



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ  
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

**Εφαρμογή Αλγορίθμων Μηχανικής Μάθησης και  
Πειραματική Αποτίμηση**

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

του

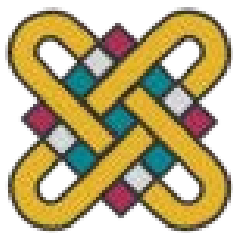
**ΠΑΤΙΑΗ ΟΔΥΣΣΕΑ**

(ΑΕΜ: 2799)

**Επιβλέπων : Νικόλαος Δημόκας**  
**Επίκουρος Καθηγητής**

Καστοριά Οκτώβριος - 2023

Η παρούσα σελίδα σκοπίμως παραμένει λευκή



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

# Εφαρμογή Αλγορίθμων Μηχανικής Μάθησης και Πειραματική Αποτίμηση

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

του

**ΠΑΤΙΑΗ ΟΔΥΣΣΕΑ**

(ΑΕΜ: 2799 )

Επιβλέπων : **Νικόλαος Δημόκας**  
**Επίκουρος Καθηγητής**

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την **02-10-2023**

.....  
Δημόκας Νικόλαος  
Επίκουρος Καθηγητής

.....  
Ιωάννης Βαρδάκας  
Αναπληρωτής Καθηγητής

.....  
Σπυρίδων Νικολάου  
Λέκτορας

Καστοριά **Οκτώβριος - 2023**

Copyright © 2023 – ΠΑΤΙΛΗΣ ΟΔΥΣΣΕΑΣ

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

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν αποκλειστικά τον συγγραφέα και δεν αντιπροσωπεύουν τις επίσημες θέσεις του Πανεπιστημίου Δυτικής Μακεδονίας.

Ως συγγραφέας της παρούσας εργασίας δηλώνω πως η παρούσα εργασία δεν αποτελεί προϊόν λογοκλοπής και δεν περιέχει υλικό από μη αναφερόμενες πηγές.

## Ευχαριστίες

Με την παρούσα πτυχιακή εργασία ολοκληρώνονται οι σπουδές μου στο Τμήμα Πληροφορικής της σχολής Θετικών Επιστημών του Πανεπιστημίου Δυτικής Μακεδονίας.

Θα ήθελα να ευχαριστήσω ιδιαίτερα τον επιβλέποντα καθηγητή μου και καθηγητή της παρούσας εργασίας κ.Νικόλαο Δημόκα για την επιστημονική και συμβουλευτική καθοδήγηση που μου προσέφερε κατά την διάρκεια εκπόνησης της εργασίας.

Τέλος θα ήθελα να ευχαριστήσω του γονείς μου για την στήριξη που μου παρείχαν καθόλη την διάρκεια των σπουδών μου.

## Περίληψη

---

Η Μηχανική Μάθηση είναι ένα από τα πιο δημοφιλή υποπεδία της Τεχνητής Νοημοσύνης. Οι έννοιες μηχανικής μάθησης χρησιμοποιούνται σχεδόν παντού, όπως στην Υγεία, στα Οικονομικά, στις Υποδομές, στο Μάρκετινγκ και σε πολλά άλλα. Η Μηχανική Μάθηση είναι ένας κλάδος της Τεχνητής Νοημοσύνης που επιτρέπει στις μηχανές να μαθαίνουν και να βελτιώνονται αυτόματα από την εμπειρία. Ορίζεται ως το αντικείμενο που δίνει στους υπολογιστές τη δυνατότητα να μαθαίνουν χωρίς να είναι ρητά προγραμματισμένοι. Είναι αρκετά διαφορετικό από τον παραδοσιακό προγραμματισμό. Αν και η μηχανική μάθηση εξελίσσεται συνεχώς με τόσες πολλές νέες τεχνολογίες, εξακολουθεί να χρησιμοποιείται σε διάφορους κλάδους. Η μηχανική μάθηση είναι σημαντική επειδή δίνει στις επιχειρήσεις μια άποψη για τις τάσεις στη συμπεριφορά των πελατών και τα λειτουργικά επιχειρηματικά πρότυπα, καθώς και υποστηρίζει την ανάπτυξη νέων προϊόντων. Πολλές από τις κορυφαίες εταιρείες του σήμερα, όπως το Facebook, η Google και η Uber, κάνουν τη μηχανική μάθηση κεντρικό μέρος των λειτουργιών τους. Η μηχανική μάθηση έχει γίνει ένας σημαντικός ανταγωνιστικός παράγοντας διαφοροποίησης για πολλές εταιρείες.

**Λέξεις Κλειδιά:** μηχανική μάθηση, αλγόριθμοι κατηγοριοποίησης, βιβλιοθήκη λογισμικού, διεπαφή προγράμματος εφαρμογής

## Abstract

---

Machine Learning is one of the most popular subfields of Artificial Intelligence. Machine learning concepts are used almost everywhere like in Healthcare, Finance, Infrastructure, Marketing and many more. Machine Learning is a branch of Artificial Intelligence that allows machines to learn and improve automatically from experience. It is defined as the field of study that gives computers the capability to learn without being explicitly programmed. It is quite different than traditional programming. Although machine learning is constantly evolving with so many new technologies, it is still used in various industries. Machine learning is important because it gives businesses insight into trends in customer behavior and operational business patterns, as well as supporting new product development. Many of today's top companies, such as Facebook, Google and Uber, are making machine learning a central part of their operations. Machine learning has become an important competitive differentiator for many companies.

