε 1 άτομο

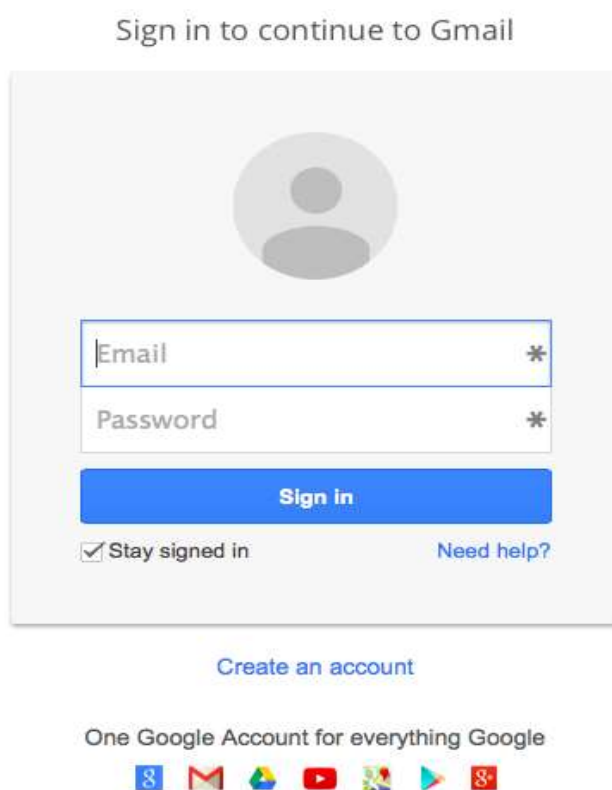
δ) Δεν υπήρχαν φαινόμενα bullying

ΥΠΗΡΧΑΝ ΑΤΟΜΙΚΑ ΦΑΙΝΟΜΕΝΑ ΑΣΚΗΣΗΣ BULLYING;

- α) Έχετε πέσει θύμα bullying
- β) Έχετε ασκήσει bullying
- γ) Τίποτα από τα δύο

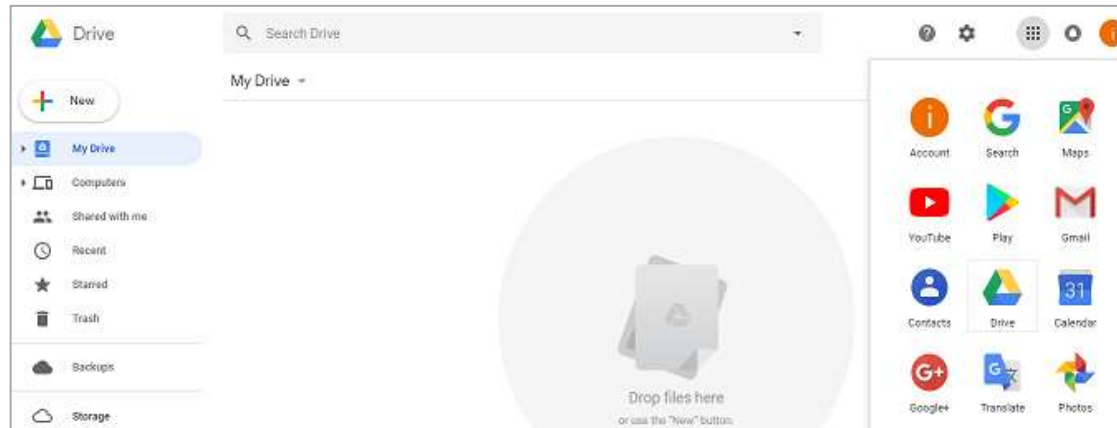
4.2 Ηλεκτρονικό ερωτηματολόγιο

Στη παρούσα πτυχιακή εργασία δημιουργήσαμε ένα ηλεκτρονικό ερωτηματολόγιο με τη βοήθεια της Google και των εργαλείων που μας παρέχει. Η Google Forms αποτελεί ένα σημαντικό και πολύ βοηθητικό εργαλείο για τη δημιουργία ερωτηματολογίων για να διεξάγουμε στατιστικές αναλύσεις και έρευνες. Αρχικά για να δημιουργήσουμε ένα τέτοιο ερωτηματολόγιο στη Google forms θα πρέπει να έχουμε κάποιο λογαριασμό στη Google ή να φτιάξουμε ένα. Στη συνέχεια αρκεί να συνδεθούμε με το λογαριασμό μας και από τις εφαρμογές που μας παρέχει η Google να επιλέξουμε την Google drive.



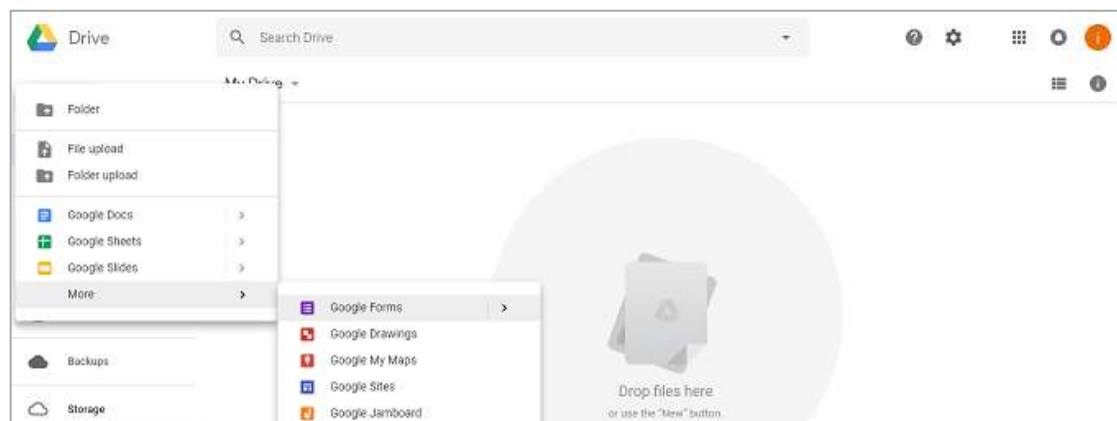
Εικόνα 4.2 : Σύνδεση σε ένα λογαριασμό Google

Στη συνέχεια αφού έχουμε δημιουργήσει ένα λογαριασμό και συνδεθούμε σε αυτόν επιλέγουμε ανάμεσα στις τόσες πολλές εφαρμογές που μας δίνει η Google την εφαρμογή Drive όπως βλέπουμε και στην εικόνα.



