

ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΗΠΕΙΡΟΥ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ Τ.Ε.



ΤΕΧΝΟΛΟΓΙΚΟ  
ΕΚΠΑΙΔΕΥΤΙΚΟ  
ΙΔΡΥΜΑ  
ΤΕΙ ΗΠΕΙΡΟΥ

### ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Λειτουργικά συστήματα κινητών συσκευών και διερεύνηση κριτηρίων  
επιλογής με χρήση Μηχανικής Μάθησης



ΓΚΙΟΚΑΣ ΙΩΑΝΝΗΣ

ΑΜ i116

Εποπτεύων καθηγητής

Γιανακκέας Νικόλαος



ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΗΠΕΙΡΟΥ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ Τ.Ε.



ΤΕΧΝΟΛΟΓΙΚΟ  
ΕΚΠΑΙΔΕΥΤΙΚΟ  
ΙΔΡΥΜΑ  
ΤΕΙ ΗΠΕΙΡΟΥ

## **ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

Λειτουργικά συστήματα κινητών συσκευών και διερεύνηση κριτηρίων  
επιλογής με χρήση Μηχανικής Μάθησης

**Γκιόκας Ιωάννης**

**AM i116**

Επιβλέπων καθηγητής

Γιαννακέας Νικόλαος

- Άρτα 2018 -



Εγκρίθηκε από τριμελή εξεταστική επιτροπή

Άρτα, Φεβρουάριος 2018

## **ΕΠΙΤΡΟΠΗ ΑΞΙΟΛΟΓΗΣΗΣ**

1. Επιβλέπων καθηγητής

Γιαννακέας Νικόλαος

2. Μέλος επιτροπής

Τζάλλας Αλέξανδρος

3. Μέλος επιτροπής

Τσίπουρας Μάρκος

ΟΠροϊστάμενος του Τμήματος

© Γκιόκας Ιωάννης , 2018.  
Με επιφύλαξη παντός δικαιώματος. Allrightsreserved.

## 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.

Γκιόκας Ιωάννης

Υπογραφή

## ΕΥΧΑΡΙΣΤΙΕΣ

---

Θα ήθελα να ευχαριστήσω θερμά τον καθηγητή μου κ. Γιαννακέα Νικόλαο για την εμπιστοσύνη που μου έδειξε, την υπομονή και τη σωστή καθοδήγηση κατά την υλοποίηση της παρούσας πτυχιακής εργασίας .



# ΠΕΡΙΛΗΨΗ

---

Η παρούσα εργασία μελετά ποιοτικά τις επιλογές των λειτουργικών συστημάτων σε φορητές συσκευές. Επίσης γίνεται χρήση μεθόδων ταξινόμησης με σκοπό να διερευνηθούν τα κριτήρια επιλογής των λειτουργικών συστημάτων από τους χρήστες.

Στο πρώτο κεφάλαιο της παρούσας πτυχιακής εργασίας θα περιγράψουμε τα βασικά λειτουργικά συστήματα. Πιο συγκεκριμένα το Android, το Ios και το Windowsmobile και θα κάνουμε μια σύντομη ιστορική αναδρομή στα λειτουργικά συστήματα και τα στατιστικά στοιχεία για τη διάχυση τους στην αγορά. Το κεφάλαιο θα κλείσει με την σύγκριση των λειτουργικών συστημάτων. Στο δεύτερο κεφάλαιο θα μελετήσουμε τις ευφυείς μέθοδοι μηχανικής μάθησης. Αρχικά θα ορίσουμε τη μηχανική μάθηση και το datamining. Έπειτα θα μελετήσουμε τις μεθόδους ταξινόμησης K-nearest neighbor, Naïve Bayes classifier, Decision trees και Support Vector Machines – SVM. Στο τρίτο κεφάλαιο θα γνωρίσουμε το πρόγραμμα weka το οποίο θα χρησιμοποιήσουμε στην εργασία μας. Στο τέταρτο κεφάλαιο θα κάνουμε εξαγωγή χαρακτηριστικών και ένα ηλεκτρονικό ερωτηματολόγιο με το googleforms. Στο πέμπτο κεφάλαιο θα παρουσιάσουμε τα στατιστικά στοιχεία του δείγματος και τα αποτελέσματα ταξινόμηση του K-κοντινότερος γείτονας, του Μπευζιανού ταξινομητή, τα Δένδρα απόφασης και Μηχανές Διανυσμάτων υποστήριξης. Στο τελευταίο κεφάλαιο θα παρουσιάσουμε τα συμπεράσματα που προέκυψαν από την εργασία μας.

# ABSTRACT

---

This studies qualitatively investigates the operating system on portable devices. In addition classification methods have been used to predict user beliefs about each operating system and explore selection criteria.

In the first chapter of this dissertation we will describe the basic operating systems. More specifically, Android, iOS and Windows Mobile and we will make a brief historical review of operating systems and statistics for their market diffusion. The chapter will close by comparing operating systems. In the second chapter we will study intelligent methods of mechanical learning. Initially we will define mechanical learning and data mining. Then we will study the classification methods K-nearest neighbor, Naive Bayes classifier, Decision trees and Support Vector Machines - SVM. In the third chapter we will know the weka program that we will use in our work. In the fourth chapter we will export attributes and an electronic questionnaire with google forms. In the fifth chapter we will present the sample statistics and the classification results of the K-Nearest Neighbor, Beggar Classifier, Decision Trees, and Support Vector Machines. In the final chapter we will present the conclusions that have emerged from our work.

# ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

---

Κεφάλαιο 1: Περιγραφή Βασικών λειτουργικών συστημάτων.....	13
1.1 Εισαγωγή.....	13
1.2 Android.....	14
1.3 Ios.....	16
1.4 Windowsmobile.....	17
1.5 Ιστορική Αναδρομή λειτουργικών.....	18
1.6 Διάχυση στην αγορά (στατιστικά στοιχεία χρηστών).....	20
1.7 Ποιοτική σύγκριση.....	21
Κεφάλαιο 2: Ευφυείς μέθοδοι Μηχανικής Μάθησης.....	24
2.1 Γενικά για μηχανική μάθηση και datamining.....	24
2.2.Μέθοδοι ταξινόμησης.....	25
2.2.1 Κ-κοντινότερος γείτονας (KNN = K-nearestneighbor).....	25
2.2.2 Μπευζιανόςταξινομητής (NaïveBayesclassifier).....	28
2.2.3 Δένδρα απόφασης (Decisiontrees).....	31
2.3.4 Μηχανές Διανυσμάτων υποστήριξης (SupportVectorMachines - SVM)....	32
ΚΕΦΑΛΑΙΟ 3: WEKA.....	36
3.1 Εισαγωγή.....	36
3.2 Το περιβάλλον Weka.....	36
3.3 Οπτικοποίηση και Συσταδοποίηση δεδομένων.....	39
ΚΕΦΑΛΑΙΟ 4:ΥΛΟΠΟΙΗΣΗ ΠΕΙΡΑΜΑΤΩΝ.....	41
4.1 Εξαγωγή χαρακτηριστικών.....	41
4.2 Ηλεκτρονικό ερωτηματολόγιο.....	42
ΚΕΦΑΛΑΙΟ 5:ΑΠΟΤΕΛΕΣΜΑΤΑ.....	55
5.1 Στατιστικά στοιχεία δείγματος.....	55

5.2 Αποτελέσματα ταξινόμηση.....	56
ΚΕΦΑΛΑΙΟ6:ΣΥΜΠΕΡΑΣΜΑΤΑ.....	61
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	63

# Κεφάλαιο 1: Περιγραφή Βασικών λειτουργικών συστημάτων

---

## 1.1 Εισαγωγή

Με τον όρο λειτουργικό σύστημα (OperatingSystem) καλείται το λογισμικό του υπολογιστή το οποίο ως βασικό του ρόλο έχει τη διαχείριση καθώς και τον συντονισμό εργασιών. Για παράδειγμα τη διαχείριση πόρων. Αποτελείται από ένα μεσολοβητικό επίπεδο λογικής διασύνδεσης ανάμεσα σε υλικό και λογισμικό. Με την βοήθεια του οποίου οι εφαρμογές αντιλαμβάνονται έμμεσος τον υπολογιστή. Όλοι υπολογιστές καθώς και οι κινητές συσκευές χρησιμοποιούν το κατάλληλο λογισμικό. [1]

Λειτουργικό κινητής συσκευής είναι ένα λειτουργικό σύστημα το οποίο χρησιμοποιείται όχι απαραίτητα μόνο σε κινητές συσκευές αλλά και σε tablets, smartwatches, φορητούς υπολογιστές. Βέβαια αρχικά σχεδιάστηκε για επιτραπέζιους υπολογιστές και με την πάροδο των χρόνων εξελίχθηκε και εφαρμόστηκε σε κινητές συσκευές. Θα πρέπει να προσθέσουμε ότι τα λειτουργικά συστήματα κινητής τηλεφωνίας εκτός από χαρακτηριστικά λειτουργικού συστήματος προσωπικών υπολογιστών συνδυάζουν και :

- Bluetooth
- Wifi
- GPS
- Wifi Protected Access
- Ανάγνωση ομιλίας
- Κάμερα
- Βίντεο
- Συσκευή εγγραφή φωνής
- Συσκευή αναπαραγωγή φωνής
- Single frames

## 1.2 Λειτουργικό Android

Με τον όρο Android αναφερόμαστε σε ένα λειτουργικό σύστημα κινητών συσκευών το οποίο έχει αναπτυχθεί από την Google. Το συγκεκριμένο λογισμικό έχει ως βάση τα Linux.



*Εικόνα 1.1: Λογότυπο του Λειτουργικού Android*

Αποτελεί λογισμικό ανοικτού κώδικα δηλαδή υπάρχει η δυνατότητα να τροποποιηθεί ο πηγαίος κώδικας από ειδικούς του είδους. Μεγάλες εταιρίες λόγου χάρη η Motorola, Samsung το έχουν ως βάση για να φτιάξουν διάφορες εφαρμογές. Βέβαια θα πρέπει να επισημάνουμε ότι το Android δεν αναφέρεται μόνο σε κινητά τηλέφωνα αλλά και σε tablet λόγου χάρη το Nexus 7, κάμερες, κονσόλες παιχνιδιών, ρολόγια και γυαλιά. Ο ανταγωνισμός ανάμεσα στις εταιρίες οι οποίες προσπαθούν να προσελκύσουν ολοένα και περισσότερους πελάτες διαφοροποιεί το Android από συσκευή σε συσκευή και από εταιρία σε εταιρία. Υπάρχει η δυνατότητα αλλαγής εμφάνισης του γραφικού περιβάλλοντος σε περίπτωση που το επιθυμούμε και αυτό μπορεί να πραγματοποιηθεί εάν αντικαταστήσουμε τον launcher ή ολόκληρο το λειτουργικό σύστημα με μια customROM λόγου χάρη CyanogenModROM. Όσον αφορά τα τεχνικά χαρακτηριστικά αυτά διαφέρουν από συσκευή σε συσκευή αλλά πάντα έχουν το ίδιο λογισμικό δηλαδή το Android. Οι συσκευές αυτές υποστηρίζουν:

- ✓ Κάμερα
- ✓ GPS
- ✓ Bluetooth
- ✓ NFC
- ✓ Επιταχυνσιόμετρο

- ✓ Πυξίδα
- ✓ Live Wallpaper
- ✓ Widget

Κατά κύριο λόγο οι εφαρμογές Android είναι γραμμένες σε γλώσσα προγραμματισμού Java αλλά υποστηρίζουν και άλλες γλώσσες όπως C#, ObjectiveC, Ruby. [2]

Το Android έχει αφήσει το στίγμα του στην αγορά αλλά παρόλα αυτά έχει και πλεονεκτήματα και μειονεκτήματα.

### **Τα πλεονεκτήματα του Android**

- Μεγάλος αριθμός χρηστών και στις αναπτυσσόμενες χώρες
- Αρκετά απλές εφαρμογές
- Μπορούν να τρέξουν πολλές εφαρμογές
- Δυνατότητα περιήγησης
- Δυνατότητα χρήσης του Facebook
- Δυνατότητα ακούσματος ενός τραγουδιού
- Δυνατότητα δωρεάν λήψης online εφαρμογών
- Δυνατότητα άριστης υποστήριξης λογισμικού
- Διαθεσιμότητα ενημερώσεων και βελτίωσης της απόδοσης
- Δυνατότητα λειτουργίας ως δρομολογητών
- Δυνατότητα εγκατάστασης τροποποιημένου ROM
- Βέλτιστα γραφικά
- Δυνατότητα ενσωμάτωσης υπηρεσιών Google

### **Τα μειονεκτήματα του Android**

- Η δυσκολία σχεδιασμού της εφαρμογής
- Η ποιότητα εφαρμογών στο PlayStore είναι χειρότερη από το AppStore
- Η απαίτηση ενεργής σύνδεσης στο Διαδίκτυο
- Παρέχουν ιούς

- Πολλές διαφημίσεις στην οθόνη από εφαρμογές που έχουμε εγκαταστήσει
- Οι εφαρμογές έχουν επιπτώσεις στην μπαταρία
- Οι πολλές εφαρμογές κάνουν αργό το κινητό
- Υψηλό κόστος[3]

### 1.3 Λειτουργικό Ios

Το iOS αποτελεί ένα λογισμικό κινητών συσκευών το οποίο αναπτύχθηκε από την εταιρία Apple Inc.



*Εικόνα 1.2: Λογότυπο του λειτουργικού Ios*

Η πρώτη του εμφάνιση πραγματοποιήθηκε το 2007 για το iPhone αλλά υποστηρίζει και άλλες συσκευές λόγω χάρη:

- ✓ iPod touch
- ✓ iPad
- ✓ Apple TV

Εν αντιθέσει με άλλα λογισμικά κινητών συσκευών όπως για παράδειγμα Android , WindowsPhone της Microsoft η Apple δεν παρέχει τη δυνατότητα χρήσης του εν λόγω λογισμικού από εταιρίες που δεν ανήκουν στην Apple. Το 2013 το AppStore της Apple περιείχε περισσότερες από 775000 εφαρμογές iOS, 300000 από τις οποίες ήταν συμβατές με iPad. [4]

#### **Τα πλεονεκτήματα του**

- Η εκπληκτική απόδοση
- Η λιγότερη θερμότητα
- Η καλύτερη εμπειρία παιχνιδιού
- Η εξαιρετική ασφάλεια



- Η πολυεπεξεργασία
- Η εξαιρετική ενημέρωση και ψυχαγωγία
- Οι γρήγορες ρυθμίσεις
- Η σάρωση δακτυλικού αποτυπώματος
- Η εξαιρετική φωτογραφική μηχανή

#### **Τα μειονεκτήματα του**

- Η απουσία ευελιξίας
- Το κόστος
- Δεν αποτελεί ανοικτή πηγή
- Απουσία εναλλακτικών ήχων κλήσης
- Οι πολύ μεγάλες εφαρμογές
- Η υποστήριξη μιας μόνο SIM'
- Τα πολύ ακριβά έξοδα επισκευής
- Δεν υποστηρίζει NFC[5]

#### **1.4 Windows mobile**

Η εταιρία Microsoft εκτός από λογισμικό σύστημα Windows το οποίο χρησιμοποιείται σε ηλεκτρονικούς υπολογιστές έφτιαξε μια έκδοση Windows που χρησιμοποιείται από κινητές συσκευές.