***Key Words: machine learning, classification algorithms, library, API***

## Πίνακας Περιεχομένων

---

### Contents

Εισαγωγή.....	1
1. Αλγόριθμοι Κατηγοροποίησης .....	2
<b>1.1 Logistic Regression</b> .....	2
1.1.1 Διαισθηση .....	2
1.1.2 Μαθηματική Διατύπωση.....	2
1.1.3 Εκπαίδευση και Βελτιστοποίηση .....	3
1.1.4 Εφαρμογές .....	3
<b>1.2 Decision Trees</b> .....	4
1.2.1 Διαισθηση .....	4
1.2.2 Κατασκευή Δέντρων Απόφασης .....	4
1.2.3 Τεχνικές Κλαδέματος (Pruning Techniques) .....	4
1.2.4 Εφαρμογές .....	5
<b>1.3 Random Forest</b> .....	6
1.3.1 Διαισθηση .....	6
1.3.2 Διαδικασία Κατασκευής.....	6
1.3.3 Εκτίμηση Σημαντικότητας Χαρακτηριστικών .....	7
1.3.4 Εφαρμογές .....	7
<b>1.4 Naïve Bayes</b> .....	8
1.4.1 Διαισθηση .....	8
1.4.2 Μαθηματικά Θεμέλια .....	8
1.4.3 Διαδικασία Εκπαίδευσης .....	8
1.4.4 Χειρισμός Συνεχών και Κατηγορικών Δεδομένων .....	9
1.4.5 Εφαρμογές .....	9
2. OpenCV και Tensorflow Lite .....	12
<b>2.1 OpenCV</b> .....	12
2.1.1 Αρχιτεκτονική και Σχεδίαση .....	12
2.1.2 APIs και Χρήση .....	15
2.1.3 Εφαρμογές στο Πραγματικό Κόσμο .....	17



<b>2.2 TensorFlow Lite</b> .....	19
<b>2.2.1 Αρχιτεκτονική και Σχεδίαση</b> .....	19
<b>2.2.2 APIs και Χρήση</b> .....	21
<b>2.2.3 Εφαρμογές στο Πραγματικό Κόσμο</b> .....	22
<b>3. ML Kit</b> .....	25
<b>3.1 Αρχιτεκτονική και Σχεδίαση</b> .....	25
<b>3.2 APIs και Χρήση</b> .....	26
<b>3.2 Εφαρμογές στον Πραγματικό Κόσμο</b> .....	28
<b>4. Υλοποίηση Εφαρμογής</b> .....	31
<b>4.1 Λειτουργικό Σύστημα Android</b> .....	31
<b>4.1.1 Βασικά Χαρακτηριστικά</b> .....	31
<b>4.2 Υλοποίηση Εφαρμογής</b> .....	32
<b>4.2.1 UI</b> .....	33
<b>4.2.2 Capture</b> .....	33
<b>4.2.3 Detect</b> .....	34
<b>4.2.4 Translator</b> .....	36
<b>4.2.5 Object Detector</b> .....	43
Συμπεράσματα.....	48
Βιβλιογραφία.....	49

## Λίστα Εικόνων

---

Εικόνα 1: Main Activity .....	33
Εικόνα 2: Capture code .....	33
Εικόνα 3: Detect code .....	34
Εικόνα 4: Το αποτέλεσμα του button Detect .....	35
Εικόνα 5: Translator Activity .....	36
Εικόνα 6: Επιλογή Γλώσσας .....	37
Εικόνα 7: Translate code .....	37
Εικόνα 8: Αποτέλεσμα Μετάφρασης στα Ελληνικά .....	39
Εικόνα 9: Αποτέλεσμα Μετάφρασης στα Κορεάτικα .....	40
Εικόνα 10: Mic code .....	40
Εικόνα 11: Microphone View .....	41
Εικόνα 12: Speaker Code .....	42
Εικόνα 13: Object Detector .....	43
Εικόνα 14- Object Detection code .....	44
Εικόνα 15- Ανάλυση Εικόνας .....	46

## Εισαγωγή

---

Η παρούσα εργασία αφορά τη μελέτη ορισμένων αλγορίθμων μηχανικής μάθησης και την αξιοποίησή τους προκειμένου να υλοποιήσουμε χρήσιμες εφαρμογές για τη βελτίωση της καθημερινότητά μας.

Η εργασία αποτελείται από τέσσερα κεφάλαια. Το πρώτο κεφάλαιο αναφέρεται στους αλγόριθμους κατηγοριοποίησης. Οι αλγόριθμοι κατηγοριοποίησης είναι μία κατηγορία αλγορίθμων μηχανικής μάθησης που χρησιμοποιείται για την κατηγοριοποίηση ή την επισήμανση δεδομένων σε διακριτές κλάσεις ή κατηγορίες με βάση τα χαρακτηριστικά εισόδου. Αυτοί οι αλγόριθμοι εντάσσονται στη μάθηση με επίβλεψη και συνεπώς μαθαίνουν από δεδομένα εκπαίδευσης (που γνωρίζουμε ήδη την κατηγορία/ ετικέτα στην οποία ανήκουν) τα οποία στη συνέχεια χρησιμοποιούν προκειμένου να προβλέψουν τις ετικέτες κλάσης νέων, άγνωστων δεδομένων.