Εικόνα 4.3: Δημιουργία ερωτηματολογίου

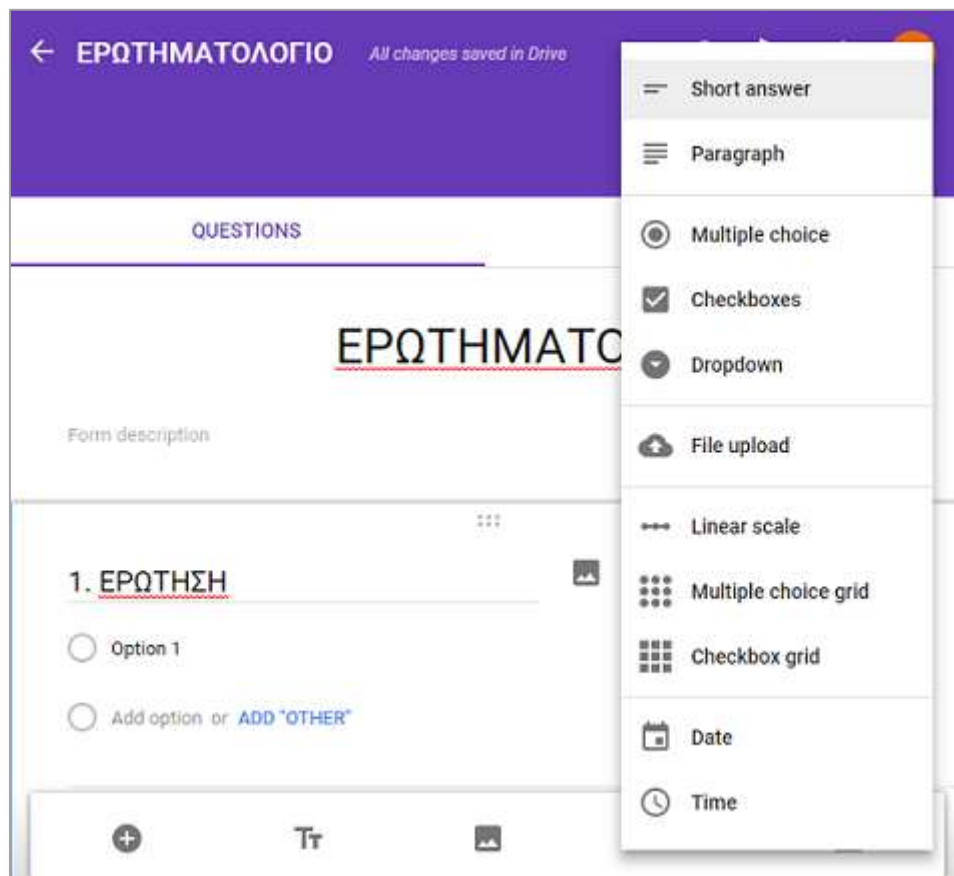
Αφού συνδεθούμε στο drive.google.com με το gmail μας επιλέγουμε την δημιουργία φόρμας και της δίνουμε το όνομα που επιθυμούμε πχ. Ερωτηματολόγιο. Όλες οι αλλαγές που κάνουμε πάνω στο ερωτηματολόγιο μας αποθηκεύονται αυτόματα.



Εικόνα 4.4 : Δημιουργία ερωτηματολογίου

Έχοντας δώσει τίτλο στο ερωτηματολόγιο μας στη συνέχεια μπορούμε να δώσουμε μια σύντομη περιγραφή. Στη συνέχεια θα αρχίσουμε να γράφουμε τις ερωτήσεις μας. Σημαντικό κομμάτι αποτελεί ο τύπος των απαντήσεων μας ανάλογα με την ερώτηση που έχουμε θέσει. Ο τύπος των απαντήσεων μας μπορεί να είναι πολλαπλής επιλογής,

κείμενο, πλαίσιο ελέγχου και άλλα. όπως βλέπουμε και στη παρακάτω εικόνα. Για την προσθήκη επόμενης ερώτησης επιλέγουμε το εικονίδιο + .



Εικόνα 4.5: Δημιουργία ερωτηματολογίου

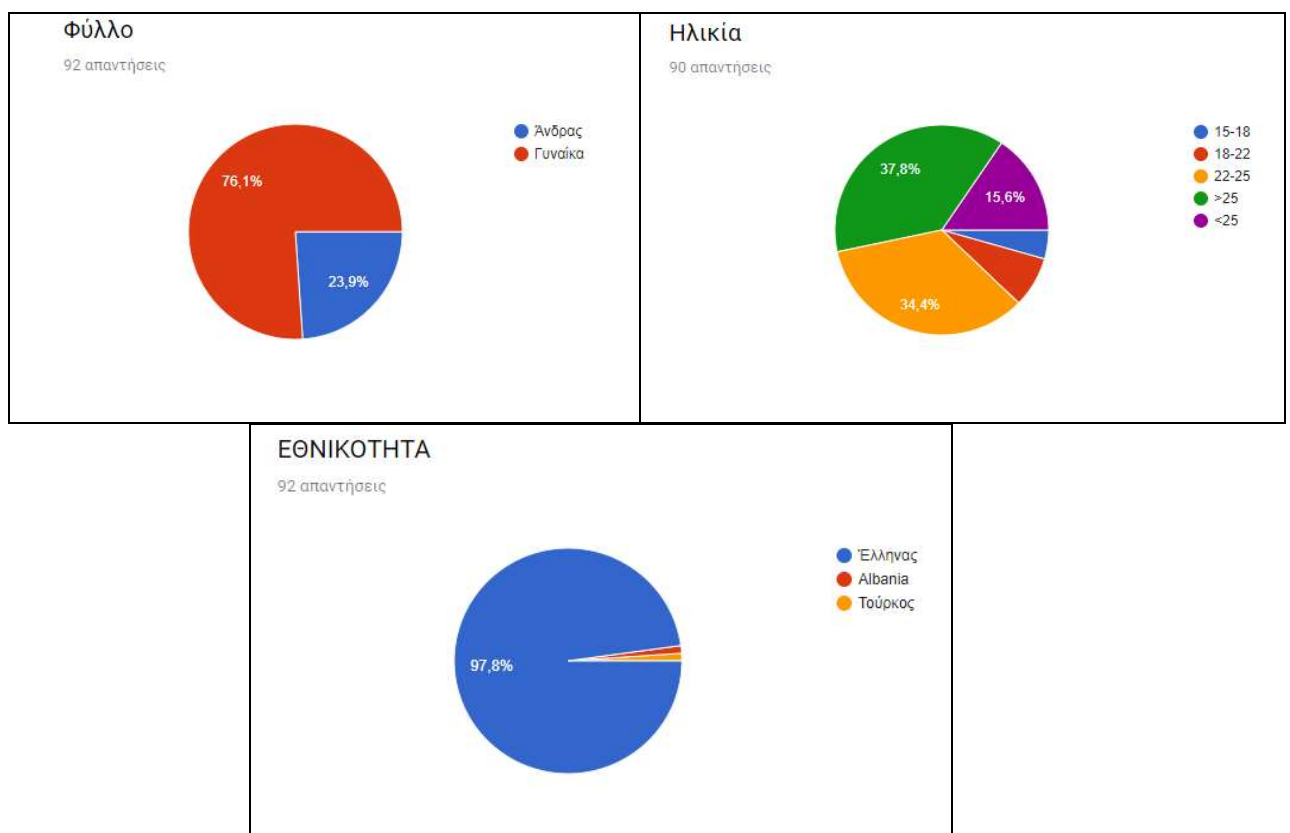
Ολοκληρώνοντας την παρακάτω διαδικασία με τις ερωτήσεις και τις πιθανές απαντήσεις που μπορεί να δώσει κάποιος το ερωτηματολόγιο μας ανέβηκε στο μέσο κοινωνικής δικτύωσης facebook. Το ερωτηματολόγιο μας αποτελείτε από 27 ερωτήσεις χωρισμένες ανά κατηγορία και απαντήθηκε από 92 χρήστες [34].