*Εικόνα 1.3: Λογότυπο του λειτουργικού Windowsmobile*

Η παραπάνω έκδοση έχει πολλές ομοιότητες με τα:

- ✓ Microsoft Office
- ✓ Media Player
- ✓ Internet Explorer

Οι κινητές συσκευές οι οποίες δύναται να υποστηρίξουν φαίνονται στη σελίδα : <https://www.microsoft.com/el-gr/windows/windows-10-mobile-upgrade> και οι νεότερες εκδόσεις καλούνται WindowsPhone.[6]

### **Τα πλεονεκτήματα του**

- Η καλύτερη επίδοση
- Η ασφάλεια των εφαρμογών
- Οι καλύτερες επιλογές προσαρμογής
- Η απουσία προβλήματος κακόβουλου λογισμικού

### **Τα μειονεκτήματα του**

- Η αναμονή
- Κενά και παραλήψεις
- Η προσιτότητα
- Όχι Google
- Η απουσία παιχνιδιών
- Bluetooth HID
- Απουσία έμπνευσης[7][8]

## **1.5 Ιστορική Αναδρομή λειτουργικών**

Στη συνέχεια θα παρουσιάσουμε την ιστορική εξέλιξη των λογισμικών κινητών συσκευών από το 1973 έως και το 2017.



Εικόνα 1.4: Λειτουργικά κινητών συσκευών

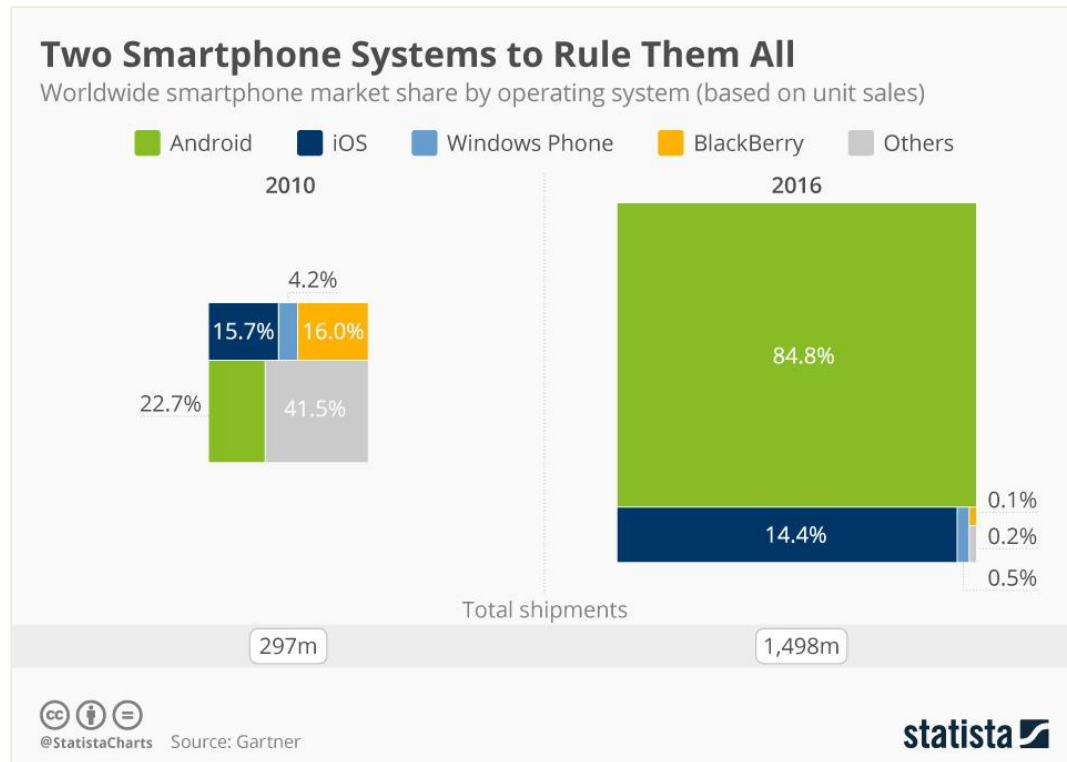
- 1973-1993: Το πρώτο Smartphone, χρησιμοποιούνται ενσωματωμένα συστήματα για τον έλεγχο της λειτουργίας.
- 1994: Παρουσιάστηκε το πρώτο Smartphone , το IBMSimon το οποίο διέθετε χαρακτηριστικά οθόνης αφής , email και PDA
- 1996: ΤοPalmPilot 1000 εισάγει προσωπικό ψηφιακό βοηθό με το κινητό λειτουργικό σύστημα PalmOS
- 1998: Η SymbianLtd έχει αναπτύξει λειτουργικό σύστημα Symbian το οποίο χρησιμοποιήθηκε από πολλές μάρκες κινητών, λόγω χάρη η NOKIA
- 1999: Η πλατφόρμα NOKIAS40 παρουσιάζεται επίσημα
- 2000: Το Symbian αποτελεί το πρώτο σύγχρονο λειτουργικό σύστημα σε ένα Smartphone με την εκκίνηση του EricssonR380
- 2001: Το Kyocera 6035 αποτελεί το πρώτο Smartphone με PalmOS
- 2002: Παρουσιάζονται τα πρώτα smartphones των WindowsCE της Microsoft και το BlackBerry
- 2005: Η NOKIA παρουσιάζει το MaemoOS στο πρώτο tablet του N770 στο Διαδίκτυο
- 2007: ΤοiPhoneApple εισάγεται ως iPod κινητό τηλέφωνο και InternetCommunicator
- 2008: Το OHA κυκλοφορεί το Android 1.0 με το HTCdream ως το πρώτο τηλέφωνο Android

- 2009: Το Palm εισάγει το webOS με το PalmPre. Κοντά στο 2012 οι συσκευές WebOS διακόπηκαν
- Την ίδια χρονιά η Samsung ανακοίνωσε το λειτουργικό σύστημα Bada με την εισαγωγή του SamsungS8500
- 2010: Κυκλοφόρησαν τα WindowsPhoneOS τα οποία δεν ήταν συμβατά με το προηγούμενο λειτουργικό
- 2011: Η MeeGo και Mobinπαρουσιάζεται με το NokiaN9, ηSamsung , η Intel, LinuxFoundation
- 2012: Το Mozillaανακοίνωσε ότι το BoottoGecko καλείται πλέον FirefoxOS
- Την ίδια χρονιά η Apple κυκλοφόρησε το iOS 6
- 2013: ΗBlackberry κυκλοφόρησε το λειτουργικό σύστημα για smartphonesBlackBerry 10
- Την ίδια χρονιά η Apple κυκλοφόρησε το iOS 7 και η Google το KitKat 4.4
- 2014: Η Microsoft κυκλοφορεί το Windows Phone 8.1και η BlackBerry κυκλοφορεί το BlackBerry 10 .3 με ενσωμάτωση στο Amazon Appstore
- 2015:Η Apple κυκλοφορεί το iOS 9, η Google κυκλοφορεί το Android 6.0 και η Microsoft κυκλοφορεί τα Windows 10 Mobile .
- 2016: ΗMicrosoft κυκλοφόρησε την ενημερωμένη έκδοση για την επέκταση των WindowsMobile 10 και η Apple ανακοίνωσε το iOS 10
- 2017: Η Samsung ξεκίνησε επίσημα το Android, ηMicrosoft κυκλοφόρησε την ενημερωμένη έκδοση για τα Windows 10 Mobile Creators.
- Η Google κυκλοφόρησε το Android 8.0 και η Apple εισάγει τα iPhone 8 , το iPhoneX και το iOS 11 . Τέλος η Microsoft ανήγγειλε την ανάπτυξη τωνWindows 10 Mobile. [9]

## **1.6Διάχυση στην αγορά**

Το 2010 εμφανίστηκαν τα πρώτα έξυπνα τηλέφωνα στην αγορά. Από τη μια πλευρά η BlackBerry βρισκόταν σε μια πολύ ισχυρή θέση, από την άλλη πλευρά το Symbian τροφοδοτούσε εκατομμύρια συσκευές και από την άλλη πλευρά η Samsung ερχόταν να κάνει τη δική της πρόταση στην αγορά. Μετά από 7 χρόνια οι ρόλοι έχουν αντιστραφεί και πλέον την κυριαρχία έχουν δύο μόνο πλατφόρμες. Η iOS της

Apple και Android της Google. Σύμφωνα με επίσημα στατιστικά στοιχεία του τρέχοντος έτους η Gartner στο 99% κυριαρχούν οι παραπάνω πλατφόρμες. Από την άλλη πλευρά η BlackBerry και Microsoft τείνουν να χαθούν από την αγορά. Στην παρακάτω εικόνα παρουσιάζουμε αναλυτικότερα τα στατιστικά δεδομένα:



Εικόνα1.5: Two smartpnone Systems to Rule Them All [10]

## 1.7 Ποιοτική σύγκριση

**Android:** το συγκεκριμένο λογισμικό κατέχει το 75% στο μερίδιο της αγοράς και έχει τη μεγαλύτερη ποικιλία εφαρμογών. Οι εταιρίες που υποστηρίζουν Android είναι :

- ✓ H LG
- ✓ H Sony
- ✓ HTC
- ✓ Motorola
- ✓ Lenovo



Εικόνα 1.6: Σύγκριση λογισμικών κινητών συσκευών

Όσον αφορά το εύρος των τιμών θα πρέπει να επισημάνουμε ότι καλύπτει μεγάλη γκάμα τιμών από 80 ευρώ έως και 500 ευρώ. Οπότε αναφέρεται σε μεγάλο αγοραστικό κοινό. Το παραπάνω λογισμικό αποτελεί εξαιρετική επιλογή για κάποιον χρήστη που επιθυμεί το τηλέφωνο του από πλευρά του υπολογιστή και επίσης επιθυμεί αλλαγές εμφάνισης .

**IOS:** το συγκεκριμένο λογισμικό είτε έχει φανατικούς οπαδούς είτε όχι. Κατακτά το 45% στην Αμερική και το 14% είναι το λειτουργικό που έκανε γνωστά τα έξυπνα τηλέφωνα στο ευρύ κοινό. Το εν λόγω λογισμικό είναι κλειστό και ουσιαστικά δεν επιτρέπει πρόσβαση στα βασικά στοιχεία του λειτουργικού. Η δομή της διεπαφής έχει πολλές ομοιότητες με το Android και επιπλέον βασίζεται σε εικονίδια ομαδοποιημένα σε σελίδες. Το κόστος του ξεκινά από 650 ευρώ και φτάνει τα 1100 ευρώ. Το εν λόγω λογισμικό παρά το υψηλό κόστος του αποτελεί μια πολύ καλή επιλογή για οποιον επιθυμεί μια εξαιρετική ποιότητα κατασκευής.

**WindowsPhones:** το συγκεκριμένο λογισμικό αποτελεί κατάλληλο εργαλείο για χρήστες που πρώτη φορά έρχονται σε επαφή με Smartphones και έχουν μια ευκολία με υπολογιστές που χρησιμοποιούν windows. Αποτελεί ένα εργαλείο που είναι ιδιαίτερα εύκολο στην εκμάθηση. Έχουν εξαιρετικές φωτογραφικές δυνατότητες και αποτελούν συσκευές με πολύ υψηλή ποιότητα. Το 95% της αγοράς το έχουν οι συσκευές NOKIA/Microsoft και το κόστος του δεν είναι ιδιαίτερα υψηλό.

Σύμφωνα με τα παραπάνω ανάλογα τις δυνατότητες, την ευκολία χρήσης, το κόστος και το τι αναζητά ο κάθε χρήστης ο ίδιος θα καταλήξει στο τι είναι καλύτερο τελικά για τον ίδιο και τι τον εξυπηρετεί στη δουλειά του. Ο ανταγωνισμός φέρει πάντα τα επιθυμητά θετικά αποτελέσματα διότι μέσα από αυτόν όλοι γίνονται καλύτεροι προσπαθώντας πάντα να προσελκύσουν νέους πελάτες και να αυξήσουν τις πωλήσεις τους το βέλτιστο δυνατό.

Στο περιοδικό PCworld παρουσιάζεται η παρακάτω εικόνα 1.6 η οποία συγκρίνει τα τρία λογισμικά (βλ. WindowsPhone, iOS ,Android) και αναφέρει ποια είναι τα υπέρ και τα κατά κάθε λογισμικού:

	 Windows Phone 7	 iOS (iPhone)	 Android
Developer	Microsoft	Apple	Google
Copy/Paste	x	✓	✓
Multitasking	x	✓	✓
Flash Support	x	x	✓
Silverlight Support	x	x	x
HTML5 Support	x	✓	✓
Unified Inbox	x	✓	✓
Exchange Support	✓	✓	✓
Threaded Email	x	✓	✓
Visual Voicemail	x	✓	✓
Video Calling	x	✓	✓ Third Party App
Universal Search	x	✓	✓
Internet Tethering	x	✓	✓
Removable Storage	x	x	✓
Facebook Integration	✓	x (Third Party App)	✓ (Third Party Integration)
Twitter Integration	x	x (Third Party App)	✓ (Third Party Integration)
Folders	Hubs	✓	✓
Apps Organization	Alphabetical	Customizable	Customizable
App Store	1,000+ Apps	300,000+ Apps	90,000+ Apps
Microsoft Office Support	Built-In	Third Party App	Third Party App
Widgets	Tiles on Home Screen	x	✓
Media Sync	Zune Software Mac & PC	iTunes Mac & PC	Direct File Transfer + Third Party Software
X-Box Live Integration	Built-In	Via Third Party App	Via Third Party App

Εικόνα 1.7: Σύγκριση Λογισμικών κινητών συσκευών

# Κεφάλαιο 2: Ευφυείς μέθοδοι Μηχανικής Μάθησης

---

## 2.1 Γενικά για μηχανική μάθηση και datamining

Με τον όρο μηχανική μάθηση αναφερόμαστε σε έναν τομέα της ΤΝ που ασχολείται με προγράμματα τα οποία μαθαίνουν μέσα από τις εμπειρίες τους. Η βασική ιδέα της μηχανικής μάθησης αναφέρει ότι υπάρχει ο πράκτορας ο οποίος ότι προσλαμβάνει από το περιβάλλον δύναται να το χρησιμοποιήσει με απώτερο σκοπό να αποφασίσει τις ενέργειες που θα κάνει ώστε να λύσει ένα συγκεκριμένο πρόβλημα και να βελτιώσει τις μελλοντικές ενέργειες.

Οι κατηγορίες μηχανικής μάθησης είναι :

- Μάθηση με επίβλεψη
  - Μάθηση από παρατήρηση εισόδων και εξόδων παραδειγμάτων
- Μάθηση χωρίς επίβλεψη
  - Μάθηση που δεν γνωρίζουμε τις εξόδους των παραδειγμάτων
- Ενισχυτική μάθηση
  - Μάθηση μέσω ενίσχυση – επιβράβευση[11]

Με τον όρο εξόρυξη δεδομένων αναφερόμαστε στην εξεύρεση μιας πληροφορίας ή ακόμη και προτύπων μεγάλων βάσεων δεδομένων με χρήση των κατάλληλων αλγορίθμων ομαδοποίησης ή κατηγοριοποίησης καθώς επίσης και των αρχών :

- Στατιστικής
- Τεχνητής νοημοσύνης
- Μηχανικής μάθησης

των συστημάτων βάσεων δεδομένων. Βασικό μέλημα της εξόρυξης δεδομένων η πληροφορία η οποία θα εξαχθεί καθώς και τα πρότυπα να αποτελούνται από μια κατανοητή δομή προς τον άνθρωπο με σκοπό να βοηθήσουν στο να πάρει σωστές αποφάσεις. [12]



## 2.2 Μέθοδοι ταξινόμησης

### 2.2.1 K-κοντινότερος γείτονας

Ο αλγόριθμος K-κοντινότερος γείτονας ή αλλιώς όπως είναι γνωστός KNN αποτελεί έναν πολύ απλό αλγόριθμο έχει την ιδιότητα να αποθηκεύει τις διαθέσιμες περιπτώσεις και να κάνει ταξινόμηση των καινούριων περιπτώσεων με βάση κάποιο μέτρο ομοιότητας λόγω χάρη την απόσταση. Ο εν λόγω αλγόριθμος πρωτοεμφανίστηκε κοντά στο 1970 που χρησιμοποιήθηκε για στατιστική εκτίμηση στην αναγνώριση προτύπων.