Το δεύτερο κεφάλαιο αναλύει τις βιβλιοθήκες OpenCV και Tensorflow Lite. Η OpenCV (Open Source Computer Vision Library) και η TensorFlow Lite είναι δύο δημοφιλείς βιβλιοθήκες που χρησιμοποιούνται στον τομέα της όρασης υπολογιστών και της μηχανικής μάθησης, αντίστοιχα. Το OpenCV χρησιμοποιείται ευρέως σε ακαδημαϊκή έρευνα, εμπορικές εφαρμογές και σε προσωπικά έργα(project) λόγω της εκτεταμένης λειτουργικότητας, της ευκολίας χρήσης και της ενεργής υποστήριξης της κοινότητας. Το TensorFlow Lite χρησιμοποιείται ευρέως για ένα ευρύ φάσμα εφαρμογών, συμπεριλαμβανομένων εφαρμογών για κινητές συσκευές, έξυπνων συσκευών, ρομποτικής και άλλων ενσωματωμένων συστημάτων, όπου απαιτούνται δυνατότητες μηχανικής εκμάθησης σε πραγματικό χρόνο ή εκτός σύνδεσης (offline).

Το τρίτο κεφάλαιο αναλύει την βιβλιοθήκη μηχανικής μάθησης του λειτουργικού συστήματος Android, ML Kit. Το ML Kit είναι ένα SDK (Kit ανάπτυξης λογισμικού) για κινητά που παρέχεται από την Google και φέρνει δυνατότητες μηχανικής μάθησης σε εφαρμογές Android και iOS. Επιτρέπει στους προγραμματιστές να ενσωματώνουν λειτουργίες μηχανικής μάθησης στις εφαρμογές τους για κινητά χωρίς να απαιτούν εις βάθος γνώση αλγορίθμων ή μοντέλων μηχανικής μάθησης. Το ML Kit αφαιρεί την πολυπλοκότητα της μηχανικής μάθησης και παρέχει απλά API, διευκολύνοντας τους προγραμματιστές να αξιοποιήσουν τη μηχανική μάθηση στις εφαρμογές τους.

Τέλος στο τέταρτο κεφάλαιο δίνεται η περιγραφή της εφαρμογής που δημιουργήθηκε χρησιμοποιώντας μια μίξη των βιβλιοθηκών ML Kit, TensorFlow Lite και OpenCV.

## 1. Αλγόριθμοι Κατηγοριοποίησης

Οι αλγόριθμοι κατηγοροποίησης στην μηχανική μάθηση [1, 2] χρησιμοποιούνται για την αντιστοίχιση ετικετών ή κατηγοριών για την εισαγωγή δεδομένων με βάση τα χαρακτηριστικά ή τα γνωρισματά τους. Αυτοί οι αλγόριθμοι αναλύουν μοτίβα και σχέσεις μεταξύ των δεδομένων για να δημιουργήσουν μοντέλα που μπορούν με ακρίβεια να κατηγοροποιήσουν καινούργια instances. Σε αυτό το κεφάλαιο αναλύονται τεσσσερείς αλγόριθμοι που χρησιμοποιούνται πιο συχνά:

1. Logistic Regression
2. Decision Trees
3. Random forest
4. Naïve Bayes

### 1.1 Logistic Regression

Η λογιστική παλινδρόμηση[1-2] είναι ένας ευρέως χρησιμοποιούμενος αλγόριθμος κατηγοροποίησης στη μηχανική μάθηση και τη στατιστική. Είναι ένας εποπτευόμενος αλγόριθμος μάθησης που μοντελοποιεί τη σχέση μεταξύ των μεταβλητών εισόδου και την πιθανότητα να ανήκουν σε μια συγκεκριμένη τάξη. Σε αντίθεση με τη γραμμική παλινδρόμηση, η οποία προβλέπει συνεχείς τιμές, η λογιστική παλινδρόμηση προβλέπει την πιθανότητα ενός δυαδικού ή κατηγορικού αποτελέσματος. Σε αυτή την περιγραφή θα διερευνήσουμε τη διαίσθηση πίσω από την λογιστική παλινδρόμηση, τη μαθηματική της διατύπωση και τις εφαρμογές της.

#### 1.1.1 Διαίσθηση

Η διαίσθηση (μια συνάρτηση που μετρά την απόδοση ενός μοντέλου μηχανικής μάθησης για τα δοθείσα δεδομένα) πίσω από την λογιστική παλινδρόμηση βρίσκεται στη μοντελοποίηση της σχέσης μεταξύ των μεταβλητών εισόδου και της πιθανότητας να ανήκουν σε μια συγκεκριμένη κλάση. Ο αλγόριθμος στοχεύει να βρει την καλύτερη προσαρμοσμένη γραμμή (ή υπερεπίπεδο σε υψηλότερες διαστάσεις) που διαχωρίζει τα σημεία δεδομένων διαφορετικών κατηγοριών στο feature space[8] (n-διαστάσεις όπου ζουν οι μεταβλητές). Το κάνει μετατρέποντας την έξοδο μιας γραμμικής παλινδρόμησης σε πιθανότητα χρησιμοποιώντας μια μη γραμμική συνάρτηση που ονομάζεται λογιστική συνάρτηση ή σιγμοειδής συνάρτηση.