ΚΕΦΑΛΑΙΟ 5

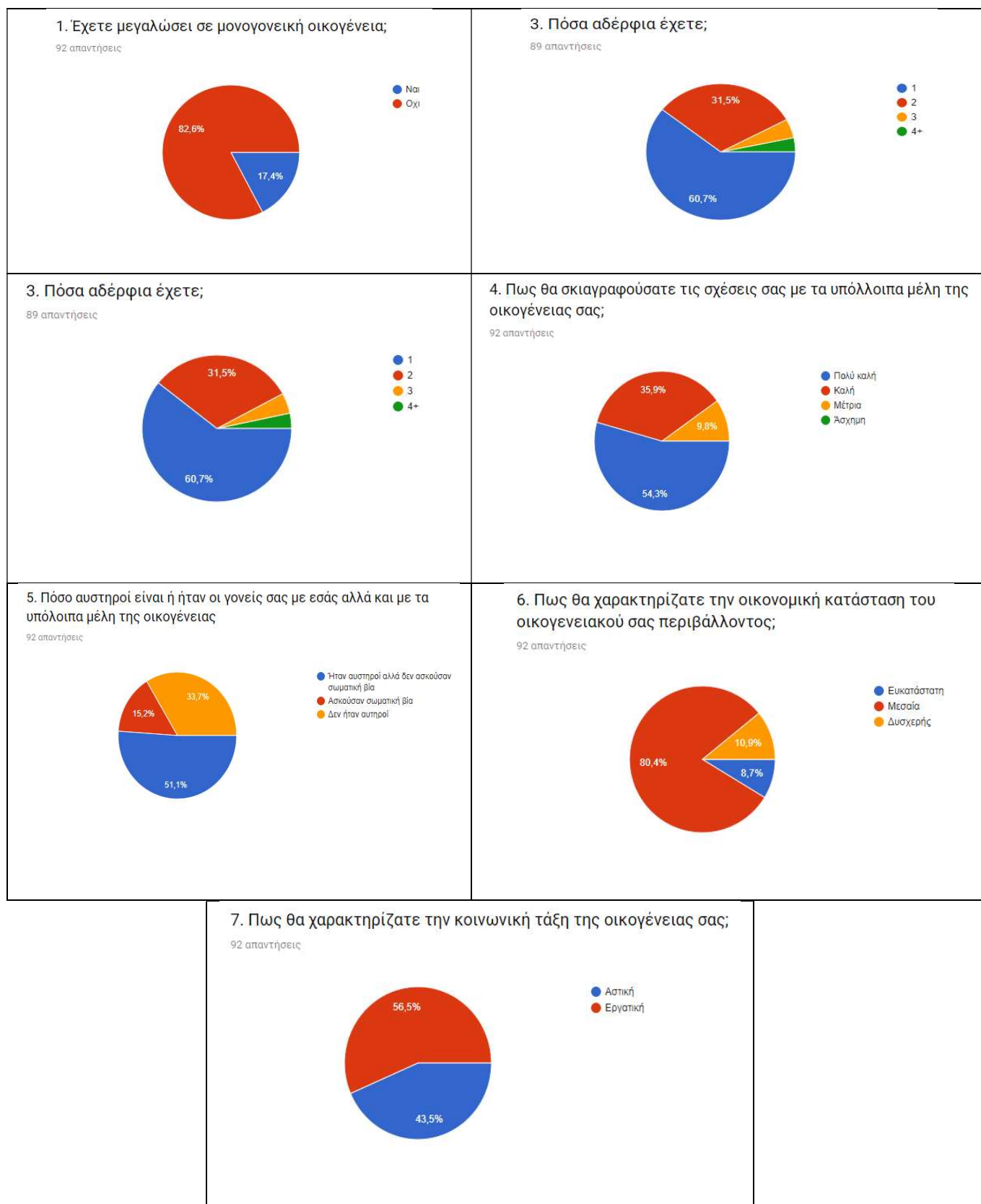
Αποτελέσματα

5.1 Στατιστικά στοιχεία δείγματος

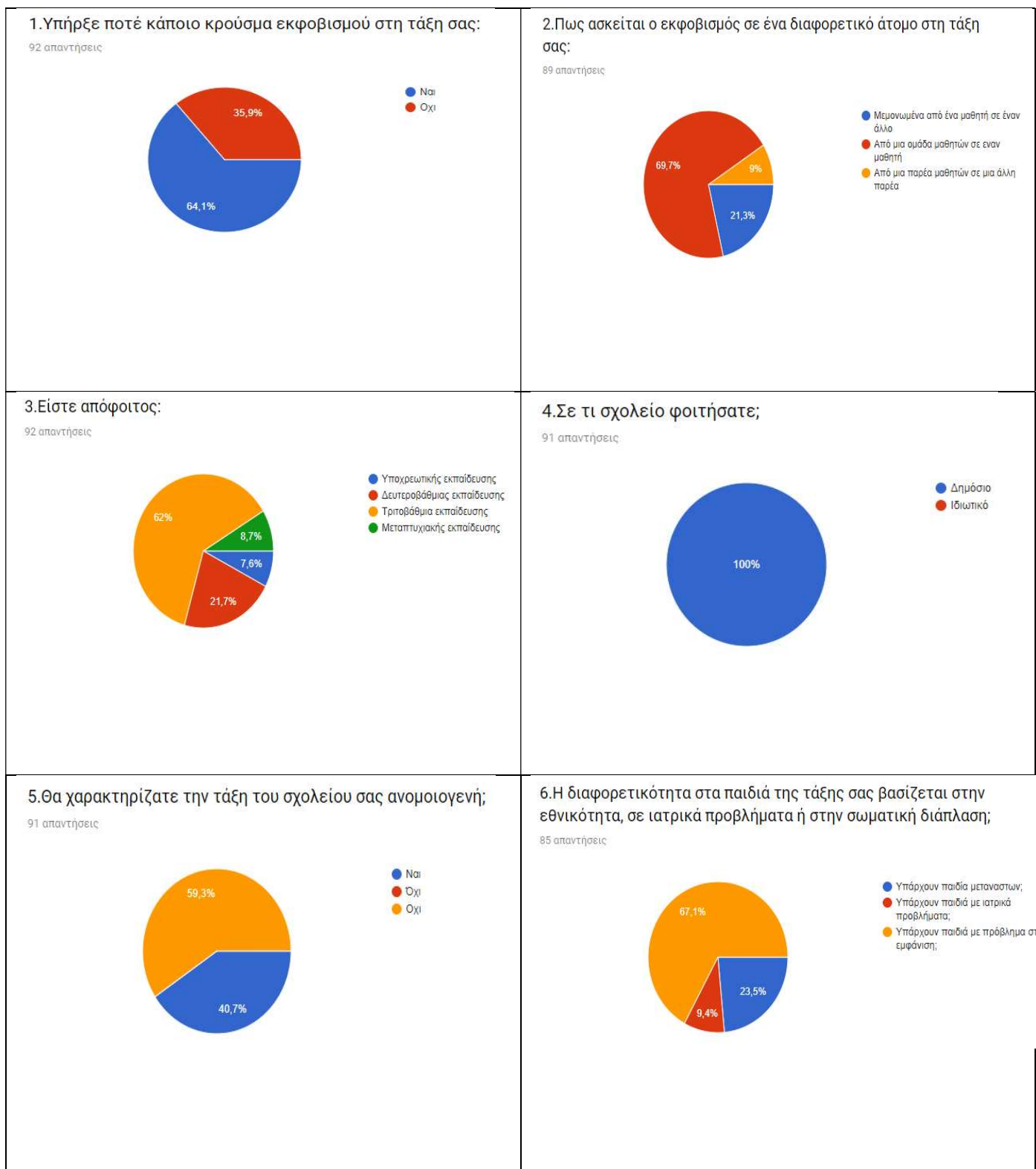
Παρακάτω θα σας παρουσιάσουμε τα δείγματα που συγκεντρώθηκαν