#### Πώς λειτουργεί ο KNN;

Μια υπόθεση κατατάσσει τους πλησιέστερους γείτονες K οι οποίοι μετριοούνται με τη χρήση μιας συνάρτησης απόστασης. Χαρακτηριστικά παραδείγματα τέτοιων αποστάσεων αποτελούν οι:

- Euclidean
- Manhattan
- Minkowski

Οι αποστάσεις ορίζονται ως εξής:

$$\text{Euclidean} \quad \sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad \text{Εξ. (1)}$$

$$\text{Manhattan} \quad \sum_{i=1}^k |x_i - y_i| \quad \text{Εξ. (2)}$$

$$\text{Minkowski} \quad \left( \sum_{i=1}^k (|x_i - y_i|)^q \right)^{1/q} \quad \text{Εξ. (3)}$$

Τα παραπάνω μέτρα απόστασης αφορούν μόνο συνεχείς μεταβλητές ενώ η απόσταση Hamming χρησιμοποιείται στις κατηγορικές μεταβλητές και σε περίπτωση μεταβλητών 0,1 όταν υπάρχει ένα μίγμα κατηγορικών μεταβλητών στο σύνολο δεδομένων.

### Hamming Distance

$$D_H = \sum_{i=1}^k |x_i - y_i|$$

$$x = y \Rightarrow D = 0$$

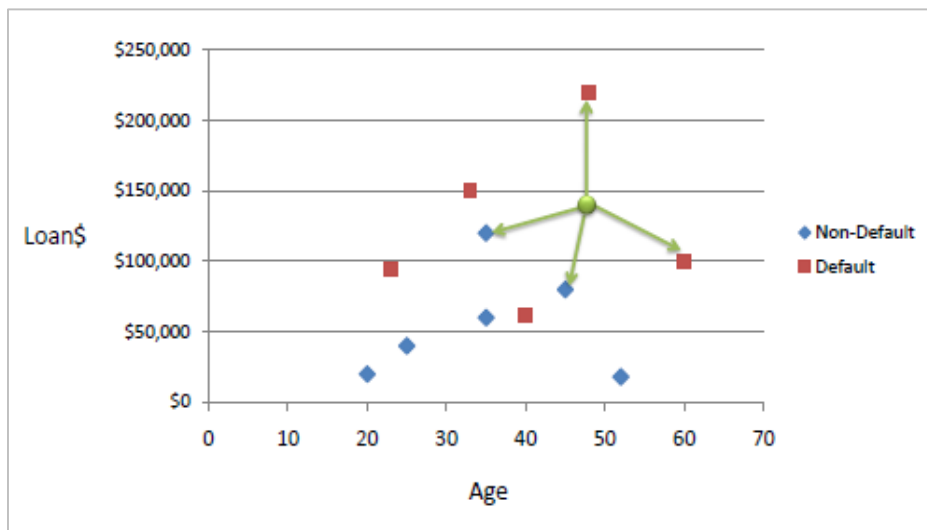
$$x \neq y \Rightarrow D = 1$$

X	Y	Distance
Male	Male	0
Male	Female	1

Εικόνα 2.1: HammingDistance

Η επιλογή της βέλτιστης τιμής K είναι η ακριβέστερη και συνήθως γίνεται με την πρώτη επιθεώρηση δεδομένων. Η συγκεκριμένη επιλογή ελαχιστοποιεί τον θόρυβο. Θα πρέπει να αναφέρουμε ότι η βέλτιστη τιμή K για τα περισσότερα σύνολα δεδομένων βρίσκεται ανάμεσα σε 3-10.

Στη συνέχεια θα περιγράψουμε ένα παράδειγμα του αλγορίθμου για δεδομένα που σχετίζονται με πιστωτική προεπιλογή. Πιο συγκεκριμένα η ηλικία, το δάνειο αποτελούν δύο μεταβλητές πρόβλεψης και η προεπιλογή είναι ο στόχος.



Εικόνα 2.2: ExampleKNN

Με τη χρήση Euclidean απόστασης μπορούμε να ταξινομήσουμε μια άγνωστη περίπτωση και συγκεκριμένα την περίπτωση όπου η ηλικία έχει την τιμή 48 και το

δάνειο την τιμή \$ 142,000. Σε περίπτωση που K=1 τότε ο πλησιέστερος γείτονας αποτελεί η τελευταία περίπτωση.

$$D = \text{Sqrt}[(48-33)^2 + (142000-150000)^2] = 8000.01 \gg \text{Προεπιλογή} = Y \quad \text{Εξ. (4)}$$

Age	Loan	Default	Distance
25	\$40,000	N	102000
35	\$60,000	N	82000
45	\$80,000	N	62000
20	\$20,000	N	122000
35	\$120,000	N	22000
52	\$18,000	N	124000
23	\$95,000	Y	47000
40	\$62,000	Y	80000
60	\$100,000	Y	42000
48	\$220,000	Y	78000
33	\$150,000	Y	8000
48	\$142,000	?	

Euclidean Distance

$$D = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

Εικόνα 2.3: Example KNN

Με K=3 μπορούμε να παρατηρήσουμε ότι υπάρχουν δύο Default=Y και ένα Default=N όπου και πάλι ο πλησιέστερος είναι η προεπιλογή =Y.

Ο υπολογισμός των μέτρων της απόστασης μειονεκτεί στο ότι οι μεταβλητές έχουν διαφορετικές κλίμακες μέτρησης ή υπάρχει ένα μίγμα αριθμητικών και κατηγορικών μεταβλητών. Στην παρακάτω εικόνα παρουσιάζουμε ένα τυποποιημένο σετ εκπαίδευσης:

Age	Loan	Default	Distance
0.125	0.11	N	0.7652
0.375	0.21	N	0.5200
0.625	0.31	N	0.3160
0	0.01	N	0.9245
0.375	0.50	N	0.3428
0.8	0.00	N	0.6220
0.075	0.38	Y	0.6669
0.5	0.22	Y	0.4437
1	0.41	Y	0.3650
0.7	1.00	Y	0.3861
0.325	0.65	Y	0.3771
0.7	0.61	?	

Standardized Variable

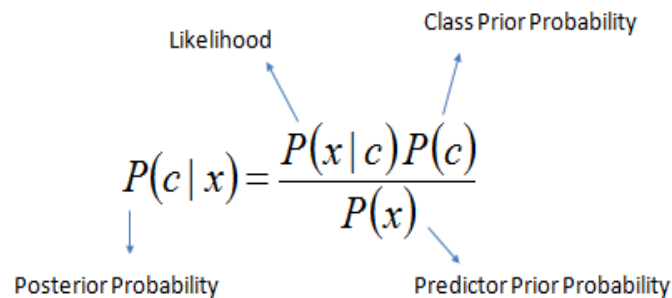
$$X_s = \frac{X - Min}{Max - Min}$$

Εικόνα 2.4: ExampleKNN

Όπως μπορούμε να παρατηρήσουμε με τη χρήση της τυποποιημένης απόστασης στο ίδιο σετ εκπαίδευσης ο κοντινότερος γείτονας είναι διαφορετικός.[13]

### 2.2.2 Μπευζιανός ταξινομητής

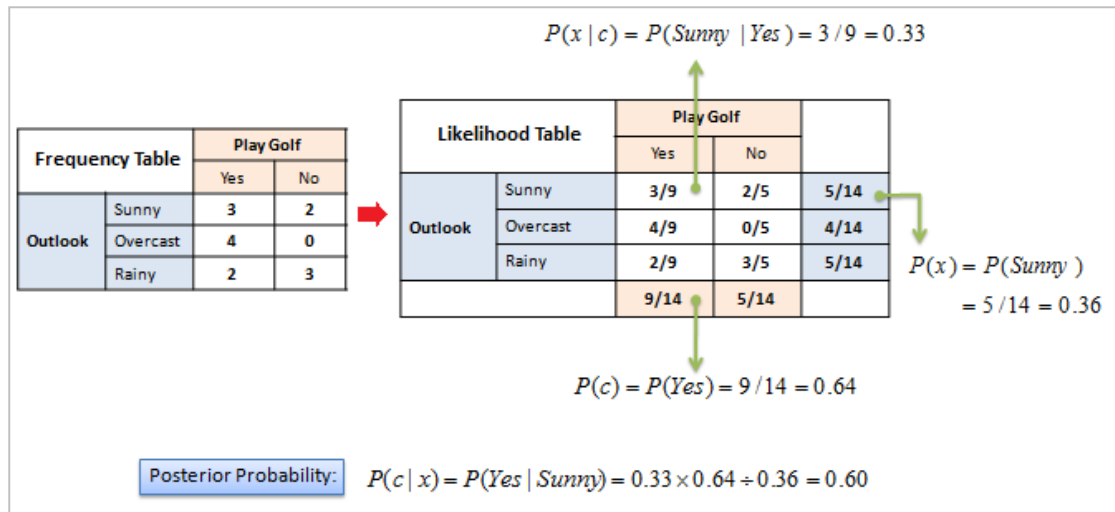
Ο Naïve Bayesian αποτελεί έναν ταξινομητή ο οποίος βασίζεται στο θεώρημα του Bayes. Ο εν λόγω αλγόριθμος δύναται να υλοποιηθεί πολύ εύκολα χωρίς να χρειάζονται ιδιαίτερες παραμετροποιήσεις. Αποτελεί ένα εξαιρετικό εργαλείο ταξινόμησης για μεγάλα σύνολα δεδομένων. Το θεώρημα του Bayes υπολογίζεται η πιθανότητα  $P(c|x)$  από  $P(c)$ ,  $P(x)$  και  $P(x|c)$ . Υπάρχει η τιμή ενός προγνώστη  $x$  σε μια δεδομένη κλάση  $c$  που είναι ανεξάρτητη από τις τιμές άλλων προγνωστικών.

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


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

Εικόνα 2.5: Πιθανότητα Bayes

Η πιθανότητα  $P(c|x)$  δύναται να υπολογιστεί εύκολα με τη δημιουργία ενός πίνακα συχνοτήτων για κάθε χαρακτηριστικό έναντι του στόχου. Έπειτα μετασχηματίζονται οι πίνακες συχνοτήτων σε πίνακες πιθανοτήτων και στο τέλος χρησιμοποιούμε την Naïve Bayesian εξίσωση για να υπολογίσουμε την πιθανότητα  $P(c|x)$  κάθε κατηγορίας. Η τάξη με την υψηλότερη Posterior πιθανότητα αποτελεί το αποτέλεσμα της πρόβλεψης.



Εικόνα 2.6: ExampleNaïveBayes

### Το πρόβλημα μηδενικής συχνότητας

Αν προσθέσουμε 1 στην αρίθμηση για κάθε συνδυασμό αξιών –κλάσης χαρακτηριστικού σε περίπτωση που δεν εμφανίζεται η τιμή χαρακτηριστικού (Outlook=Overcast) με κάθε τιμή κλάσης (PlayGolf=όχι).

### Αριθμητικοί προγνώστες

Οι αριθμητικές μεταβλητές θα πρέπει να μετατραπούν σε κατηγορίες πριν δημιουργηθούν οι πίνακες συχνοτήτων. Διαφορετικά μπορούμε να χρησιμοποιήσουμε τη διανομή της αριθμητικής μεταβλητής για να έχουμε καλή εικόνα για τη συχνότητα.

Η συνάρτηση πυκνότητας πιθανότητας για την κανονική κατανομή ορίζεται από δύο παραμέτρους :

- Την μέση απόκλιση
- Την τυπική απόκλιση

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{Mean} \quad \text{Εξ (5)}$$

$$\sigma = \left[ \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2 \right]^{0.5} \quad \text{Standard deviation} \quad \text{Εξ (6)}$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Normal distribution

Εξ (7)

### Παράδειγμα

		Humidity								Mean	StDev	
Play Golf	yes	86	96	80	65	70	80	70	90	75	79.1	10.2
	no	85	90	70	95	91					86.2	9.7

Εικόνα 2.7: Example Naive Bayes

$$P(\text{humidity} = 74 \mid \text{play} = \text{yes}) = \frac{1}{\sqrt{2\pi}(10.2)} e^{-\frac{(74-79.1)^2}{2(10.2)^2}} = 0.0344 \quad \text{Εξ (8)}$$

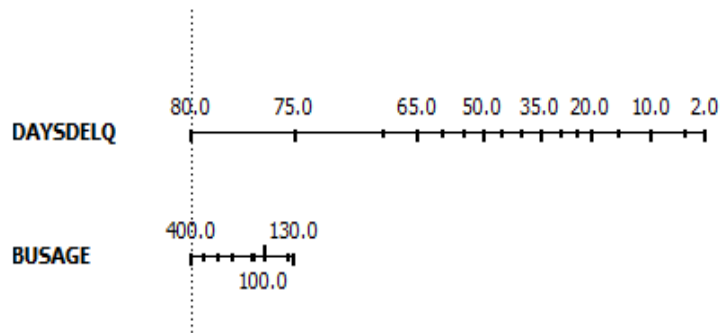
$$P(\text{humidity} = 74 \mid \text{play} = \text{no}) = \frac{1}{\sqrt{2\pi}(9.7)} e^{-\frac{(74-86.2)^2}{2(9.7)^2}} = 0.0187 \quad \text{Εξ (9)}$$

### Προβλέψεις συμβολής

Το κέρδος πληροφοριών του Κοποnenko ως ένα άθροισμα πληροφοριών το οποίο συνεισφέρει σε κάθε χαρακτηριστικό, δύναται να δώσει μια εξήγηση για το πώς οι τιμές των προγνωστικών δύναται να επηρεάζουν την πιθανότητα κλάσης.

$$\text{Log}_2 P(c/x) - \text{Log}_2 P(c) \quad \text{Εξ (10)}$$

Η συμβολή των προγνωστικών μπορεί να απεικονιστεί με την καταγραφή των γραφημάτων. Οι αναλογίες πιθανών καταστάσεων του γραφήματος για κάθε τιμή κάθε προγνωστικό. Τα μήκη των γραμμών αντιστοιχούν στα ποσοστά των λόγων πιθανότητας κάτι που υποδηλώνει τη σημασία του προγνωστικού. [14]