#### 1.1.2 Μαθηματική Διατύπωση

Ας εξετάσουμε ένα πρόβλημα δυαδικής ταξινόμησης με δύο κατηγορίες: 0 και 1. Η λογιστική παλινδρόμηση μοντελοποιεί τη σχέση μεταξύ των χαρακτηριστικών εισόδου ( $x$ ) και της πιθανότητας να ανήκει στην κλάση 1 ( $p$ ). Η λογιστική συνάρτηση ορίζεται ως εξής:  $p=1/(1+\exp(-z))$  όπου  $z$  είναι ένας γραμμικός συνδυασμός των χαρακτηριστικών εισόδου και των αντίστοιχων βαρών τους(συντελεστές):  $z=w_0+w_1x_1+w_2x_2+\dots+w_nx_n$ . Τα βάρη  $w$  και η τομή  $w_0$  είναι οι παράμετροι του μοντέλου λογιστικής παλινδρόμησης. Ο

στόχος του αλγορίθμου είναι να βρει τις βέλτιστες τιμές για αυτές τις παραμέτρους που μεγιστοποιούν την πιθανότητα των παρατηρούμενων δεδομένων.

### 1.1.3 Εκπαίδευση και Βελτιστοποίηση

Για την εκπαίδευση ενός μοντέλου λογιστικής παλινδρόμησης, χρησιμοποιείται ένας αλγόριθμος βελτιστοποίησης για την εκτίμηση των καλύτερων τιμών για τις παραμέτρους. Η εκτίμηση μέγιστης πιθανότητας (MLE) χρησιμοποιείται συνήθως για το σκοπό αυτό. Το MLE εκτιμά τις παραμέτρους που μεγιστοποιούν την πιθανότητα των παρατηρούμενων δεδομένων του μοντέλου.

Υπάρχουν διάφοροι αλγόριθμοι βελτιστοποίησης που μπορούν να χρησιμοποιηθούν για την εύρεση των βέλτιστων τιμών παραμέτρων, όπως η μέθοδος κλίσης, η μέθοδος του Νεύτωνα ή η στοχαστική κλίση. Αυτοί οι αλγόριθμοι ενημερώνουν επαναληπτικά τις τιμές των παραμέτρων με βάση τη διαβάθμιση της αντικειμενικής συνάρτησης, η οποία είναι συνήθως η αρνητική πιθανότητα καταγραφής.

### 1.1.4 Εφαρμογές

Η λογιστική παλινδρόμηση βρίσκει εφαρμογές σε διάφορους τομείς, όπως η υγειονομική περίθαλψη, τα οικονομικά, το μάρκετινγκ και οι κοινωνικές επιστήμες. Μερικές κοινές εφαρμογές περιλαμβάνουν:

1. Διάγνωση Ασθενειών : Η λογιστική παλινδρόμηση μπορεί να χρησιμοποιηθεί για τη διάγνωση της νόσου με βάση τα αποτελέσματα των ιατρικών εξετάσεων και τα χαρακτηριστικά του ασθενούς. Εκπαιδύοντας ένα μοντέλο λογιστικής παλινδρόμησης σε ιστορικά δεδομένα, μπορεί να προβλέψει την πιθανότητα ένας ασθενής να έχει μια συγκεκριμένη ασθένεια, βοηθώντας τους γιατρούς στη λήψη αποφάσεων.
2. Ανάλυση Συναισθημάτων: Η λογιστική παλινδρόμηση εφαρμόζεται στην ανάλυση συναισθήματος για να ταξινομήσει το κείμενο ως θετικό ή αρνητικό συναίσθημα. Με την εκπαίδευση σε δεδομένα με ετικέτα, το μοντέλο μπορεί να μάθει να εντοπίζει το συναίσθημα σε κριτικές πελατών, αναρτήσεις στα μέσα κοινωνικής δικτύωσης ή άλλα δεδομένα κειμένου, βοηθώντας στην κατανόηση της κοινής γνώμης.
3. Εκτίμηση Πιστωτικού Κινδύνου: Η λογιστική παλινδρόμηση χρησιμοποιείται στην αξιολόγηση πιστωτικού κινδύνου για την πρόβλεψη της πιθανότητας αθέτησης ενός δανείου από έναν πελάτη. Αναλύοντας διάφορα χαρακτηριστικά, όπως το πιστωτικό ιστορικό, το εισόδημα και τα δημογραφικά στοιχεία, το μοντέλο μπορεί να εκτιμήσει την πιθανότητα αθέτησης υποχρεώσεων και να βοηθήσει τους δανειστές να λαμβάνουν τεκμηριωμένες αποφάσεις.
4. Πρόβλεψη Απόκλισης Πελατών: Η λογιστική παλινδρόμηση χρησιμοποιείται στην πρόβλεψη απόκλισης πελατών, όπου ο στόχος είναι να εντοπιστούν πελάτες που είναι πιθανό να σταματήσουν να χρησιμοποιούν ένα προϊόν ή μια υπηρεσία.