από τις απαντήσεις δόθηκαν στο ερωτηματολόγιο της εργασίας.



Εικόνα 5.1: Γενικά στατιστικά στοιχεία



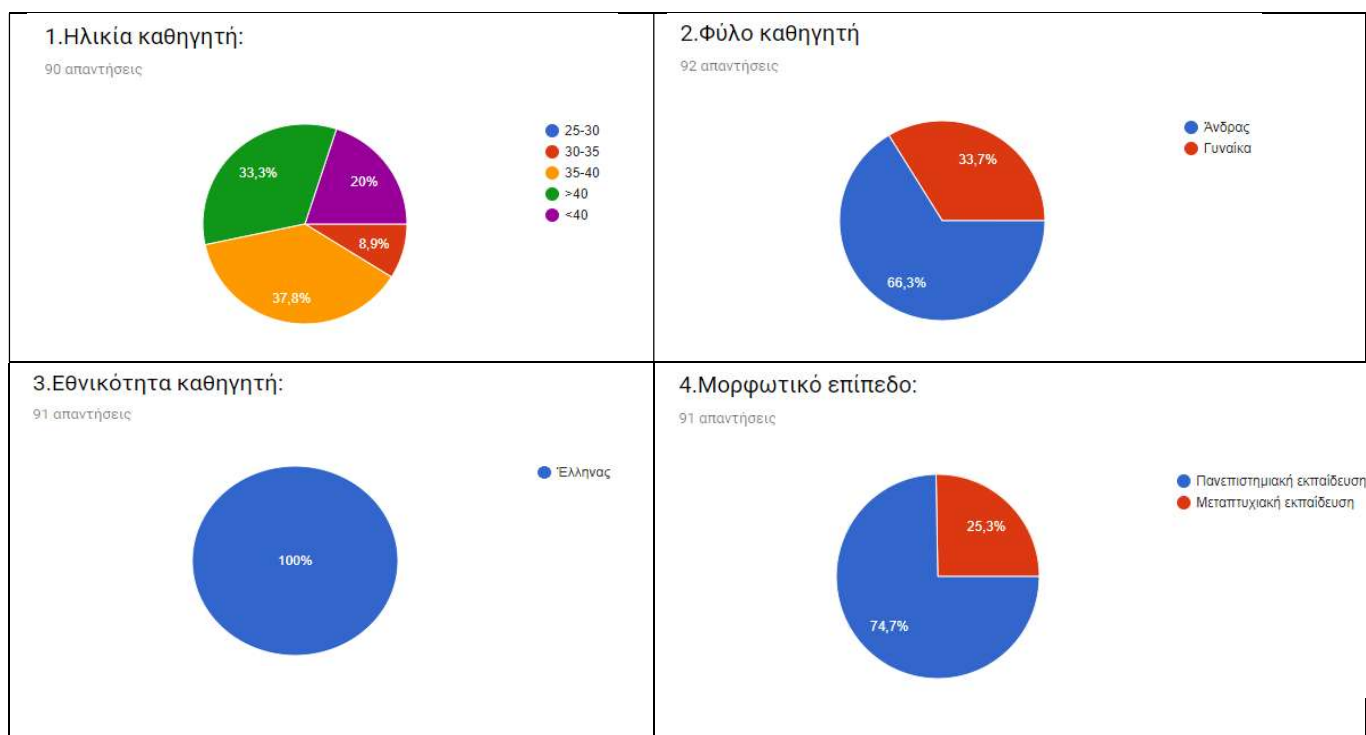
Εικόνα 5.2: Στατιστικά στοιχεία οικογένειας