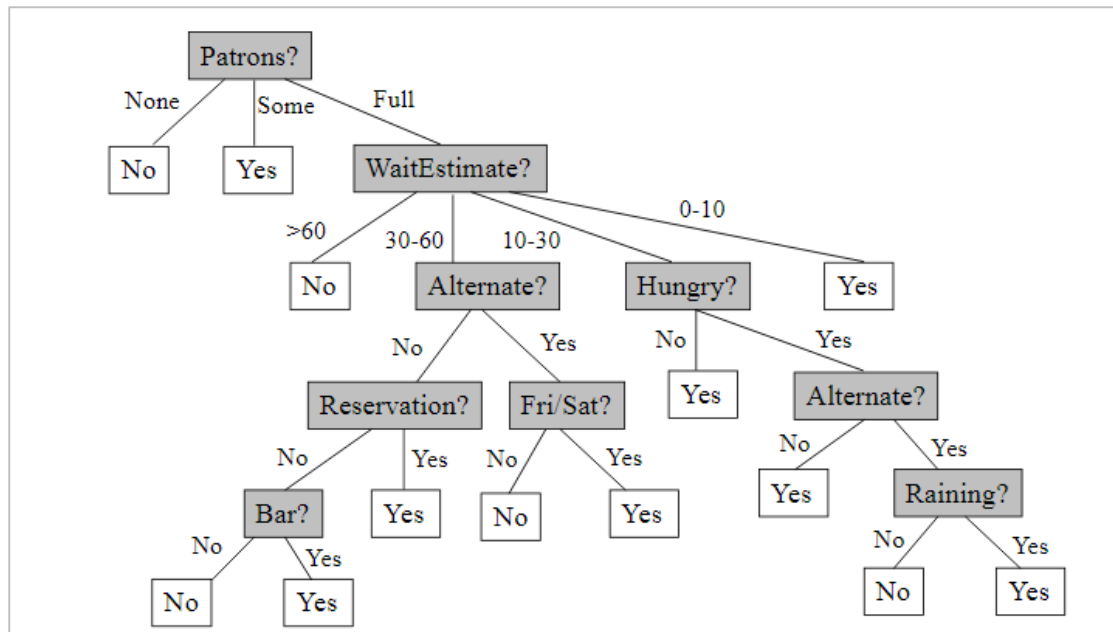


Εικόνα 2.8: Γράφημα

### 2.2.3 Δένδρα απόφασης

Τα δέντρα απόφασης αποτελούν έναν πολύ απλό αλγόριθμο μάθησης. Ως είσοδο έχουν ένα αντικείμενο ή μια κατάσταση η οποία περιγράφεται από ένα σύνολο ιδιοτήτων και επιστρέφει ναι ή όχι. Ουσιαστικά τα δέντρα απόφασης αναπαριστούν Boolean συναρτήσεις όπου κάθε εσωτερικός κόμβος αντιστοιχεί σε τεστ της τιμής μιας από τις ιδιότητες. Στη συνέχεια παρουσιάζουμε ένα παράδειγμα δέντρου αποφάσεων όπου έχουμε τα παρακάτω γνωρίσματα:

- Alternate
- Bar
- Fri/Sat
- Hungry
- Patrons
- Price
- Raining
- Reservation
- Type
- WaitEstimate .[15]



Εικόνα 2.11: Example Δέντρου Αποφάσεων

### 2.3.4 Μηχανές Διανυσμάτων υποστήριξης

Οι μηχανές διανυσμάτων υποστήριξης γνωστές με την ορολογία SVM αποτελούν έναν από τους πολύ γνωστούς αλγορίθμους ταξινόμησης. Αποτελούν μια τεχνική μάθησης υπό εποπτεία. Σε περίπτωση που έχουμε ένα σύνολο δεδομένων με χαρακτηριστικά και ετικέτες κλάσης χρησιμοποιούμε SVM. Όταν έχουμε ένα σύνολο δεδομένων το οποίο όπως προαναφέραμε αποτελείται από χαρακτηριστικά γνωρίσματα καθώς και ετικέτες ο ταξινομητής SVM δημιουργεί ένα μοντέλο για πρόβλεψη των τάξεων για νέα παραδείγματα. Σε αυτή την περίπτωση αναθέτει νέα παραδείγματα / σημεία δεδομένων σε μια από τις κλάσεις. Σε περίπτωση που αποτελεί μόνο από δύο κλάσης αναφερόμαστε σε δυαδική ταξινομητή SVM. Τα είδη των ταξινομητών SVM είναι:

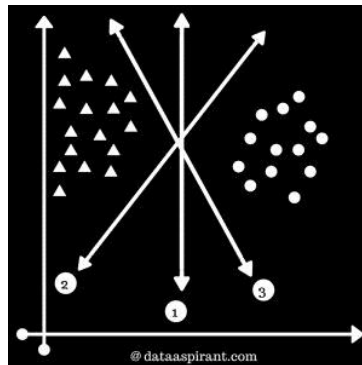
- ✓ Γραμμικός ταξινομητής SVM
- ✓ Μη γραμμικός ταξινομητής SVM

Στη συνέχεια θα παρουσιάσουμε κάποιες περιπτώσεις για να γνωρίσουμε την επιλογή του καλύτερου υπερπληρωμένου για την ταξινόμησης. Θα παρουσιάσουμε δεδομένα από 2 κατηγορίες. Οι κλάσεις αποτελούνται από τρίγωνο και κύκλο.



### 1<sup>η</sup> περίπτωση

Στην εικόνα έχουμε δεδομένα από 2 κλάσεις . Θέλουμε να βρούμε τον βέλτιστο διαχωρισμό για τις δύο κατηγορίες. Στην εικόνα στα δεξιά μπορείτε να δείτε ποια περίπτωση ταιριάζει καλύτερα. Στον ταξινομητή SVM μπορούμε να μεγιστοποιήσουμε την απόσταση ανάμεσα σε υπερπληθυσμό και πλησιέστερου σημείου δεδομένων.

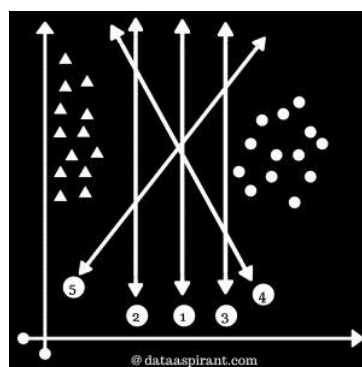


Εικόνα 2.10: Πρώτη περίπτωση

Δηλαδή το περιθώριο. Έχοντας ως δεδομένο ότι το πρώτο όριο απόφασης μεγιστοποιεί την απόσταση ανάμεσα στις κλάσεις αριστερά και δεξιά. Κατά συνέπεια το μέγιστο περιθώριο είναι το 1<sup>ο</sup>.

### 2<sup>η</sup> περίπτωση

Στην παραπάνω εικόνα παρουσιάζουμε δεδομένα από 2 διαφορετικές κατηγορίες. Καθώς τα δεδομένα κάθε τάξης διανέμονται είτε αριστερά είτε δεξιά στόχος είναι η επιλογή υπερπλήσιασπου μπορεί να χωρίσει τις κατηγορίες με το μέγιστο περιθώριο.



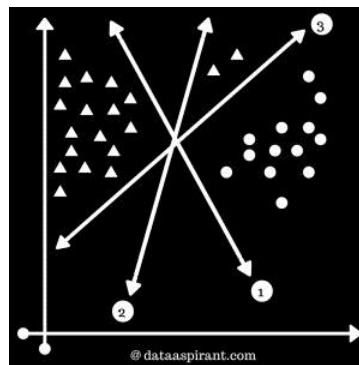
Εικόνα 2.11: Δεύτερη περίπτωση

Βέβαια μόνο το πρώτο όριο απόφασης εμφανίζει το μέγιστο περιθώριο ανάμεσα σε τρίγωνο και κύκλο.

Στην παραπάνω εικόνα μπορούμε να διακρίνουμε δύο κλάσεις και ψαχνουμε πάλι να βρούμε τον καλύτερο διαχωρισμό. Τα δεδομένα όπως μπορούμε να παρατηρήσουμε δεν κατανέμονται ομοιόμορφα αριστερά και δεξιά. Επίσης κάποια από τα τρίγωνα βρίσκονται δεξιά.

### **3<sup>η</sup> περίπτωση**

Θα μπορούσαμε να βρούμε το βέλτιστο διαχωρισμό αγνοώντας κάποια σημεία αλλά αυτό δεν είναι σωστό. Το SVM έχει ως στόχο να βρει το βέλτιστο διαχωρισμό αλλά πρώτα πρέπει να διορθώσει τα λάθη.



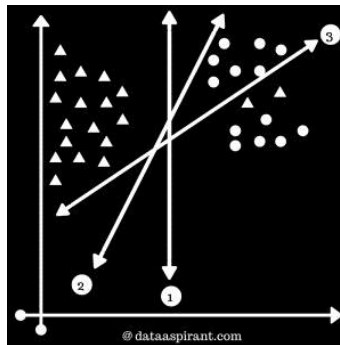
Εικόνα 2.12: Τρίτη περίπτωση

Το πρώτο όριο χωρίζει μερικά τρίγωνα από κύκλους αλλά όχι όλα. Από την άλλη πλευρά το δεύτερο όριο απόφασης χωρίζει τα δεδομένα παρόμοια με το πάνω όριο είναι πολύ μεγαλύτερο από το προηγούμενο. Τέλος το τρίτο όριο απόφασης διαχωρίζει όλα τα τρίγωνα από όλους τους κύκλους. Άρα η SVM θα επιλέξει το τρίτο όριο.

### **4<sup>η</sup> περίπτωση**

Στην παρακάτω εικόνα μπορούμε να γνωρίσουμε τα υψηλά επίπεδα SVM. Έπειτα θα χωρίσουμε με το βέλτιστο τρόπο τις δύο κατηγορίες. Και εδώ παρατηρούμε ότι τα δεδομένα δεν κατανέμονται ομοιόμορφα δεξιά και αριστερά. Κάποια από τα τρίγωνα βρίσκονται δεξιά. Τα δεδομένα αυτά αν υπήρχαν στον πραγματικό κόσμο θα αντιπροσώπευαν ακραίες περιπτώσεις. Η SVM αγνοεί ακραίες περιπτώσεις στην

εικόνα παρατηρούμε ότι 2 τρίγωνα βρίσκονται ανάμεσα στην ομάδα των κύκλων. Αυτά τα τρίγωνα είναι υπερβολικά υψηλά.

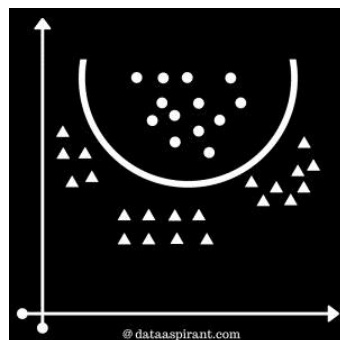


Εικόνα 2.13: Τέταρτη περίπτωση

Βέβαια η SVM αγνοεί τα τρίγωνα σε αυτή την περίπτωση. Το πρώτο και δεύτερο όριο απόφασης αποτελούν τις κλάσεις διαχωρισμού αλλά το πρώτο όριο απόφασης δείχνει το μέγιστο περιθώριο ανάμεσα στα όρια .

### 5<sup>η</sup> περίπτωση

Στην εικόνα τα δεδομένα δεν μπορούν να είναι χωρισμένα από κάθε ευθεία γραμμή δηλαδή δεν διαχωρίζονται γραμμικά αλλά μπορούμε να χρησιμοποιήσουμε διαφορετικούς τύπους. Το όριο απόφασης το οποίο χωρίζει και τις δύο κλάσεις και μοιάζει με παραβολή [16].



Εικόνα 2.14: Πέμπτη περίπτωση

# ΚΕΦΑΛΑΙΟ 3

## WEKA

---

### 3.1 Εισαγωγή

Το λογισμικό Weka με την ακριβή ορολογία Waikato Environment for Knowledge Analysis αποτελεί ένα εύχρηστο λογισμικό για την μηχανική μάθηση και την εξόρυξη δεδομένων. Η ονομασία του προέρχεται από το Πανεπιστήμιο του Waikato που βρίσκεται στην Νέα Ζηλανδία. Πιο συγκεκριμένα υπάρχει ένα πουλί στην Νέα Ζηλανδία και καλείται Weka.



Εικόνα 3.1: Το Weka

Ανήκει στην κατηγορία ελεύθερου λογισμικού δηλαδή διατίθεται δωρεάν, δημοσίως σύμφωνα με άδεια του GNU (General Public License). Με χρήση της συγκεκριμένης άδειας υπάρχει δυνατότητα ελεύθερης τροποποίησης του λογισμικού. Δεν είναι τυχαία ένα από τα πιο δημοφιλή λογισμικά για την εξόρυξη δεδομένων. Μερικές από τις σημαντικές δυνατότητες :

1. Ποικιλία μεθόδων κατηγοριοποίησης, παλινδρόμηση, ανάλυση συστάδων και κανόνες συσχέτισης
2. Δυνατότητα προεπεξεργασίας των δεδομένων
3. Δυνατότητα χρήσης εργαλείων οπτικοποίησης
4. Είναι λογισμικό ανοικτού κώδικα

5. Δυνατότητα εγκατάστασης σε διαφορετικές πλατφόρμες υλικού και λογισμικού

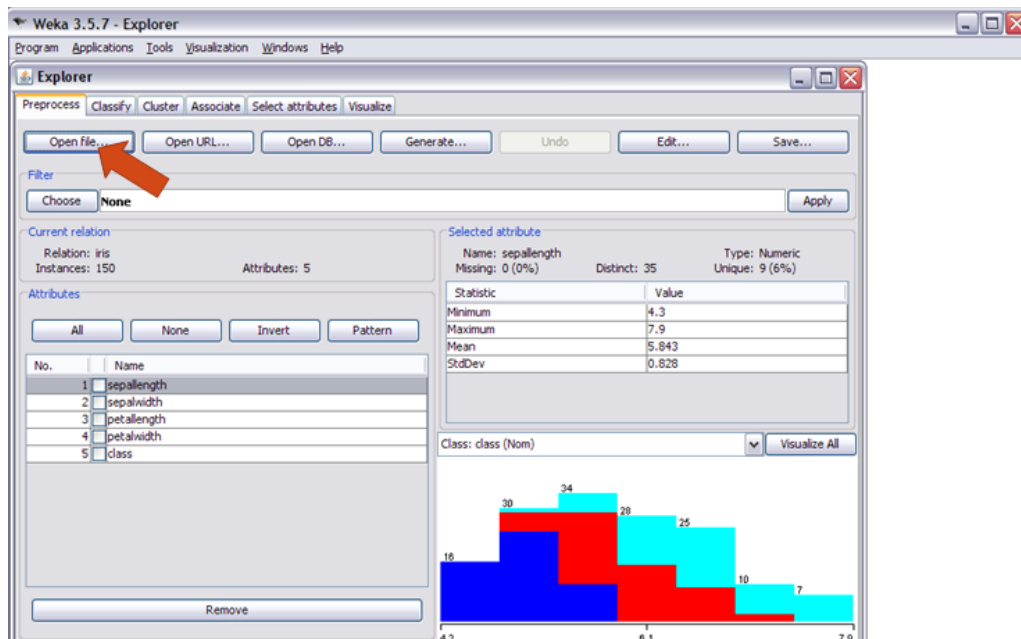
Το Weka διατίθεται σε δυο διαφορετικές εκδόσεις:

- Η σταθερή (stable)
- Η έκδοση που απευθύνεται σε προγραμματιστές

Το Weka διατίθεται για εγκατάσταση από την ιστοσελίδα <http://www.cs.waikato.ac.nz/ml/weka/>. Επίσης θα προσθέσουμε ότι για το περιβάλλον των Windows σε περίπτωση που κάποια έκδοση της Java δεν είναι ήδη εγκατεστημένη το εκτελέσιμο που θα αποθηκευτεί είναι η έκδοση που περιλαμβάνει την JavaVM 1.6.