Αναλύοντας τη συμπεριφορά των πελατών, τις μετρήσεις αφοσίωσης και τις δημογραφικές πληροφορίες, το μοντέλο μπορεί να εκτιμήσει την πιθανότητα απόρριψης και να ενεργοποιήσει προληπτικές στρατηγικές διατήρησης.

## 1.2 Decision Trees

Τα δέντρα αποφάσεων [1-2] είναι ισχυρά και ερμηνεύσιμα μοντέλα μηχανικής μάθησης που μπορούν να χρησιμοποιηθούν τόσο για εργασίες ταξινόμησης όσο και για εργασίες παλινδρόμησης. Παρέχουν μια γραφική αναπαράσταση μιας διαδικασίας λήψης αποφάσεων με αναδρομική κατάτμηση του χώρου εισόδου με βάση τις τιμές των χαρακτηριστικών εισόδου. Σε αυτήν την περιγραφή, θα διερευνήσουμε τη διαίσθηση πίσω από τα δέντρα αποφάσεων, τα κριτήρια κατασκευής και διαχωρισμού τους, τις τεχνικές κλαδέματος και τις εφαρμογές τους σε διάφορους τομείς.

### 1.2.1 Διαίσθηση

Η διαίσθηση πίσω από τα δέντρα αποφάσεων πηγάζει από τις ανθρώπινες διαδικασίες λήψης αποφάσεων. Φανταστείτε να παίρνετε μια απόφαση εξετάζοντας μια σειρά ερωτήσεων και προϋποθέσεων. Κάθε ερώτηση ή προϋπόθεση βοηθά στον περιορισμό των πιθανών αποτελεσμάτων μέχρι να ληφθεί μια τελική απόφαση. Τα δέντρα αποφάσεων μιμούνται αυτή τη διαδικασία κατασκευάζοντας μια δομή που μοιάζει με δέντρο, όπου κάθε εσωτερικός κόμβος αντιπροσωπεύει μια απόφαση ή μια συνθήκη και κάθε κόμβος φύλλου αντιπροσωπεύει μια τελική απόφαση ή αποτέλεσμα.

### 1.2.2 Κατασκευή Δέντρων Απόφασης

Η κατασκευή ενός δέντρου αποφάσεων περιλαμβάνει την αναδρομική κατάτμηση του χώρου εισόδου με βάση τις τιμές των χαρακτηριστικών εισόδου. Σε κάθε βήμα, ο αλγόριθμος επιλέγει το καλύτερο χαρακτηριστικό και μια τιμή κατωφλίου για να χωρίσει τα δεδομένα. Αυτό το κριτήριο διαχωρισμού καθορίζεται με βάση ένα μέτρο ακαθαρσίας [8, 9] (Η impurity function μετρά την έκταση καθαρότητας για μια περιοχή που περιέχει σημεία δεδομένων από πιθανώς διαφορετικές κατηγορίες) ή κέρδους πληροφοριών. Τα κοινά μέτρα ακαθαρσίας περιλαμβάνουν την ακαθαρσία Gini [10, 11] και την εντροπία [12] (είναι ένα μέτρο του επίπεδου διαταραχής ή αβεβαιότητας σε ένα δεδομένο σύνολο δεδομένων ή σύστημα).

Ο στόχος είναι να βρεθούν διαχωρισμοί που ελαχιστοποιούν την ακαθαρσία, με αποτέλεσμα καθαρά ή ομοιογενή υποσύνολα δεδομένων σε κάθε κόμβο φύλλου. Υπάρχουν διάφοροι αλγόριθμοι για την κατασκευή δέντρων αποφάσεων, όπως ID3, C4.5, CART και Random Forest. Αυτοί οι αλγόριθμοι διαφέρουν ως προς τα κριτήρια διαχωρισμού, τον χειρισμό των τιμών που λείπουν και τις τεχνικές κλαδέματος (Pruning techniques).

### 1.2.3 Τεχνικές Κλαδέματος (Pruning Techniques)

Τα δέντρα αποφάσεων έχουν την τάση να προσαρμόζονται υπερβολικά στα δεδομένα εκπαίδευσης, με αποτέλεσμα πολύπλοκα δέντρα που μπορεί να μην γενικεύονται καλά σε μη ορατά δεδομένα. Το κλάδεμα είναι μια τεχνική [13] που χρησιμοποιείται για την αντιμετώπιση της υπερπροσαρμογής μειώνοντας το μέγεθος του δέντρου. Δύο κοινές τεχνικές κλαδέματος είναι το προ-κλάδεμα (pre-pruning) και το μετα-κλάδεμα (post-pruning).

Το προ-κλάδεμα περιλαμβάνει τον καθορισμό κριτηρίων διακοπής κατά τη διαδικασία κατασκευής του δέντρου, όπως ο περιορισμός του μέγιστου βάθους του δέντρου, η απαίτηση ελάχιστου αριθμού δειγμάτων σε κόμβους φύλλων ή η απαίτηση ελάχιστης βελτίωσης στη μέτρηση ακαθαρσιών για τη διάσπαση ενός κόμβου. Το μετα-κλάδεμα, γνωστό και ως κλάδεμα πολυπλοκότητας κόστους, περιλαμβάνει την ανάπτυξη ενός πλήρους δέντρου και στη συνέχεια την επαναληπτική αφαίρεση κόμβων που οδηγούν σε ελάχιστη υποβάθμιση της απόδοσης σε ένα σύνολο επικύρωσης [14] (Validation set είναι ένα σύνολο δεδομένων που χρησιμοποιούνται για την εκπαίδευση της τεχνητής νοημοσύνης (Artificial Intelligence - AI) με στόχο την εύρεση και τη βελτιστοποίηση του καλύτερου μοντέλου για την επίλυση ενός δεδομένου προβλήματος).