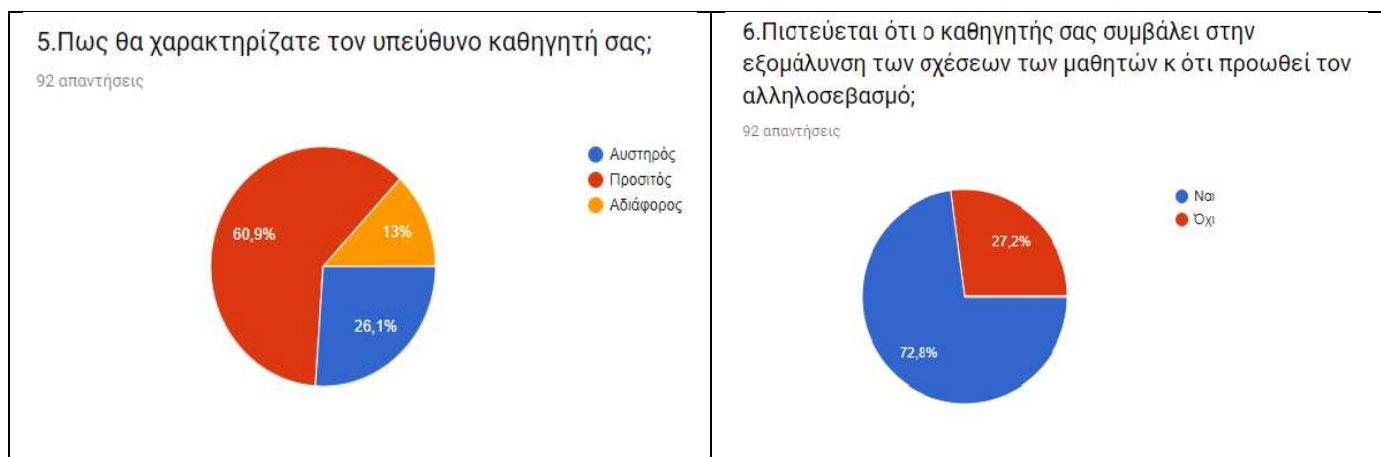
Εικόνα 5.3: Στατιστικά στοιχεία εκπαίδευσης



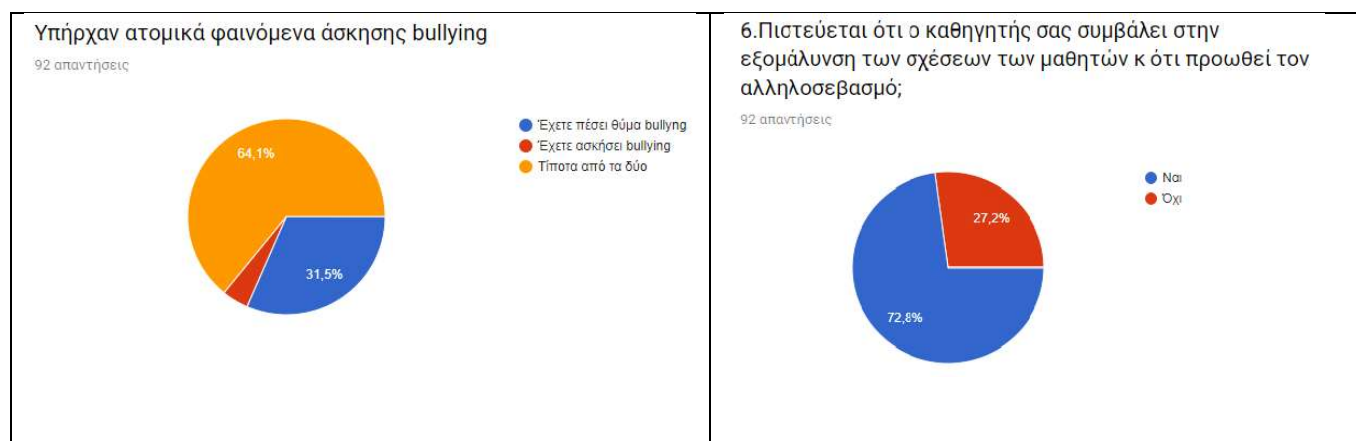
Εικόνα 5.4: Στατιστικά στοιχεία για τη περιοχή



Εικόνα 5.5: Στατιστικά στοιχεία υπεύθυνου καθηγητή



Εικόνα 5.6: Στατιστικά στοιχεία υπεύθυνου καθηγητή



Εικόνα 5.7: Στατιστικά στοιχεία δυο κλάσεων

5.2 Αποτελέσματα ταξινόμησης

Παρακάτω βλέπουμε αναλυτικά τα αποτελέσματα της ταξινόμησης για κάθε έναν αλγόριθμο 3 κλάσεων και 2 κλάσεων

5.2.1 K-κοντινότερος γείτονας (KNN = K-nearest neighbor)

```

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      172           93.4783 %
Incorrectly Classified Instances    12           6.5217 %
Kappa statistic                    0.8208
Mean absolute error                 0.0693
Root mean squared error             0.2297
Relative absolute error             18.6689 %
Root relative squared error         53.4248 %
Total Number of Instances          184

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                0,964   0,156   0,950     0,964   0,957     0,821   0,928    0,961    Yes
                0,844   0,036   0,884     0,844   0,864     0,821   0,928    0,850    No
Weighted Avg.   0,935   0,126   0,934     0,935   0,934     0,821   0,928    0,934

=== Confusion Matrix ===

  a  b  <-- classified as
134  5 |  a = Yes
  7 38 |  b = No

```

Εικόνα 5.8: K-κοντινότερος γείτονας 2 κλάσεων (ναι – όχι)

```

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      162           92.5714 %
Incorrectly Classified Instances    13           7.4286 %
Kappa statistic                    0.8423
Mean absolute error                 0.0594
Root mean squared error             0.1983
Relative absolute error             18.9811 %
Root relative squared error         50.3394 %
Total Number of Instances          175
Ignored Class Unknown Instances      9

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                0,943   0,094   0,958     0,943   0,950     0,840   0,848    0,849    N_to_1
                0,848   0,049   0,800     0,848   0,824     0,782   0,919    0,838    1_to_1
                0,950   0,006   0,950     0,950   0,950     0,944   0,999    0,995    N_to_N
Weighted Avg.   0,926   0,076   0,928     0,926   0,926     0,841   0,878    0,863

=== Confusion Matrix ===

  a  b  c  <-- classified as
115  7  0 |  a = N_to_1
  4 28  1 |  b = 1_to_1
  1  0 19 |  c = N_to_N

```

Εικόνα 5.9: K-κοντινότερος γείτονας 3 κλάσεων (N-to-1, 1_to_1, N_to_N)

5.2.2 Μπευζιανός ταξινομητής (Naive Bayes classifier)

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      147          79.8913 %
Incorrectly Classified Instances    37          20.1087 %
Kappa statistic                    0.4163
Mean absolute error                 0.2769
Root mean squared error             0.3916
Relative absolute error             74.6159 %
Root relative squared error         91.067 %
Total Number of Instances          184

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                0,899   0,511   0,845     0,899   0,871     0,421   0,765    0,895    Yes
                0,489   0,101   0,611     0,489   0,543     0,421   0,765    0,500    No
Weighted Avg.   0,799   0,411   0,787     0,799   0,791     0,421   0,765    0,799