### 3.2. Το περιβάλλον Weka

Ανοίγοντας το πρόγραμμα του Weka πηγαίνοντας στο menu Application-Explorer-OpenFile μπορούν να επιλεγούν ένα σύνολο από δεδομένα όπου θα εφαρμοστούν τεχνικές που αφορούν Preprocess, Classify, Cluster, Associate, Selectattribute και Visualize.



Εικόνα 3.2: Το περιβάλλον Weka

Αρκεί να επιλέξουμε ένα σύνολο δεδομένων ώστε να μπορούν να εφαρμοστούν γραφικά και στατιστικά τα δεδομένα. Σε περίπτωση που στο σύνολο δεδομένων υπάρχει η κλάση στην οποία ταξινομούνται τα δεδομένα που ανήκουν στην ίδια κλάση έχουν το ίδιο χρώμα. Τα αρχεία του weka έχουν κατάληξη .arff σαν επέκταση και παραδείγματα τέτοιων αρχείων βρίσκονται στον φάκελο C:\ Programfiles \ Weka-3-5\data. Επίσης στο Weka υπάρχει η δυνατότητα να χρησιμοποιηθούν δεδομένα από ένα url ή από μια sql βάση. Στη συνέχεια θα παρουσιάσουμε ένα παράδειγμα του Weka. Ανοίγοντας το Weka ο χρήστης έχει να ξεκινήσει τις κύριες εφαρμογές του :

- ✓ Explorer
- ✓ Experimenter
- ✓ Knowledge flow

**OExplorer** μπορεί να χαρακτηριστεί ως η πιο δημοφιλή επαφή. Εκεί ο χρήστης έχει την δυνατότητα εκτελεί όλες τις κύριες διεργασίες της εξόρυξης δεδομένων.

**OExperimenter** αποτελεί ένα περιβάλλον διεξαγωγής πειραμάτων. Εκεί αξιολογούνται οι μέθοδοι κατηγοριοποίησης και παλινδρόμησης .

**To Knowledgeflow** έχει τον ίδιο ρόλο με τον Explorer αλλά σε διαφορετική διεπαφή.

Το παράδειγμα που φαίνεται στην παραπάνω εικόνα είναι αρχείο με επέκταση ARFF.

Στην πρώτη σειρά βλέπουμε το @relationheart-diseases-simplified. Η λέξη @relation προϋπάρχει από το όνομα του στον πίνακα δεδομένων με όνομα heart-diseases- simplified. Ακολουθούν τα πεδία που δηλώνονται με @attribute όνομα πεδίου και τύπου πεδίου.

```

@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 3.3: Παράδειγμα αρχείου με επέκταση ARFF

Στο συγκεκριμένο παράδειγμα έχουμε έξι πεδία: Age, Sex, Chest pain, Cholesterol, Exercise induced angina, Class.

Μετά από τα πεδία ακολουθούν τα δεδομένα. Δηλώνονται με @data και οι τιμές που χωρίζονται με κόμμα. Για παράδειγμα στην άσκηση μας με τα δεδομένα είναι :