#### 1.2.4 Εφαρμογές

Τα δέντρα αποφάσεων βρίσκουν εφαρμογές σε διάφορους τομείς λόγω της απλότητας, της ερμηνευσιμότητας και της ικανότητάς τους να χειρίζονται τόσο κατηγορικά όσο και αριθμητικά χαρακτηριστικά. Μερικές κοινές εφαρμογές περιλαμβάνουν:

##### 1. Εκτίμηση Πιστωτικού Κινδύνου

Τα δέντρα αποφάσεων μπορούν να χρησιμοποιηθούν για την αξιολόγηση πιστωτικού κινδύνου, όπου ο στόχος είναι να προβλεφθεί η πιθανότητα αθέτησης ενός δανείου από έναν πελάτη. Αναλύοντας τα χαρακτηριστικά των πελατών, το ιστορικό απασχόλησης, το εισόδημα και άλλους σχετικούς παράγοντες, ένα μοντέλο δέντρου αποφάσεων μπορεί να ταξινομήσει τους πελάτες σε κατηγορίες υψηλού ή χαμηλού κινδύνου.

##### 2. Διάγνωση Ασθενειών

Τα δέντρα αποφάσεων χρησιμοποιούνται σε συστήματα ιατρικής διάγνωσης για να βοηθήσουν τους γιατρούς να λαμβάνουν τεκμηριωμένες αποφάσεις. Λαμβάνοντας υπόψη τα συμπτώματα των ασθενών, τα αποτελέσματα ιατρικών εξετάσεων και άλλα κλινικά δεδομένα, ένα δέντρο αποφάσεων μπορεί να προτείνει πιθανές διαγνώσεις ή επιλογές θεραπείας.

##### 3. Τμηματοποίηση Πελατών

Τα δέντρα αποφάσεων μπορούν να βοηθήσουν στην τμηματοποίηση πελατών, η οποία περιλαμβάνει τη διαίρεση των πελατών σε ξεχωριστές ομάδες με βάση τα χαρακτηριστικά, τις συμπεριφορές ή τις προτιμήσεις τους. Αναλύοντας τα δημογραφικά στοιχεία πελατών, το ιστορικό αγορών και τη συμπεριφορά περιήγησης, τα δέντρα αποφάσεων μπορούν να προσδιορίσουν τμήματα για στοχευμένες καμπάνιες μάρκετινγκ ή εξατομικευμένες προτάσεις.

##### 4. Ανίχνευση Απάτης

Τα δέντρα αποφάσεων χρησιμοποιούνται σε συστήματα ανίχνευσης απάτης για τον εντοπισμό δυνητικά δόλιων συναλλαγών ή δραστηριοτήτων. Αναλύοντας μοτίβα

συναλλαγών, συμπεριφορά χρήστη και άλλα σχετικά δεδομένα, τα δέντρα αποφάσεων μπορούν να επισημάνουν ύποπτες συναλλαγές για περαιτέρω διερεύνηση.

### 1.3 Random Forest

Το Random Forest [2] είναι μια ισχυρή και ευέλικτη μέθοδος μάθησης συνόλου που συνδυάζει πολλαπλά δέντρα αποφάσεων για τη βελτίωση της ακρίβειας πρόβλεψης και τη μείωση της υπερπροσαρμογής. Χρησιμοποιείται ευρέως τόσο για εργασίες κατηγοριοποίησης όσο και για εργασίες παλινδρόμησης. Σε αυτήν την περιγραφή, θα διερευνήσουμε τη διαίσθηση πίσω από το Random Forest, τη διαδικασία κατασκευής του, την εκτίμηση της σημασίας των χαρακτηριστικών και τις εφαρμογές του σε διάφορους τομείς.

#### 1.3.1 Διαίσθηση

Η διαίσθηση πίσω από το Random Forest πηγάζει από την ιδέα του συνδυασμού πολλαπλών δέντρων αποφάσεων για να γίνουν συλλογικές προβλέψεις. Κάθε δέντρο απόφασης στο Random Forest εκπαιδεύεται σε ένα τυχαίο υποσύνολο των δεδομένων εκπαίδευσης και ένα υποσύνολο των διαθέσιμων χαρακτηριστικών. Στη συνέχεια, οι προβλέψεις από τα μεμονωμένα δέντρα συνδυάζονται για να ληφθεί η τελική πρόβλεψη. Αυτή η προσέγγιση συνόλου βοηθά στη μείωση της μεροληψίας και της διακύμανσης του μοντέλου, οδηγώντας σε βελτιωμένη απόδοση γενίκευσης [14, 15] (Η γενίκευση αναφέρεται στην ικανότητα του μοντέλου σας να προσαρμόζεται σωστά σε νέα, προηγουμένως αόρατα δεδομένα, που προέρχονται από την ίδια κατανομή με αυτή που χρησιμοποιήθηκε για τη δημιουργία του μοντέλου).