=== Confusion Matrix ===

  a  b  <-- classified as
125 14 |  a = Yes
 23 22 |  b = No
```

Εικόνα 5.10 : Μπευζιανός ταξινομητής (Naive Bayes classifier) 2 κλάσεων (ναι - όχι)

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      121          69.1429 %
Incorrectly Classified Instances    54          30.8571 %
Kappa statistic                    0.243
Mean absolute error                 0.2451
Root mean squared error             0.3755
Relative absolute error             78.3663 %
Root relative squared error         95.3026 %
Total Number of Instances          175
Ignored Class Unknown Instances      9

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                0,861   0,679   0,745     0,861   0,798     0,211   0,617    0,752    N_to_1
                0,273   0,077   0,450     0,273   0,340     0,240   0,697    0,446    I_to_1
                0,350   0,045   0,500     0,350   0,412     0,357   0,894    0,522    N_to_N
Weighted Avg.   0,691   0,493   0,661     0,691   0,668     0,233   0,664    0,668

=== Confusion Matrix ===

  a  b  c  <-- classified as
105 11  6 |  a = N_to_1
 23  9  1 |  b = I_to_1
 13  0  7 |  c = N_to_N
```

Εικόνα 5.11: Μπευζιανός ταξινομητής (Naive Bayes classifier) 3 κλάσεων (N_to_1, I_to_1, N_to_N)

5.3.2 Δένδρα απόφασης (Decision trees)

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      163           88.587 %
Incorrectly Classified Instances    21           11.413 %
Kappa statistic                    0.6791
Mean absolute error                 0.1413
Root mean squared error             0.3175
Relative absolute error             38.0783 %
Root relative squared error         73.8395 %
Total Number of Instances          184

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                0,942    0,289    0,910     0,942    0,926      0,681    0,901     0,945     Yes
                0,711    0,058    0,800     0,711    0,753      0,681    0,901     0,774     No
Weighted Avg.   0,886    0,232    0,883     0,886    0,884      0,681    0,901     0,903

=== Confusion Matrix ===

  a  b  <-- classified as
131  8  |  a = Yes
 13 32  |  b = No
```

Εικόνα 5.12: Δένδρα απόφασης (Decision trees) 2 κλάσεων (ναι - όχι)

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      135           77.1429 %
Incorrectly Classified Instances     40           22.8571 %
Kappa statistic                    0.4716
Mean absolute error                 0.1604
Root mean squared error             0.333
Relative absolute error             51.27 %
Root relative squared error         84.5306 %
Total Number of Instances          175
Ignored Class Unknown Instances      9

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                0,893    0,453    0,820     0,893    0,855      0,474    0,854     0,902     N_to_1
                0,394    0,077    0,542     0,394    0,456      0,360    0,833     0,555     1_to_1
                0,650    0,032    0,722     0,650    0,684      0,647    0,909     0,584     N_to_N
Weighted Avg.   0,771    0,334    0,756     0,771    0,760      0,472    0,856     0,800

=== Confusion Matrix ===

  a  b  c  <-- classified as
109  9  4  |  a = N_to_1
 19 13  1  |  b = 1_to_1
  5  2 13  |  c = N_to_N
```

Εικόνα 5.13: Δένδρα απόφασης (Decision trees) 3 κλάσεων (N_to_1, 1_to_1, N_to_N)

5.4.2 Μηχανές Διανυσμάτων υποστήριξης (Support Vector Machines – SVM)

```

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      160          86.9565 %
Incorrectly Classified Instances    24          13.0435 %
Kappa statistic                    0.6416
Mean absolute error                 0.1304
Root mean squared error             0.3612
Relative absolute error             35.1439 %
Root relative squared error         83.9957 %
Total Number of Instances          184

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                0,921   0,289   0,908     0,921   0,914     0,642   0,816    0,896    Yes
                0,711   0,079   0,744     0,711   0,727     0,642   0,816    0,600    No
Weighted Avg.   0,870   0,238   0,868     0,870   0,869     0,642   0,816    0,823

=== Confusion Matrix ===

  a  b  <-- classified as
128 11 |  a = Yes
 13 32 |  b = No

```

Εικόνα 5.14: Μηχανές Διανυσμάτων υποστήριξης (Support Vector Machines – SVM) 2 κλάσεων(ναι - όχι)

```

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      135          77.1429 %
Incorrectly Classified Instances    40          22.8571 %
Kappa statistic                    0.4716
Mean absolute error                 0.1604
Root mean squared error             0.333
Relative absolute error             51.27 %
Root relative squared error         84.5306 %
Total Number of Instances          175
Ignored Class Unknown Instances      9

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                0,893   0,453   0,820     0,893   0,855     0,474   0,854    0,902    N_to_1
                0,394   0,077   0,542     0,394   0,456     0,360   0,833    0,555    1_to_1
                0,650   0,032   0,722     0,650   0,684     0,647   0,909    0,584    N_to_N
Weighted Avg.   0,771   0,334   0,756     0,771   0,760     0,472   0,856    0,800

=== Confusion Matrix ===

  a  b  c  <-- classified as
109  9  4 |  a = N_to_1
 19 13  1 |  b = 1_to_1
  5  2 13 |  c = N_to_N

```

Εικόνα 5.15: Μηχανές Διανυσμάτων υποστήριξης (Support Vector Machines – SVM) 3 κλάσεων (N_to_1, 1_to_1, N_to_N)

Τα παραπάνω πειράματα συνοψίζονται στον πίνακα 5.1, στον οποίο συγκεντρώθηκαν όλες οι τιμές ακρίβειας για όλους τους αλγορίθμους ταξινόμησης, και για τα δύο προβλήματα (2 κλάσεων και τριών κλάσεων).