```

@data

63, male, typ_angina, 233, no, not-present

67, male, asympt, 286, yes, present

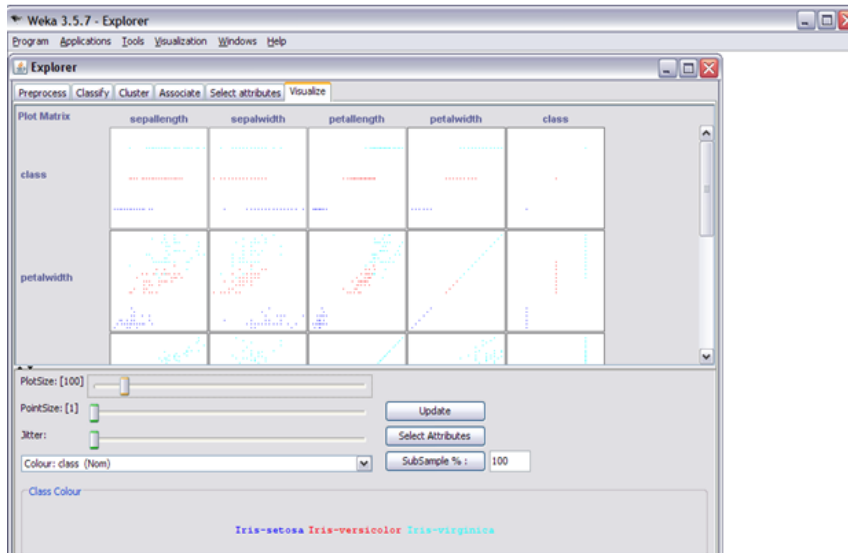
67, male, asympt, 229, yes, present

38, female, non_angina, ?, no, not_present

```

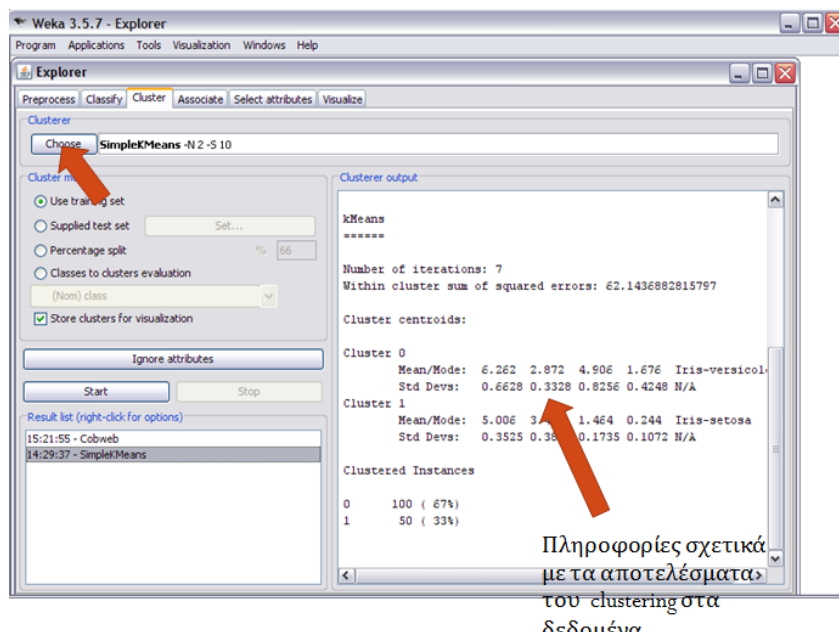
### 3.3 Οπτικοποίηση και Συσταδοποίηση δεδομένων

Στην εικόνα 3.4 βλέπουμε στο menu την καρτέλα visualize. Πατώντας στην παραπάνω καρτέλα υπάρχει η δυνατότητα εμφάνισης της γραφικής παράστασης κάθε γνώρισματος σε συνάρτηση με κάθε άλλο γνώρισμα.



Εικόνα 3.4: Οπτικοποίηση και Συσταδοποίηση δεδομένων

Αρχικά επιλέγεται ένα σύνολο δεδομένων που θα γίνει συσταδοποίηση δηλαδή εύρεση ομάδων με ίδια χαρακτηριστικά αυτό μπορεί να επιτευχθεί από την καρτέλα classified στο Weka 3.5.7 Explorer. Πατώντας στην καρτέλα cluster επιλέγεται ένας αλγόριθμος ο οποίος είναι υπεύθυνος για την συσταδοποίηση και πατώντας το κουμπί start θα αρχίσει να εκτελείται. Μερικοί από τους αλγορίθμους συσταδοποίησης :Cobweb, Dbscan, Em, Farthestfirst, Optics, Simplekmeans (k-means) και Xmeans[17].



Εικόνα 3.5: classified στο Weka 3.5.7 Explorer



# ΚΕΦΑΛΑΙΟ 4

## ΥΛΟΠΟΙΗΣΗ ΠΕΙΡΑΜΑΤΩΝ

---

### 4.1 Εξαγωγή χαρακτηριστικών

Για να εξαχθούν τα χαρακτηριστικά ενός προβλήματος μηχανικής μάθησης απαντήθηκαν οι παρακάτω ερωτήσεις σε ηλεκτρονικό ερωτηματολόγιο (Πίνακας 4.1). Η έρευνα διεξάγεται στο πλαίσιο πτυχιακής εργασίας του Τμήματος Μηχανικών Πληροφορικής του Ανώτατου Τεχνολογικού Εκπαιδευτικού Ιδρύματος Ηπείρου.

<p><b><u>1) Φύλο</u></b> Αντρας Γυναίκα</p> <p><b><u>2) Τι ηλικία έχετε;</u></b> &lt;18 18 με 35 35 με 50 50+</p> <p><b><u>3) Θα επιλέγατε κινητό τηλέφωνο σύμφωνα με το λογισμικό του;</u></b> Ναι Όχι</p> <p><b><u>4) Είστε προχωρημένος χρήστης στα κινητά τηλέφωνα;</u></b> Όχι το χρησιμοποιώ μόνο για τηλεφωνο Όχι γνωρίζω τα βασικά Ναι χρησιμοποιώ αρκετές εφαρμογές Ναι μπορώ να προγραμματίσω ο ίδιος εφαρμογές</p>	<p><b><u>10) Κάθε ποτέ αλλάζετε κινητό τηλέφωνο;</u></b> Κάθε χρόνο Κάθε δύο χρόνια Κάθε 3 χρόνια + Ανάλογα με το χρόνο ζωής του κινητού μου</p> <p><b><u>11) Αλλάζετε λόγω μόδας, τεχνολογίας η προβλήματος;</u></b> Το αλλάζω λόγω μόδας Το αλλάζω λόγω τεχνολογίας Το αλλάζω όποτε χαλάει και δεν συμφέρει να το φτιάξω η δεν φτιάχνεται</p> <p><b><u>12) Είστε άτομο το οποίο ακολουθεί γενικότερα τις τάσεις της μόδας?</u></b> Ναι Όχι</p> <p><b><u>13) Είστε υπέρ του ανοιχτού κώδικα που χρησιμοποιούν κάποια λογισμικά η του κλειστού κώδικα?</u></b> Είμαι υπέρ του ανοιχτού κώδικα Είμαι υπέρ του κλειστού κώδικα</p>
---	--

<p><b><u>5)Θα επιλέγατε κινητό τηλέφωνο ανάλογα με την ευκολία χρήσης του λογισμικού;</u></b></p> <p>Ναι Όχι</p>	<p><b><u>14) Χρησιμοποιείται το κινητό σας για να παίζεται παιχνίδια?</u></b></p> <p>Όχι Ναι λίγο Ναι πολύ</p>
<p><b><u>6) Θα επιλέγατε κινητό τηλέφωνο από την εξωτερική του εμφάνιση;</u></b></p> <p>Ναι Όχι</p>	<p><b><u>15) Χρησιμοποιείτε το κινητό σας ως ατζέντα?</u></b></p> <p>Ναι Όχι</p>
<p><b><u>7)Θα σας επηρέαζαν το κόστος στην επιλογή του κινητού σας;</u></b></p> <p>Ναι Όχι</p>	<p><b><u>16) Χρησιμοποιείτε τις παροχές που δίνουν στην εποχή μας οι υπηρεσίες Cloud (Αποθήκευση αρχείων και φωτογραφιών, ημερολόγιο κ.ο.κ)</u></b></p> <p>Ναι Όχι Άλλο</p>
<p><b><u>8)Ποιο ήταν το πρώτο λογισμικό που χρησιμοποιήσατε;</u></b></p> <p>Ios Android Windows</p>	<p><b><u>17)Είστε χρήστης iosandroid η windows?</u></b></p> <p>Ios Android Windows</p>
<p><b><u>9)Έχετε παρέμβει ποτέ στο κώδικα του λογισμικού σας η σκοπεύετε να παρέμβετε</u></b></p> <p>Έχω παρέμβει και σκοπεύω να το ξανακάνω Δεν έχω παρέμβει και ούτε έχω σκοπό Έχω παρέμβει αλλά δεν το ξανακάνω Δεν έχω παρέμβει αλλά σκέφτομαι να το κάνω</p>	

Πίνακας 4.1: οι ερωτήσεις από τις οποίες προήλθαν τα χαρακτηριστικά

## 4.2 Ηλεκτρονικό ερωτηματολόγιο

Στην πτυχιακή εργασία διεξαγάγαμε μια έρευνα όπου χρησιμοποιήσαμε ένα ερωτηματολόγιο του Googleforms. Για να δημιουργήσουμε ένα ερωτηματολόγιο στο Googleforms αρκεί να μπούμε στην διεύθυνση <https://forms.google.com>.

### **Τι μπορούμε να κάνουμε με τις φόρμες του Google?**

Με τις φόρμες αυτές μπορείτε να:

- Διαχειριστείτε καταχωρήσεις εκδηλώσεων
- Δημιουργήσετε μια γρήγορη δημοσκόπηση
- Δημιουργήσετε και να αναλύσετε έρευνες από το κινητό σας
- Δημιουργήσετε και να αναλύσετε έρευνες στο πρόγραμμα περιήγησης ιστού

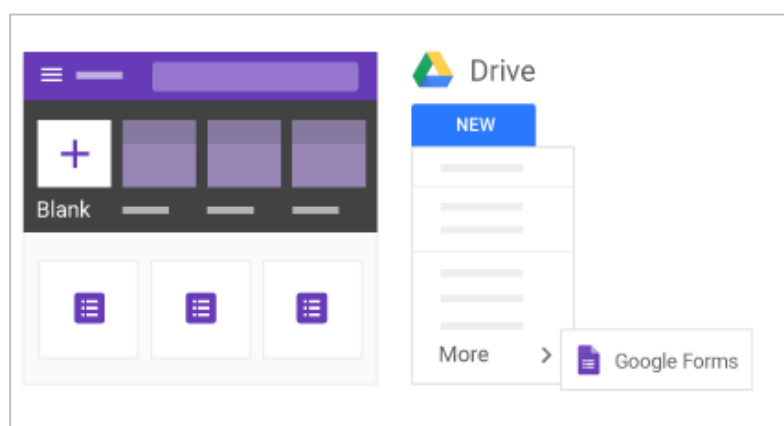
Θα πρέπει να τονίσουμε ότι δεν απαιτείται ειδικό λογισμικό για να πραγματοποιήσουμε όλα τα παραπάνω. Τα αποτελέσματα προκύπτουν άμεσα και παρουσιάζονται με τα κατάλληλα γραφήματα. Απαραίτητη προϋπόθεση για να ξεκινήσουμε με τις φόρμες του Google είναι:

- Η δημιουργία λογαριασμού G Suite
- Ένα χρονοδιάγραμμα 10 λεπτών

### **1<sup>ο</sup> Βήμα : Δημιουργία φόρμας**

Αρχικά θα πρέπει να δημιουργήσουμε μια νέα φόρμα και να προσθέσουμε τις ερωτήσεις τις οποίες επιθυμούμε. Η αποθήκευση των αλλαγών επιτυγχάνεται αυτόματα από τις φόρμες. Σε αυτό το σημείο ουσιαστικά μαθαίνουμε:

- Τη δημιουργία νέας φόρμας
- Την προσθήκη ερωτήσεων
- Την επεξεργασία ερωτήσεων



*Εικόνα 4.1: Βήματα παρουσίασης του GoogleForms*

**1. Για να δημιουργήσουμε νέα φόρμα θα πρέπει να επιλέξουμε μια επιλογή :**

- Στο forms.google.com κάνουμε κλικ στο κουμπί Blank ή επιλέγουμε template.
- Στο drive.google.com κάνουμε κλικ στο κουμπί New – More έπειτα δίπλα στο Google Forms στο δεξί βέλος και κάνουμε κλικ Blank form ή από το template.

## 2. Θα πρέπει να δώσετε όνομα στη φόρμα σας :

Κάντε κλικ στην επάνω αριστερή γωνία χωρίς τίτλο ή στο όνομα της φόρμας πρότυπου και πληκτρολογήστε το όνομα το οποίο επιθυμείτε.

## 3. Έπειτα κάντε μια από τις παρακάτω ενέργειες:

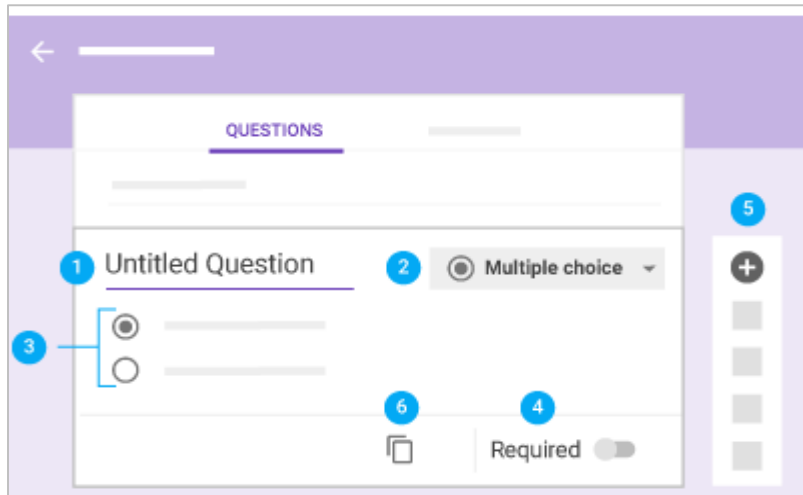
- Προσθήκη μιας περιγραφής: Κάτω από το όνομα της φόρμας προσθέστε το κείμενο σας.
- Αλλάξτε το χρώμα φόντου ή κάντε κλικ στην παλέτα χρώμα παλέτας και έπειτα επιλέξτε ένα χρώμα ή ένα θέμα που σας αρέσει προσθέστε ερωτήσεις.

Σε περίπτωση που χρησιμοποιήσετε ένα πρότυπο μπορείτε να μεταβείτε στην κατάλληλη ενότητα όπου σας δίνεται η δυνατότητα επεξεργασίας της φόρμας. Επίσης έχετε τη δυνατότητα να επιστρέψετε οποιαδήποτε στιγμή σε αυτό το σημείο σε περίπτωση που επιθυμείτε να προσθέσετε ή να επεξεργαστείτε υπάρχουσες ερωτήσεις. Κάντε κλικ στην ερώτηση Untitled και πατήστε enter your question. Σε αυτό το σημείο λαμβάνετε προτάσεις βάση του τύπου ερωτήματος σας.

1. Για να αλλάξετε τον τύπο ερώτησης σας κάντε κλικ στο κάτω βέλος του προγράμματος
2. Για να προσθέσετε επιλογές απάντησης
3. Για να καθορίσετε εάν οι χρήστες πρέπει να απαντήσουν στην ερώτηση κάντε κλικ στην επιλογή Required

### Για να προσθέσετε περισσότερες ερωτήσεις επιλέξτε μια επιλογή:

4. Για να προσθέσετε μια νέα ερώτηση κάντε κλικ στην επιλογή add\_circle
5. Για να προσθέσετε ένα αντίγραφο της υπάρχουσας ερώτησης κάντε κλικ στην επιλογή Duplicate



Εικόνα 4.2: Βήματα παρουσίασης του GoogleForms

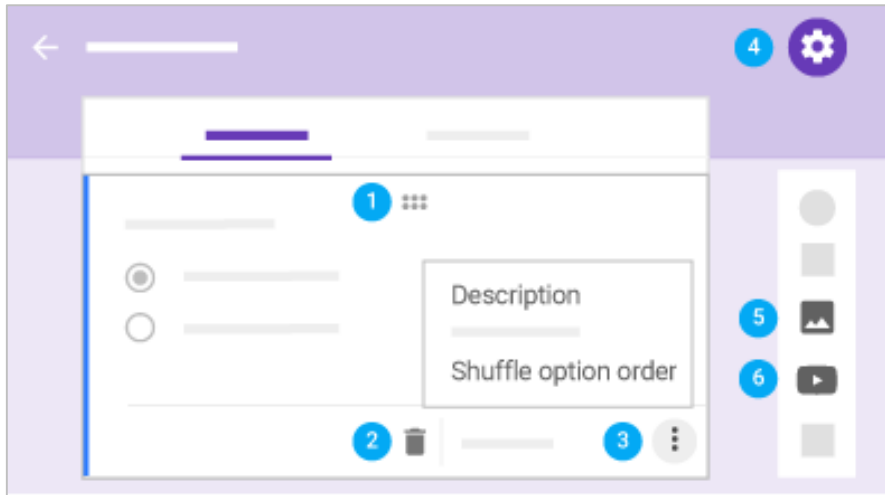
### Επεξεργασία ερωτήσεων

Μπορείτε με τη βοήθεια του GoogleForms να προσθέσετε στοιχεία όπως :

- *Εικόνα*
- *Βίντεο*
- *Ερώτηση*

Επιπλέον έχετε τη δυνατότητα να αναδιατάξετε καθώς και να διαγράψετε ερωτήσεις. Κατά τη διάρκεια της επεξεργασίας μιας ερώτησης μπορείτε να κάνετε μια από τις παρακάτω ενέργειες:

1. Σύρετε για να αναδιατάξετε μια ερώτηση. Μπορείτε επίσης να σύρετε για να αναδιατάξετε τις απαντήσεις.
2. Διαγράψτε μια ερώτηση κάντε κλικ στο κουμπί περισσότερα (More)
3. Προσθέστε μια περιγραφή ή υπόδειξη
4. Ανακατέψτε τη σειρά των απαντήσεων . Για να ανακατέψετε την παραγγελία ερωτήματος ανακατεύθυνσης κάντε κλικ στην επιλογή αποθήκευση.
5. Προσθέστε μια εικόνα σε μια ερώτηση. Επίσης υπάρχει η δυνατότητα προσθήκης εικόνων σε απαντήσεις ή φόρμες
6. Προσθέστε ένα βίντεο στο Youtube



Εικόνα 4.3: Βήματα παρουσίασης του GoogleForms

### **Επιλέξτε τις ρυθμίσεις φόρμας και την προεπισκόπηση**

Όταν τελειώσετε με την προσθήκη ερωτήσεων μπορείτε να επιλέξετε τις ρυθμίσεις φόρμας και να κάνετε επισκόπηση των αλλαγών σας πριν την αποστολή. Σε αυτή την ενότητα μαθαίνετε πώς να:

- *Επιλέξετε τις ρυθμίσεις φόρμας*
- *Προεπισκόπηση φόρμας*