#### 1.3.2 Διαδικασία Κατασκευής

Η κατασκευή ενός Random Forest περιλαμβάνει δύο βασικά βήματα: τη δημιουργία μεμονωμένων δέντρων αποφάσεων και τον συνδυασμό των προβλέψεών τους. Ακολουθεί μια επισκόπηση της διαδικασίας:

1. Bootstrapped Sampling

Το Random Forest χρησιμοποιεί μια τεχνική που ονομάζεται bootstrapped sampling ή bagging, όπου δημιουργούνται πολλαπλά υποσύνολα των δεδομένων εκπαίδευσης με τυχαία δειγματοληψία με αντικατάσταση. Κάθε υποσύνολο, γνωστό και ως δείγμα εκκίνησης, έχει το ίδιο μέγεθος με το αρχικό σύνολο εκπαίδευσης, αλλά μπορεί να περιέχει διπλότυπα στιγμιότυπα.

2. Τυχαία Υποδειγματοληψία Χαρακτηριστικών

Εκτός από τη δειγματοληψία των δεδομένων, το Random Forest επιλέγει επίσης τυχαία ένα υποσύνολο χαρακτηριστικών που θα ληφθούν υπόψη κατά την κατασκευή κάθε δέντρου αποφάσεων. Αυτή η τυχαία υποδειγματοληψία χαρακτηριστικών βοηθά στην εισαγωγή της διαφορετικότητας μεταξύ των δέντρων και μειώνει τη συσχέτιση μεταξύ τους.

3. Δόμηση Δέντρων Απόφασης



Με τα bootstrapped δείγματα και τα τυχαία υποσύνολα χαρακτηριστικών, τα δέντρα αποφάσεων κατασκευάζονται ανεξάρτητα για κάθε υποσύνολο. Τα δέντρα μπορούν να κατασκευαστούν χρησιμοποιώντας οποιονδήποτε αλγόριθμο δέντρου αποφάσεων όπως ID3, C4.5 ή CART. Ωστόσο, μια κοινή επιλογή είναι η χρήση του αλγόριθμου CART με ένα τροποποιημένο κριτήριο διαχωρισμού που λαμβάνει υπόψη ένα τυχαίο υποσύνολο χαρακτηριστικών σε κάθε κόμβο.

#### 4. Συνδυασμός Προβλέψεων

Μόλις κατασκευαστούν όλα τα δέντρα απόφασης, οι προβλέψεις από τα μεμονωμένα δέντρα συνδυάζονται για να γίνει η τελική πρόβλεψη. Για εργασίες κατηγοριοποίησης, αυτό μπορεί να γίνει με πλειοψηφία, όπου επιλέγεται η τάξη με τις περισσότερες ψήφους σε όλα τα δέντρα. Για εργασίες παλινδρόμησης, υπολογίζεται ο μέσος όρος των προβλέψεων από διαφορετικά δέντρα για να ληφθεί η τελική πρόβλεψη.

### 1.3.3 Εκτίμηση Σημαντικότητας Χαρακτηριστικών

Το Random Forest παρέχει ένα μέτρο της σημασίας των χαρακτηριστικών, το οποίο υποδεικνύει τη σχετική συμβολή κάθε χαρακτηριστικού στην προγνωστική απόδοση του μοντέλου. Η σημασία ενός χαρακτηριστικού εκτιμάται με βάση τη μείωση της ακρίβειας ή της ακαθαρσίας του μοντέλου όταν αυτό το χαρακτηριστικό μετατίθεται τυχαία. Τα χαρακτηριστικά που έχουν μεγαλύτερο αντίκτυπο στην απόδοση του μοντέλου θεωρούνται πιο σημαντικά.

### 1.3.4 Εφαρμογές

Το Random Forest βρίσκει εφαρμογές σε διάφορους τομείς λόγω της ευελιξίας, της στιβαρότητας και της ικανότητάς του να χειρίζεται δεδομένα υψηλών διαστάσεων. Μερικές κοινές εφαρμογές περιλαμβάνουν :

#### 1. Πιστωτική Βαθμολόγηση

Το Random Forest χρησιμοποιείται στη βαθμολογία πιστοληπτικής ικανότητας για την πρόβλεψη της πιστοληπτικής ικανότητας των ατόμων με βάση το οικονομικό τους ιστορικό, τα πιστωτικά χαρακτηριστικά και τις δημογραφικές πληροφορίες τους. Η φύση του συνόλου του Random Forest βοηθά στην ακριβή εκτίμηση του πιστωτικού κινδύνου και στη μείωση των πιθανοτήτων εσφαλμένης ταξινόμησης

#### 2. Βιοπληροφορική

Το Random Forest εφαρμόζεται στη βιοπληροφορική για εργασίες όπως η ανάλυση γονιδιακής έκφρασης, η πρόβλεψη της δομής των πρωτεϊνών και ο εντοπισμός γενετικών δεικτών που σχετίζονται με την ασθένεια. Μπορεί να χειριστεί βιολογικά δεδομένα υψηλών διαστάσεων και να παρέχει πληροφορίες για πολύπλοκες βιολογικές διεργασίες.