	KNN	Naïve Bayes	Decision Trees	Support Vector Machine
	Acc (%)	Acc (%)	Acc (%)	Acc (%)
Πρόβλημα 2 κλάσεων	93,48	79,89	88,59	86,96
Πρόβλημα 3 κλάσεων	92,57	69,14	77,14	77,14

Πίνακας 5.1: Αποτελέσματα Ακρίβειας Ταξινόμησης

ΚΕΦΑΛΑΙΟ 6

Συμπεράσματα

Κατά την εκπόνηση της παρούσα πτυχιακής εργασίας διερευνήθηκε το θέμα του σχολικού εκφοβισμού-bullying και φαίνεται από τα παραπάνω ότι ο σχολικός εκφοβισμός είναι ένα φαινόμενο που συναντάται σε όλες τις εκφάνσεις της σχολικής ζωής ενός παιδιού που μπορεί να έχει ως αποτέλεσμα είτε ψυχικές είτε σωματικές επιπτώσεις, αλλά ακόμη και το θάνατο του θύματος. Στην ερώτηση εάν υπήρξε κάποιο κρούσμα εκφοβισμού στη τάξη σας το ποσοστό των 64,1% των απαντήσεων δόθηκε θετική απάντηση ενώ στην ερώτηση πως ασκείται ο εκφοβισμός το ποσοστό των 69,7% των ερωτηθέντων απάντησαν πως ο εκφοβισμός – bullying γινόταν από μια ομάδα ατόμων προς ένα άτομο. Σε αυτό το σημείο θα πρέπει να πούμε το πόσο σημαντικός είναι ο ρόλος του εκπαιδευτικού. Έτσι λοιπόν όταν ο εκπαιδευτικός αντιληφθεί ένα τέτοιο περιστατικό θα πρέπει να το αξιολογήσει και να έχει τη κατάλληλη προσέγγιση λαμβάνοντας υπόψη ότι κάθε περιστατικό είναι διαφορετικό.

Σκοπός της παρούσας πτυχιακής εργασίας ήταν μέσω της χρήσης τεχνικών μηχανικής μάθησης να διερευνήσουμε τους παράγοντες που επηρεάζουν την ανάπτυξη του φαινομένου του σχολικού εκφοβισμού – bullying. Η μηχανική μάθηση αφορά τη δημιουργία μοντέλων (προτύπων) με χρήση συνόλου δεδομένων από ένα υπολογιστικό σύστημα. Μελετήθηκε η εξόρυξη δεδομένων και βασικές κατηγορίες αλγορίθμων, όπως η ταξινόμηση η ομαδοποίηση/ συσταδοποίηση και η εξαγωγή και ανάλυση συσχετίσεων. Κάποιες από τις πιο σημαντικές μεθόδους ταξινόμησης που μελετήθηκαν ήταν ο K-Κοντινότερος Γείτονας, ο Μπευζιανός Ταξινομητής, τα Δέντρα Αποφάσεων και οι Μηχανές Διανυσμάτων Υποστήριξης. Παρουσιάστηκε επίσης το λογισμικό για την μάθηση και την εξόρυξη δεδομένων WEKA, το οποίο χρησιμοποιήσαμε για την εφαρμογή των τεχνικών ταξινόμησης που αναφέρθηκαν ανωτέρω. Το συγκεκριμένο λογισμικό είναι ανοιχτού κώδικα, δίνει πολλές δυνατότητες στους χρήστες, διαθέτει περιβάλλον φιλικό προς το χρήστη και μπορεί να εγκατασταθεί πολύ εύκολα. Για τις ανάγκες της εργασίας, δημιουργήθηκε ερωτηματολόγιο με σκοπό να σκιαγραφηθούν τα χαρακτηριστικά των ερωτηθέντων. Οι ερωτήσεις που συγκεντρώθηκαν διαχωρίστηκαν ανάλογα με το οικογενειακό

περιβάλλον, την περιοχή που μεγάλωσε ο καθένας, το σχολικό περιβάλλον και τα χαρακτηριστικά του υπεύθυνου καθηγητή της τάξης.

Το ερωτηματολόγιο απαντήθηκε από 92 άτομα διαφορετικών ηλικιών και διαφορετικής εθνικότητας. Οι απαντήσεις κατηγοριοποιήθηκαν σε τρεις κατηγορίες δηλαδή N_to_1 (από μια ομάδα ατόμων προς ένα άτομο), 1_to_1 (από ένα άτομο προς ένα άλλο άτομο), N_to_N (από μια ομάδα ατόμων προς μια άλλη ομάδα ατόμων) και σε δυο κατηγορίες ναι και όχι. Ως προς την μηχανική μάθηση έγιναν πειράματα δύο και τριών κλάσεων με σκοπό, από τις απαντήσεις των συγκεκριμένων ερωτήσεων να προβλέπεται αν είναι το σχολικό περιβάλλον ικανό να αναπτύξει φαινόμενα εκφοβισμού, και τι είδους. Από τα αποτελέσματα, καταδεικνύεται ότι υπάρχει δυνατότητα πρόβλεψης καθώς επιτυγχάνεται ποσοστό ακρίβειας το οποίο υπερβαίνει τον 90%. Ο αλγόριθμος ο οποίος αποδίδει καλύτερα είναι ο ταξινομητή του K-κοντινότερου γείτονα, γεγονός το οποίο είναι ένδειξη ότι το πρόβλημα έχει ισχυρά χαρακτηριστικά τοπικότητας.

Συμπερασματικά, γίνεται φανερό ότι θα μπορούσε ένα εκπαιδευμένο σύστημα βασισμένο σε μια σειρά ερωτήσεων, να συμβουλέψει το εκπαιδευτικό προσωπικό για την πιθανότητα ανάπτυξης φαινομένων εκφοβισμού, με σκοπό την αναδιοργάνωση των σχολικών τμημάτων με τέτοιο τρόπο για να προλαμβάνονται τα φαινόμενα αυτά.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1].ABC EUROPE'S ANTIBULLYING CAMPAIGN. Άρθρο. Διαθέσιμο στο:
<http://www.e-abc.eu/gr/sholikos-ekfovismos/>
- [2].Μαρίνα Μόσχα. Άρθρο: Μορφές σχολικού εκφοβισμού – Bullying. Διαθέσιμο στο:
https://www.askitis.gr/psychichealth/view/morphes_scholikai_ekphovismoi_bullying
- [3].Μαρίνα Μόσχα. Άρθρο.Σχολική Βία και Σχολικός Εκφοβισμός (bullying). Διαθέσιμο στο:
https://www.askitis.gr/psychichealth/view/panellinia_imeras_kata_tis_scholikas_vias_kai_toi_ekphovismoi
- [4].Κωνσταντοπούλου Φρίντα(Ψυχίατρος παιδιών και εφήβων)[2015]. Άρθρο. Σχολικός εκφοβισμός: Αίτια και συνέπειες του bullying. Διαθέσιμο στο:
<https://www.iatronet.gr/ygeia/paidiatriki/article/32729/scholikos-ekfovismos-aitia-kai-synepeies-toy-bullying.html>
- [5].Κωνσταντία Παπαδάκη(καθηγήτρια - φιλόλογος)[2017].Άρθρο. Mobbing και Bullying στα σχολεία. Διαθέσιμο στο: <http://www.haniotika-nea.gr/mobbing-ke-bullying-sta-scholia/>
- [6].Άννα Κουρουμπλή [2013]. Άρθρο. Ο σχολικός εκφοβισμός και η αντιμετώπιση του. Διαθέσιμο στο: <http://www.pemptousia.gr/2013/02/o-scholikos-ekfovismos-ke-i-antimetop/>
- [7].Υπουργείο Εθνικής Παιδείας και Θρησκευμάτων Παιδαγωγικό Ινστιτούτο: Επιθετικότητα στο σχολείο. Προτάσεις για πρόληψη και αντιμετώπιση.
<http://www.pi-schools.gr/download/lessons/social/entypa/epithet1.pdf>
- [8].ΕΡΤ[2016].Άρθρο. Παρατηρητές μια σιωπηρή δύναμη. Διαθέσιμο στο:
<https://www.ert.gr/eidiseis/ellada/paratirites-mia-siopiri-dinami/>
- [9].Κέντρο Ευρωπαϊκά Συνταγματικού Δικαίου. Επιμορφωτικό Υλικό για Στελέχη Εκπαίδευσης και Εκπαιδευτικούς. Διαθέσιμο στο: <http://stop-bullying.sch.gr/wp-content/uploads/2015/10/ekpaideutikos.pdf>
- [10]. Μηχανική Μάθηση. Διαθέσιμο στο:
<http://aibook.csd.auth.gr/include/slides/Chap18.pdf>