### **Επιλέξτε τις ρυθμίσεις φόρμας**

Σε αυτό το σημείο μπορείτε να αποφασίσετε το οποίος επιθυμείτε να έχει πρόσβαση στη φόρμα σας και αν θα συλλέγει διευθύνσεις ηλεκτρονικού ταχυδρομείου από άλλα άτομα. Από προεπιλογή η κοινή χρήση φόρμας περιορίζεται και η συλλογή μηνυμάτων ηλεκτρονικού ταχυδρομείου είναι απενεργοποιημένη.

### **Επιλέξτε τις ρυθμίσεις φόρμας:**

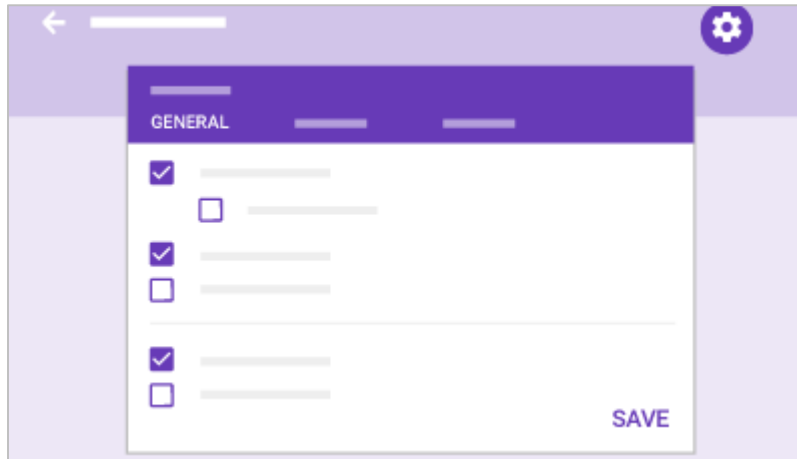
**Κάντε κλικ στο κουμπί ρυθμίσεις γενικά και επιλέξτε από τις ακόλουθες επιλογές:**

- Συλλογή διευθύνσεων ηλεκτρονικού ταχυδρομείου
- Συλλογή διευθύνσεων ηλεκτρονικού ταχυδρομείου των ερωτηθέντων
- Αποδείξεις απόκρισης

- Αποστολή αντιγράφων των απαντήσεων κατόπιν αιτήματος ή αυτόματα
- Περιορισμός στους χρήστες σας
- Κατάργηση της επιλογή αυτού του πλαισίου για τη διανομή της φόρμας στα εξωτερικά
- Περιορισμός σε μια απάντηση
- Να επιτρέπετε σε άτομα να συμπληρώσουν τη φόρμα μόνο μια φορά
- Επεξεργασία μετά την υποβολή
- Δυνατότητα αλλαγής απαντήσεων μετά την υποβολή
- Δείτε ανακεφαλαιωτικά διαγράμματα και απαντήσεις κειμένων
- Όλοι να έχουν τη δυνατότητα να δουν μια περίληψη των απαντήσεων όλων των υπολοίπων
- Σε περίπτωση ολοκλήρωσης των αλλαγών κάντε κλικ στο κουμπί αποθήκευση

Αφήστε τους χρήστες να ελέγξουν και να επεξεργαστούν τη φόρμα σας:

- Στην επάνω αριστερή γωνία κάντε κλικ στην επιλογή περισσότερα (More)
- Στην περιοχή πρόσκλησης ατόμων εισάγετε τις διευθύνσεις ηλεκτρονικού ταχυδρομείου των ατόμων που θέλετε να μοιραστείτε
- Κάντε κλικ στο κουμπί αποστολή
- Όποιοι έχετε καλέσει μπορούν να επεξεργαστούν οποιοδήποτε μέρος της φόρμας σας συμπεριλαμβανομένων των απαντήσεων και του τόπου όπου αποθηκεύονται



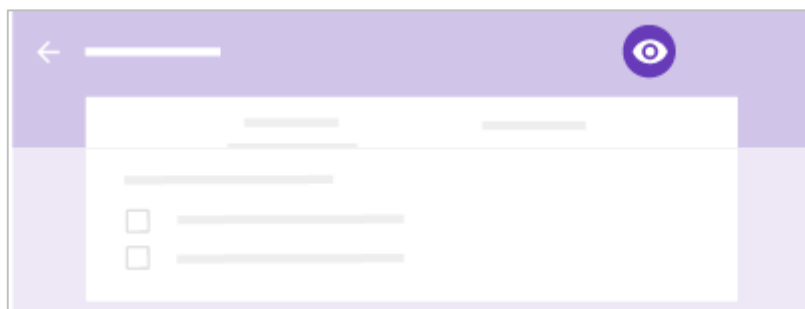
Εικόνα 4.4: Βήματα παρουσίασης του GoogleForms

### **Προεπισκόπηση της φόρμας**

Ενώ αλλάζετε τις ρυθμίσεις της φόρμας μπορείτε να κάνετε προεπισκόπηση και να δείτε τις αλλαγές. Επίσης υπάρχει αυτή η δυνατότητα και όταν τελειώσετε.

### **Προεπισκόπηση της φόρμας σας:**

Στην επάνω δεξιά γωνία κάντε κλικ στην προεπισκόπηση της προβολής. Η προεπισκόπηση ανοίγει ένα νέο παράθυρο. Για να επεξεργαστείτε τη φόρμα κάντε κλικ στην επιλογή επεξεργασία δημιουργίας ή επιστρέψτε στο παράθυρο επεξεργασίας.



Εικόνα 4.5: Βήματα παρουσίασης του GoogleForms

### **Στείλτε τη φόρμα σας**

Τώρα είστε έτοιμοι να στείλετε τη φόρμα σας και να αρχίσετε να συλλέγετε απαντήσεις. Σε αυτό το σημείο μαθαίνετε πώς να :



- *Στείλετε τη φόρμα*
- *Να κάνετε παύση*
- *Να κάνετε διακοπή της συλλογής απαντήσεων*

Μπορείτε να στείλετε τη φόρμα σας όταν είστε έτοιμοι υπάρχει η δυνατότητα αποστολής με ηλεκτρονικό ταχυδρομείο, να αντιγράψετε και να επικολλήσετε έναν σύνδεσμο σε ένα μήνυμα συνομιλίας ή σε ένα email έπειτα να την ενσωματώσετε σε έναν ιστότοπο ή να μοιραστείτε ένα σύνδεσμο στα μέσα κοινωνικής δικτύωσης.

Για να στείλετε τη φόρμα σας:

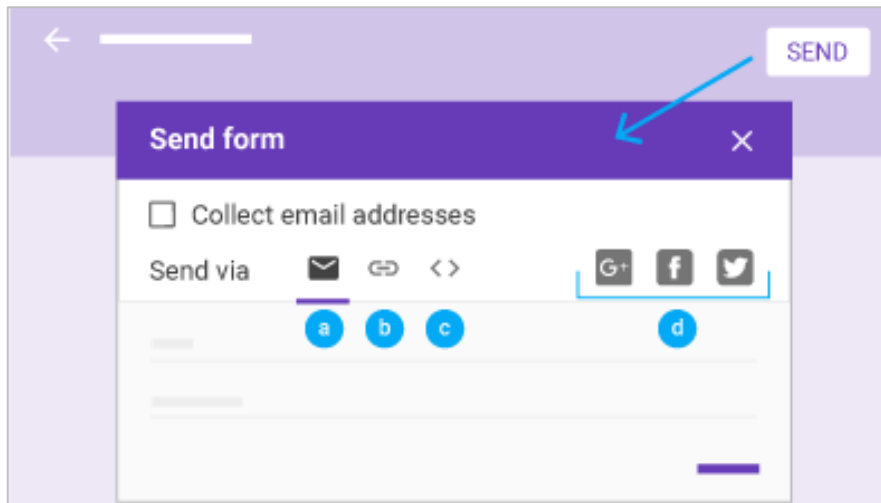
- Στην επάνω δεξιά γωνία κάντε κλικ στην επιλογή αποστολή.

Για τη συλλογή διευθύνσεων ηλεκτρονικού ταχυδρομείου επιλέξτε μια επιλογή:

- Αν το κοινό του έντυπου περιορίζεται επιλέξτε αυτόματα τη συλλογή της διεύθυνσης ηλεκτρονικού ταχυδρομείου της οργάνωσης του ερωτώμενου
- Εάν διανέμεται εξωτερικά τη φόρμα επιλέξτε συλλογή διευθύνσεων ηλεκτρονικού ταχυδρομείου

Επιλέξτε τον τρόπο αποστολής της φόρμας:

- ✓ Ηλεκτρονικό ταχυδρομείο με email κλικ . Εισάγετε τις διευθύνσεις ηλεκτρονικού ταχυδρομείου στο πεδίο προς μπορείτε να προσαρμόσετε το θέμα και το περιεχόμενο των μηνυμάτων
- ✓ Linkclickσύννεσηinsertlink . Μπορείτε να συντομεύσετε τη διεύθυνση URL κάντε κλικ στην επιλογή αντιγραφή και επικόλληση του συνδέσμου σε συζήτηση , συνομιλία ή ηλεκτρονικό ταχυδρομείο
- ✓ Περιεχόμενο ιστότοπου. Κάντε κλικ στην επιλογή ενσωμάτωση κώδικα. Μπορείτε σε αυτό το σημείο να ορίσετε διαστάσεις εσωτερικού πλαισίου. Κάντε κλικ στην επιλογή αντιγραφή και επικόλληση του HTML στον ιστότοπο ή στο ιστολόγιο σας
- ✓ Κοινωνικά μέσα. Κάντε κλικ σε ένα από τα εικονίδια των κοινωνικών μέσων.



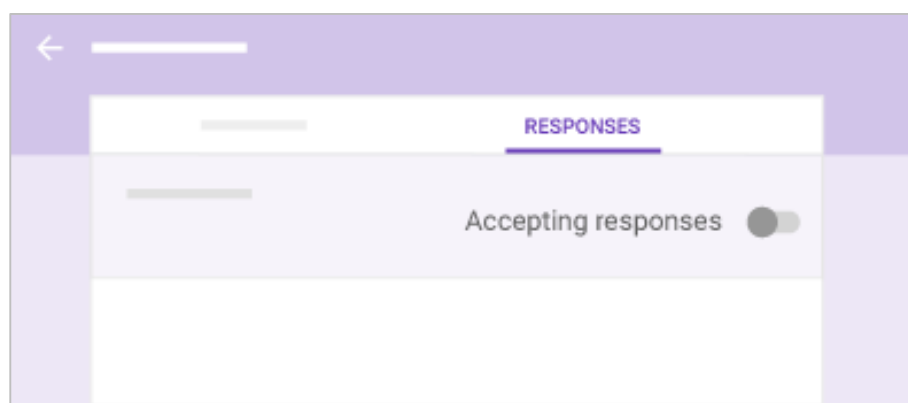
Εικόνα 4.6: Βήματα παρουσίασης του GoogleForms

### **Αναλύστε τις απαντήσεις**

Υπό την προϋπόθεση ότι έχετε στείλει τις απαντήσεις μπορείτε να τις δείτε στην καρτέλα απαντήσεις. Μπορείτε επίσης να λαμβάνετε μήνυμα ηλεκτρονικού ταχυδρομείου όταν φθάνουν οι νέες απαντήσεις.

### **Σε αυτό το σημείο θα μάθετε:**

- Πως μπορείτε να δείτε τις απαντήσεις σε έντυπα
- Πως μπορείτε να δείτε τις απαντήσεις σε φύλλα
- Πώς μπορείτε να λαμβάνετε τις απαντήσεις ως αρχείο CVS



Εικόνα 4.7: Βήματα παρουσίασης του GoogleForms

### **Πως μπορείτε να δείτε τις απαντήσεις σε έντυπα**

Στην καρτέλα απόκριση έχετε δύο δυνατότητες:

- Να δείτε τις απαντήσεις μεμονωμένες
- Να δείτε τις απαντήσεις σαν σύνοψη

Για να επιτύχετε εναλλαγή μεταξύ των επιμέρους απαντήσεων κάντε κλικ στο ηλεκτρονικό ταχυδρομείο του κάθε χρήστη.

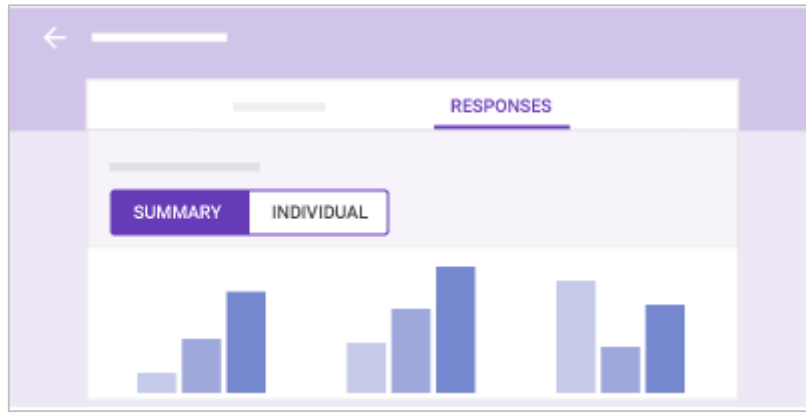
### **Πως μπορείτε να δείτε τις απαντήσεις σε φύλλα**

Για να αναλύσετε περαιτέρω τις απαντήσεις σας μπορείτε να τις στείλετε σε ένα υπολογιστικό φύλλο. Το υπολογιστικό φύλλο συνδέεται με τη φόρμα σας κατά συνέπεια οι απαντήσεις εμφανίζονται σε πραγματικό χρόνο. Για την αποστολή απαντήσεων σε νέο υπολογιστικό φύλλο πηγαίνετε:

- Στην καρτέλα απαντήσεις κλικ στο κουμπί επιλογή περισσότερα και επιλογή απάντησης προορισμού κάντε κλικ στην επιλογή δημιουργία νέου υπολογιστικού φύλλου.
- Για να αλλάξετε το όνομα , εισάγετε νέο υπολογιστικό φύλλο κάντε κλικ στην επιλογή δημιουργία

Αποστολή απαντήσεων σε ένα υπάρχον υπολογιστικό φύλλο:

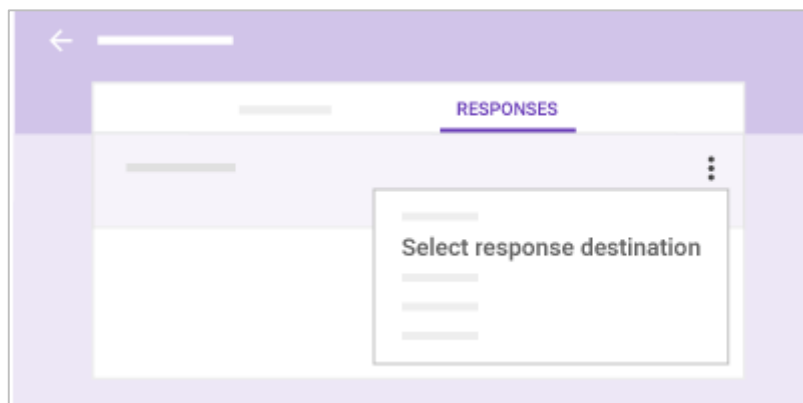
Στην καρτέλα απαντήσεις κάντε κλικ στην επιλογή περισσότερα επιλογή απάντησης προορισμού κάντε κλικ στην επιλογή υπάρχοντος υπολογιστικού φύλλου και κλικ στην επιλογή επιλέξτε υπολογιστικό φύλλο και επιλογή



Εικόνα 4.8: Βήματα παρουσίασης του GoogleForms

Για να ανοίξετε το υπολογιστικό φύλλο από τις φόρμες:

Κάντε κλικ στην επιλογή προβολή απαντήσεων σε φύλλα drive\_spreadsheet. Το υπολογιστικό φύλλο ανοίγει σε νέο παράθυρο.



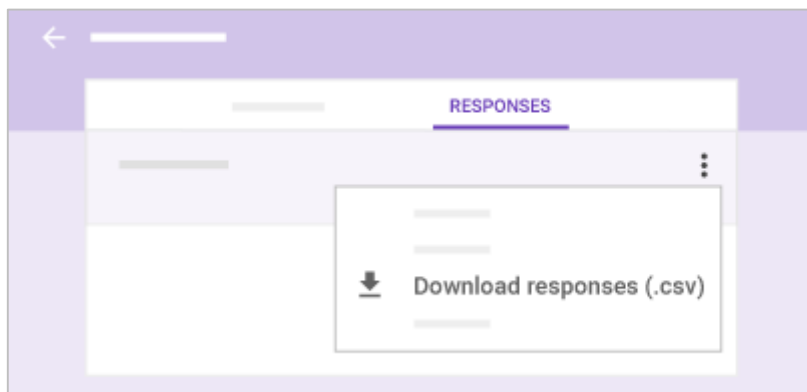
Εικόνα 4.9: Βήματα παρουσίασης του GoogleForms

Για να επιστρέψετε στις φόρμες:

Κάντε κλικ στην επιλογή φόρμα εμφάνιση σύνοψης απαντήσεων ή επιστρέψτε στο παράθυρο φόρμες

Για διακοπή αποστολής απαντήσεων σε υπολογιστικό φύλλο:

Στην καρτέλα απαντήσεις κάντε κλικ στην επιλογή περισσότερα και αποσύνδεση φόρμας. Κάντε κλικ στην επιλογή αποσύνδεση φόρμας. Κάντε κλικ στην επιλογή αποσύνδεση για επιβεβαίωση.



Εικόνα 4.10: Βήματα παρουσίασης του GoogleForms

### **Πώς μπορείτε να λαμβάνετε τις απαντήσεις ως αρχείο CVS**

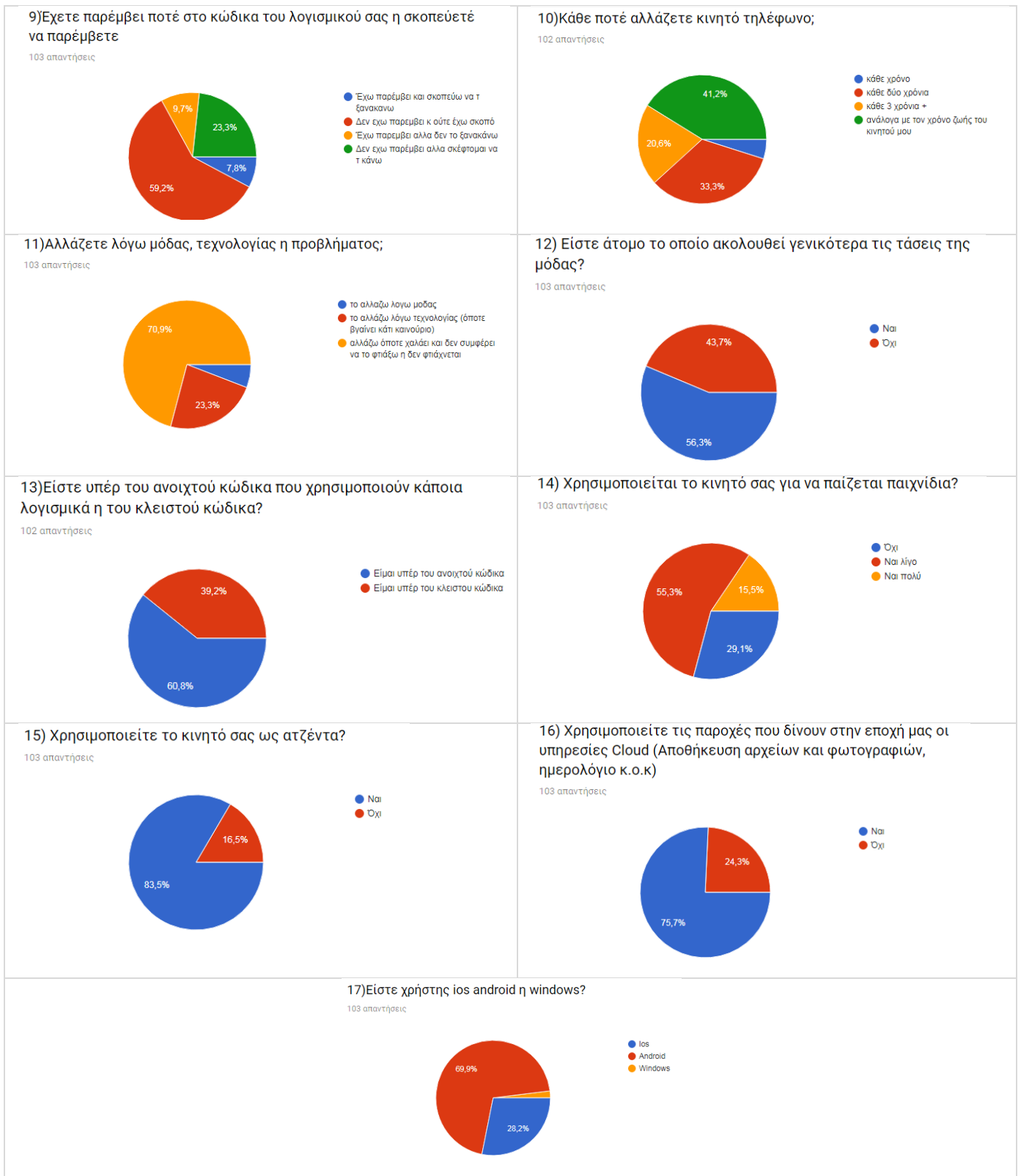
- Για να εξάγετε απαντήσεις για ανάλυση σε άλλα προγράμματα, μπορείτε να τις κατεβάσετε ως αρχείο CSV. Στην καρτέλα Απαντήσεις, κάντε κλικ στην επιλογή Περισσότερα και Λήψη απαντήσεων (.csv).[18]

# 5 ΑΠΟΤΕΛΕΣΜΑΤΑ

## 5.1 Στατιστικά στοιχεία δείγματος

Στα παρακάτω σχήματα συνοψίζεται το δείγμα το οποίο συγκεντρώθηκε από τα άτομα που συμμετείχαν στην έρευνα μας.





Εικόνα 5.1: Στατιστικά στοιχεία

## 5.2 Αποτελέσματα ταξινόμηση

### Επιλογή χαρακτηριστικών

#### K-κοντινότερος γείτονας (KNN = K-nearestneighbor)

Παρακάτω παρουσιάζονται τα αποτελέσματα της ακρίβειας πρόβλεψης για τον αλγόριθμο KNN με και χωρίς επιλογή χαρακτηριστικών:

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

Correctly Classified Instances      95          78.5124 %
Incorrectly Classified Instances    26          21.4876 %
Kappa statistic                    0.6093
Mean absolute error                 0.1668
Root mean squared error            0.3266
Relative absolute error             44.3573 %
Root relative squared error        75.4919 %
Total Number of Instances         121

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                0,483  0,109  0,583     0,483  0,528     0,400   0,822    0,611    ios
                0,847  0,306  0,803     0,847  0,824     0,550   0,862    0,894    android
                1,000  0,010  0,952     1,000  0,976     0,971   1,000    1,000    windows
Weighted Avg.   0,785  0,210  0,775     0,785  0,778     0,584   0,875    0,844

=== Confusion Matrix ===

 a  b  c  <-- classified as
14 15  0 | a = ios
10 61  1 | b = android
 0  0 20 | c = windows
```