- [11]. N Giannakeas, DI Fotiadis, “Image processing and machine learning techniques for the segmentation of cDNA microarray images, Machine Learning: Concepts, Methodologies, Tools and Applications, 817-829, 2012.
- [12]. Εξόρυξη δεδομένων. Διαθέσιμο στο: https://el.wikipedia.org/wiki/%CE%95%CE%BE%CF%8C%CF%81%CF%85%CE%BE%CE%B7_%CE%B4%CE%B5%CE%B4%CE%BF%CE%BC%CE%AD%CE%BD%CF%89%CE%BD
- [13]. V Christou, MG Tsipouras, N Giannakeas, AT Tzallas, “Hybrid Extreme Learning Machine Approach for Homogeneous Neural Networks”, Neurocomputing 311, 397-412, 2018.
- [14]. I. Tsoulos, A. T. Tzallas, D. G. Tsalikakis, N. Giannakeas, M. G. Tsipouras, I. Androulidaki, and E. Zaitseva, “NeuralGenesis: a software for distributed neural network training,” In Proceeding of the 24th Telecommunications Forum (TELFOR), Beograd, Serbia, 2016
- [15]. Εισαγωγή στην εξόρυξη δεδομένων. Διαθέσιμο στο: https://repository.kallipos.gr/bitstream/11419/2966/1/02_chapter_01.pdf
- [16]. Πίτουρας Εξόρυξη δεδομένων. Διαθέσιμο στο: <http://www.cs.uoi.gr/~pitoura/courses/dm/introspring11.pdf>
- [17]. N Giannakeas, PS Karvelis, TP Exarchos, FG Kalatzis, DI Fotiadis, “Segmentation of microarray images using pixel classification—Comparison with clustering-based methods, “Computers in biology and medicine 43 (6), 705-716, 2013.
- [18]. A. T. Tzallas, N. Giannakeas, K. Zoulis, M. G. Tsipouras, E. Glavas, K. D. Tzimourta, L. G. Astrakas and S. Konitsiotis, “EEG Classification and Short-Term Epilepsy Prognosis using Brain Computer Interface Software,” In Proc. of 30th IEEE International Symposium on Computer-Based Medical Systems (CBMS), Thessaloniki, Greece, 2017.
- [19]. K.D. Tzimourta, A. Tsilimbaris, K. Tzioukalia, A.T. Tzallas, M.G. Tsipouras, L.G. Astrakas, N. Giannakeas, EEG-based automatic sleep stage classification, Biomed J Sci Techn Res (BJSTR), vol. 7(4), pp. 1-6, 2018.
- [20]. N Giannakeas, DI Fotiadis, “An automated method for gridding and clustering-based segmentation of cDNA microarray images,” Computerized Medical Imaging and Graphics 33 (1), 40-49, 2009.

- [21]. N. Giannakeas, M.G. Tsipouras, A.T. Tzallas, K. Kyriakidi, Z.E. Tsianou, P. Manousou, A. Hall, E.C. Karvounis, V. Tsianos, E. Tsianos, “ in Proc of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3097-3100, 2015.
- [22]. Yorgos Goletsis, Themis P Exarchos, Nikolaos Giannakeas, Dimitrios I Fotiadis, “Intelligent patient profiling for diagnosis, staging and treatment selection in colon cancer,” 8th IEEE International Conference on BioInformatics and BioEngineering, 1-6, 2008.
- [23]. Tavish Srivastava[2018].Introduction to k-Nearest Neighbors. Διαθέσιμο στο: <https://www.analyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering/>
- [24]. Sunil Ray[2017]. 6 Easy Step to Learn Naïve Bayes Algorithm. Διαθέσιμο στο: <https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/>
- [25]. N Giannakeas, DI Fotiadis, “Multichannel segmentation of cDNA microarray images using the Bayes classifier,” in Proc of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3466-3469, 2007.
- [26]. Wikipedia. Naive Bayes Classifier. Διαθέσιμο στο: https://en.wikipedia.org/wiki/Naive_Bayes_classifier
- [27]. Mayur Kulkarni[2017]. Decision Trees for Classification. Διαθέσιμο στο: <https://www.xoriant.com/blog/product-engineering/decision-trees-machine-learning-algorithm.html>
- [28]. D.C. Tsouros, P.N. Smyrlis, N. Giannakeas, A.T. Tzallas, M.G. Tsipouras, “Random forests with stochastic induction of decision trees, In proc of the IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI), pp. 527-531, 2018.
- [29]. Douglas Hawks. What Is a Decision Tree – Example, Advantages & Role in Management. Διαθέσιμο στο: <https://study.com/academy/lesson/what-is-a-decision-tree-examples-advantages-role-in-management.html#lesson>
- [30]. Sunil Ray[2017]. Understanding Support Vector Machine algorithm from examples. Διαθέσιμο στο: <https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/>

- [31]. Tavish Srivastava[2014]. Support Vector Machine – Simplified. Διαθέσιμο στο: <https://www.analyticsvidhya.com/blog/2014/10/support-vector-machine-simplified/>
- [32]. N Giannakeas, PS Karvelis, DI Fotiadis, “A classification-based segmentation of cDNA microarray images using support vector machines,” in Proc of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 875-878, 2008.
- [33]. Οδηγός WEKA. Διαθέσιμο στο: https://repository.kallipos.gr/bitstream/11419/1239/2/Kef._13.pdf
- [34]. Μαριάνθη Αθανασιαδου[2014]. Δημιουργία Φόρμας στο Google Drive. Διαθέσιμο στο: <https://www.slideshare.net/athama/google-drive-39575856>