Εικόνα 5.1: K-Κοντινότερος γείτονας

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

Correctly Classified Instances      89          73.5537 %
Incorrectly Classified Instances    32          26.4463 %
Kappa statistic                    0.5083
Mean absolute error                 0.2176
Root mean squared error            0.3466
Relative absolute error             57.8686 %
Root relative squared error        80.1225 %
Total Number of Instances         121

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                0,276  0,076  0,533     0,276  0,364     0,259   0,778    0,497    ios
                0,847  0,408  0,753     0,847  0,797     0,458   0,799    0,844    android
                1,000  0,050  0,800     1,000  0,889     0,872   0,995    0,969    windows
Weighted Avg.   0,736  0,269  0,708     0,736  0,709     0,479   0,826    0,782

=== Confusion Matrix ===

 a  b  c  <-- classified as
 8 20  1 | a = ios
 7 61  4 | b = android
 0  0 20 | c = windows
```

Εικόνα 5.2: K-Κοντινότερος γείτονας με selectedattributes



Παρακάτω παρουσιάζονται τα αποτελέσματα της ακρίβειας πρόβλεψης για τον αλγόριθμο NaïveBayesclassifier με και χωρίς επιλογή χαρακτηριστικών:

### Μπευζιανόταξινομητής (NaïveBayesclassifier)

```

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

Correctly Classified Instances      95          78.5124 %
Incorrectly Classified Instances    26          21.4876 %
Kappa statistic                    0.6135
Mean absolute error                 0.1918
Root mean squared error             0.321
Relative absolute error             51.0099 %
Root relative squared error         74.1983 %
Total Number of Instances          121

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                0,517   0,120   0,577     0,517   0,545     0,413   0,811    0,556    ios
                0,833   0,286   0,811     0,833   0,822     0,552   0,831    0,850    android
                1,000   0,010   0,952     1,000   0,976     0,971   1,000    1,000    windows
Weighted Avg.   0,785   0,200   0,778     0,785   0,781     0,588   0,854    0,804

=== Confusion Matrix ===

 a  b  c  <-- classified as
15 14  0 | a = ios
11 60  1 | b = android
 0  0 20 | c = windows

```

Εικόνα5.3: Naive Bayes

```

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

Correctly Classified Instances      87          71.9008 %
Incorrectly Classified Instances    34          28.0992 %
Kappa statistic                    0.4688
Mean absolute error                 0.2659
Root mean squared error             0.3713
Relative absolute error             70.7107 %
Root relative squared error         85.8231 %
Total Number of Instances          121

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                0,172   0,076   0,417     0,172   0,244     0,138   0,715    0,394    ios
                0,861   0,449   0,738     0,861   0,795     0,439   0,749    0,801    android
                1,000   0,050   0,800     1,000   0,889     0,872   0,980    0,782    windows
Weighted Avg.   0,719   0,294   0,671     0,719   0,678     0,438   0,779    0,700

=== Confusion Matrix ===

 a  b  c  <-- classified as
 5 22  2 | a = ios
 7 62  3 | b = android
 0  0 20 | c = windows

```

Εικόνα5.4: naive Bayes με selected attributes

Παρακάτω παρουσιάζονται τα αποτελέσματα της ακρίβειας πρόβλεψης για τον αλγόριθμο Decisiontrees με και χωρίς επιλογή χαρακτηριστικών:

## Δένδρα απόφασης (Decision trees)

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

Correctly Classified Instances      95          78.5124 %
Incorrectly Classified Instances    26          21.4876 %
Kappa statistic                    0.6043
Mean absolute error                 0.1991
Root mean squared error             0.3592
Relative absolute error             52.9411 %
Root relative squared error         83.0284 %
Total Number of Instances          121

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                0,414   0,065   0,667     0,414   0,511     0,418   0,669    0,427    ios
                0,875   0,327   0,797     0,875   0,834     0,566   0,751    0,751    android
                1,000   0,040   0,833     1,000   0,909     0,895   0,988    0,893    windows
Weighted Avg.   0,785   0,216   0,772     0,785   0,769     0,585   0,771    0,697

=== Confusion Matrix ===

 a  b  c  <-- classified as
12 16  1 | a = ios
 6 63  3 | b = android
 0  0 20 | c = windows
```

Εικόνα5.5: Decision trees

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

Correctly Classified Instances      89          73.5537 %
Incorrectly Classified Instances    32          26.4463 %
Kappa statistic                    0.5437
Mean absolute error                 0.2178
Root mean squared error             0.3675
Relative absolute error             57.9086 %
Root relative squared error         84.9467 %
Total Number of Instances          121

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                0,517   0,141   0,536     0,517   0,526     0,381   0,695    0,504    ios
                0,750   0,265   0,806     0,750   0,777     0,479   0,762    0,769    android
                1,000   0,059   0,769     1,000   0,870     0,851   0,965    0,749    windows
Weighted Avg.   0,736   0,202   0,735     0,736   0,732     0,517   0,779    0,702

=== Confusion Matrix ===

 a  b  c  <-- classified as
15 13  1 | a = ios
13 54  5 | b = android
 0  0 20 | c = windows
```

Εικόνα5.6: Decision trees με selected attributes

Παρακάτω παρουσιάζονται τα αποτελέσματα της ακρίβειας πρόβλεψης για τον αλγόριθμο SupportVectorMachines - SVM με και χωρίς επιλογή χαρακτηριστικών:

## Μηχανές Διανυσμάτων υποστήριξης (Support Vector Machines - SVM)

```

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

Correctly Classified Instances      93          76.8595 %
Incorrectly Classified Instances    28          23.1405 %
Kappa statistic                    0.5743
Mean absolute error                0.2736
Root mean squared error            0.3543
Relative absolute error            72.7598 %
Root relative squared error        81.8822 %
Total Number of Instances         121

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
          0,414   0,120   0,522     0,414   0,462     0,320   0,740    0,401    ios
          0,847   0,347   0,782     0,847   0,813     0,513   0,750    0,753    android
          1,000   0,000   1,000     1,000   1,000     1,000   1,000    1,000    windows
Weighted Avg.   0,769   0,235   0,756     0,769   0,760     0,547   0,789    0,710

=== Confusion Matrix ===

  a  b  c  <-- classified as
12 17  0 | a = ios
11 61  0 | b = android
 0  0 20 | c = windows

```

Εικόνα5.7: Support Vector Machines - SVM

```

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

Correctly Classified Instances      85          70.2479 %
Incorrectly Classified Instances    36          29.7521 %
Kappa statistic                    0.4425
Mean absolute error                0.2938
Root mean squared error            0.3809
Relative absolute error            78.1314 %
Root relative squared error        88.0413 %
Total Number of Instances         121

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
          0,172   0,120   0,313     0,172   0,222     0,067   0,612    0,291    ios
          0,833   0,449   0,732     0,833   0,779     0,404   0,692    0,709    android
          1,000   0,030   0,870     1,000   0,930     0,919   0,985    0,870    windows
Weighted Avg.   0,702   0,301   0,654     0,702   0,671     0,408   0,721    0,635

=== Confusion Matrix ===

  a  b  c  <-- classified as
 5 22  2 | a = ios
11 60  1 | b = android
 0  0 20 | c = windows

```

Εικόνα5.8: Support Vector Machines - SVM με selected attributes

Στον πίνακα 5.1 παρατίθενται συγκεντρωτικά τα αποτελέσματα από τους 4 ταξινομητές και με χρήση επιλογής χαρακτηριστικών αλλά και χωρίς

Αποτελέσματα Ακρίβειας Ταξινόμησης				
	KNN (%)	BAYES (%)	SVM (%)	Decision TREES (%)
Χωρίς επιλογή Χαρακτηριστικών	<b>78.51</b>	78.51	76.86	<b>78.51</b>
Με επιλογή Χαρακτηριστικών	<b>73.55</b>	71.90	70.25	<b>73.55</b>

Πίνακας 5.1: Συγκεντρωτικός πίνακας αποτελεσμάτων

# ΚΕΦΑΛΑΙΟ 6

## ΣΥΜΠΕΡΑΣΜΑΤΑ

---

Κατά την εκπόνηση της παρούσας πτυχιακής εργασίας ορίσαμε το λειτουργικό σύστημα δηλαδή το λογισμικό του υπολογιστή, το οποίο ως βασικό του ρόλο έχει τη διαχείριση καθώς και τον συντονισμό εργασιών. Επίσης ορίσαμε το λειτουργικό κινητής συσκευής δηλαδή ένα λειτουργικό σύστημα, το οποίο χρησιμοποιείται όχι απαραίτητως μόνο σε κινητές συσκευές αλλά και σε tablets, smartwatches, φορητούς υπολογιστές. Έπειτα γνωρίσαμε τα βασικά λειτουργικά συστήματα Android, ios και Windowsmobile. Στη συνέχεια κάναμε μια σύντομη ιστορική αναδρομή στα λειτουργικά σύστημα. Το 2010 εμφανίστηκαν τα πρώτα έξυπνα τηλέφωνα στην αγορά. Από τη μια πλευρά η BlackBerry βρισκόταν σε μια πολύ ισχυρή θέση, από την άλλη πλευρά το Symbian τροφοδοτούσε εκατομμύρια συσκευές και από την άλλη πλευρά η Samsung ερχόταν να κάνει τη δική της πρόταση στην αγορά. Μετά από 7 χρόνια οι ρόλοι έχουν αντιστραφεί και πλέον την κυριαρχία έχουν δύο μόνο πλατφόρμες. Η iOS της Apple και Android της Google. Σύμφωνα με επίσημα στατιστικά στοιχεία του τρέχοντος έτους η Gartner στο 99% κυριαρχούν οι παραπάνω πλατφόρμες. Από την άλλη πλευρά η BlackBerry και Microsoft τείνουν να χαθούν από την αγορά.

Στη συνέχεια μελετήσαμε τις ευφυείς μέθοδοι μηχανικής μάθησης. Αρχικά ορίσαμε τη μηχανική μάθηση και το datamining . Ως μηχανική μάθηση ορίσαμε τη μελέτη υπολογιστικών μεθόδων που ως σκοπό έχει την απόκτηση νέας γνώσης .έπειτα γνωρίσαμε τις μεθόδους ταξινόμησης: κοντινότερο γείτονα, NaïveBayes, τα δέντρα απόφασης και τις μηχανές διανυσμάτων υποστήριξης. Στη συνέχεια γνωρίσαμε το WEKA (WaikatoEnvironmentforKnowledgeAnalysis), το οποίο αποτελεί ένα πολύ σημαντικό εργαλείο μηχανικής μάθησης και εξόρυξης δεδομένων.Τέλος εξάγαμε τα χαρακτηριστικά και δημιουργήσαμε ένα ηλεκτρονικό ερωτηματολόγιο του Googleforms και συγκεντρώσαμε τα στατιστικά στοιχεία του δείγματος και τα αποτελέσματα της ταξινόμησης.

Σύμφωνα με τα αποτελέσματα της ταξινόμησης, το ερωτηματολόγιο που καταρτήθηκε μπορεί να καθοδηγήσει με μια ακρίβεια περίπου 80% των χρήστη στο λειτουργικό που θα χρησιμοποιήσεις σε φορητές συσκευές. Το ποσοστό αυτό μπορεί να επιτευχθεί κάνοντας χρήση όλων των ερωτήσεων ως χαρακτηριστικά του προβλήματος ταξινόμησης, και εκτελώντας τον Μπεϋζιανόταξινομητή και τον ταξινομητή του κοντινότερου γείτονα. Τέλος, από τα αποτελέσματα φαίνεται ότι η επιλογή χαρακτηριστικών δεν βελτιώνει την ακρίβεια αλλά την μειώνει. Ως εκ τούτου γίνεται φανερό ότι όλες οι ερωτήσεις περιέχουν επιθυμητή πληροφορία και δεν χρειάζεται να απορρίψουμε καμιά από αυτές.

# ΒΙΒΛΙΟΓΡΑΦΙΑ

---

- [1] WikipediaΛειτουργικό σύστημα. Διαθέσιμο στο δικτυακό ιστότοπο:  
[.https://el.wikipedia.org/wiki/%CE%9B%CE%B5%CE%B9%CF%84%CE%BF%CF%85%CF%81%CE%B3%CE%B9%CE%BA%CF%8C\\_%CF%83%CF%8D%CF%83%CF%84%CE%B7%CE%BC%CE%B1](https://el.wikipedia.org/wiki/%CE%9B%CE%B5%CE%B9%CF%84%CE%BF%CF%85%CF%81%CE%B3%CE%B9%CE%BA%CF%8C_%CF%83%CF%8D%CF%83%CF%84%CE%B7%CE%BC%CE%B1)
- [2] Τι είναι το Android. Διαθέσιμο στο δικτυακό ιστότοπο:  
<http://www.doctorandroid.gr/p/iphone.html>
- [3] The advantages and disadvantages of android mobile phones .Διαθέσιμο στο δικτυακό ιστότοπο: <https://www.online-sciences.com/technology/the-advantages-and-disadvantages-of-android-mobile-phones-2/>
- [4] Wikipedia. IOS Διαθέσιμο στο δικτυακό ιστότοπο:  
<https://el.wikipedia.org/wiki/IOS>
- [5] advantages and disadvantages of iOS .Διαθέσιμο στο δικτυακό ιστότοπο:  
<https://thetechhacker.com/2015/01/08/advantages-disadvantages-apple-ios/>
- [6] Wikipedia. Windows Mobile. Διαθέσιμο στο δικτυακό ιστότοπο:  
[https://el.wikipedia.org/wiki/Windows\\_Mobile](https://el.wikipedia.org/wiki/Windows_Mobile)
- [7] What are the advantages and disadvantages of owning a windows phone?.Διαθέσιμο στο δικτυακό ιστότοπο: <https://www.quora.com/What-are-the-advantages-and-disadvantages-of-owning-a-windows-phone>
- [8] So... 10 reasons why NOT Windows Phone, and what needs to happen!.Διαθέσιμο στο δικτυακό ιστότοπο:[http://allaboutwindowsphone.com/features/item/20223\\_so\\_10\\_reasons\\_why\\_not\\_windows\\_.php](http://allaboutwindowsphone.com/features/item/20223_so_10_reasons_why_not_windows_.php)
- [9] Wikipedia. Mobile\_operating\_system. Διαθέσιμο στο δικτυακό ιστότοπο:  
[https://en.wikipedia.org/wiki/Mobile\\_operating\\_system](https://en.wikipedia.org/wiki/Mobile_operating_system)
- [10] Απόλυτη κυριαρχία για Android και iOS στην αγορά των smartphones
- [11] .Διαθέσιμο στο δικτυακό ιστότοπο: <http://gr.pcmag.com/smartphones-1/27965/feature/apolute-kuriarkhia-gia-android-kai-ios-sten-agora-ton-smartp>
- [12] Μηχανική μάθηση. Διαθέσιμο στο δικτυακό ιστότοπο:  
<http://slideplayer.gr/slide/2300475/>

- [13] Wikipedia. Εξόρυξη δεδομένων. Διαθέσιμο στο δικτυακό ιστότοπο:  
[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](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)
- [14] K Nearest Neighbors – Classification .Διαθέσιμο στο δικτυακό ιστότοπο:  
[http://www.saedsayad.com/k\\_nearest\\_neighbors.htm](http://www.saedsayad.com/k_nearest_neighbors.htm)
- [15] Naive\_bayesian. Διαθέσιμο στο δικτυακό ιστότοπο:  
[http://www.saedsayad.com/naive\\_bayesian](http://www.saedsayad.com/naive_bayesian)
- [16] svm classifier, introduction to support vector machine algorithm
- [17] . Διαθέσιμο στο δικτυακό ιστότοπο:
- [18] <http://dataaspirant.com/2017/01/13/support-vector-machine-algorithm/>
- [19] Οδηγός του WEKA. Διαθέσιμο στο δικτυακό ιστότοπο:  
[https://repository.kallipos.gr/bitstream/11419/1239/2/Kef.\\_13.pdf](https://repository.kallipos.gr/bitstream/11419/1239/2/Kef._13.pdf)
- [20] Google Forms. Διαθέσιμο στο δικτυακό ιστότοπο:  
<https://gsuite.google.com/learning-center/products/forms/get-started/>