#### 3. Τηλεπισκόπηση

Το Random Forest χρησιμοποιείται σε εφαρμογές τηλεπισκόπησης για ταξινόμηση κάλυψης γης, τμηματοποίηση εικόνων και ανίχνευση αλλαγών. Η ικανότητά του να χειρίζεται πολυφασματικά και υπερφασματικά δεδομένα το καθιστά κατάλληλο για την ανάλυση δορυφορικών εικόνων και την εξαγωγή σημαντικών πληροφοριών.

#### 4. Συστήματα Προτάσεων (Recommender Systems)

Το Random Forest μπορεί να χρησιμοποιηθεί σε συστήματα προτάσεων για την παροχή εξατομικευμένων προτάσεων στους χρήστες. Αναλύοντας τις προτιμήσεις των χρηστών, τα χαρακτηριστικά των στοιχείων και άλλα σχετικά δεδομένα, το Random Forest μπορεί να προβλέψει τις προτιμήσεις των χρηστών και να προτείνει στοιχεία που είναι πιθανό να έχουν ενδιαφέρον.

## 1.4 Naïve Bayes

Ο Naive Bayes[1-2] είναι ένας δημοφιλής και αποτελεσματικός αλγόριθμος μηχανικής μάθησης που βασίζεται στις αρχές του θεωρήματος του Bayes. Χρησιμοποιείται συνήθως για εργασίες ταξινόμησης, ειδικά σε εφαρμογές επεξεργασίας φυσικής γλώσσας και εξόρυξης κειμένου. Σε αυτήν την περιγραφή, θα διερευνήσουμε τη διαίσθηση πίσω από το Naive Bayes, τις μαθηματικές του βάσεις, τη διαδικασία εκπαίδευσης, τον χειρισμό συνεχών και κατηγορικών δεδομένων και τις εφαρμογές του σε διάφορους τομείς.

### 1.4.1 Διαίσθηση

Η διαίσθηση πίσω από το Naive Bayes πηγάζει από το θεώρημα του Bayes, το οποίο παρέχει έναν τρόπο ενημέρωσης προηγούμενων πεποιθήσεων με βάση νέα στοιχεία. Ο Naive Bayes υποθέτει ότι τα χαρακτηριστικά είναι υπό όρους ανεξάρτητα δεδομένης της μεταβλητής κλάσης. Αυτή η απλουστευτική υπόθεση επιτρέπει στον αλγόριθμο να κάνει προβλέψεις αποτελεσματικά, ακόμη και με μεγάλο αριθμό χαρακτηριστικών. Ο αλγόριθμος υπολογίζει την εκ των υστέρων πιθανότητα κάθε κλάσης δεδομένων των χαρακτηριστικών εισόδου και επιλέγει την κλάση με την υψηλότερη πιθανότητα ως πρόβλεψη.

### 1.4.2 Μαθηματικά Θεμέλια

Τα μαθηματικά θεμέλια του Naive Bayes έχουν τις ρίζες τους στη θεωρία πιθανοτήτων και στην υπό όρους πιθανότητα. Το θεώρημα του Bayes δηλώνει ότι η μεταγενέστερη πιθανότητα μιας τάξης δεδομένων των χαρακτηριστικών μπορεί να υπολογιστεί χρησιμοποιώντας την προηγούμενη πιθανότητα της κλάσης, την πιθανότητα των χαρακτηριστικών που δίνονται στην κλάση και τα στοιχεία που παρέχονται από τα χαρακτηριστικά. Μαθηματικά, μπορεί να αναπαρασταθεί ως:  $P(\text{class}|\text{features}) = (P(\text{features}|\text{class}) * P(\text{class})) / P(\text{features})$

### 1.4.3 Διαδικασία Εκπαίδευσης

Η διαδικασία εκπαίδευσης του Naive Bayes περιλαμβάνει την εκτίμηση των προηγούμενων πιθανοτήτων κάθε τάξης και των πιθανοτήτων των χαρακτηριστικών που δίνονται σε κάθε τάξη. Οι προηγούμενες πιθανότητες μπορούν να υπολογιστούν μετρώντας τις συχνότητες κάθε τάξης στα δεδομένα εκπαίδευσης. Οι πιθανότητες εκτιμώνται διαφορετικά με βάση τον τύπο των δεδομένων:

### Κατηγορικά Δεδομένα

Για τα κατηγορικά δεδομένα, οι πιθανότητες εκτιμώνται υπολογίζοντας τη συχνότητα κάθε τιμής χαρακτηριστικού σε κάθε κλάση και κανονικοποιώντας την με το συνολικό αριθμό παρουσιών σε αυτήν την κλάση. Η εξομάλυνση Laplace[16](είναι μια τεχνική εξομάλυνσης που βοηθά στην αντιμετώπιση του προβλήματος της μηδενικής πιθανότητας στον αλγόριθμο μηχανικής μάθησης Naive Bayes) εφαρμόζεται συχνά για να χειριστεί το ζήτημα των μηδενικών πιθανοτήτων.

### Συνεχή Δεδομένα

Για συνεχή δεδομένα, οι πιθανότητες τυπικά μοντελοποιούνται χρησιμοποιώντας κατανομές πιθανοτήτων όπως η Gaussian (Κανονική) κατανομή[17] (είναι μια συνεχής κατανομή πιθανοτήτων που χρησιμοποιείται ευρέως στη στατιστική μοντελοποίηση και στη Μηχανική Μάθηση). Ο μέσος όρος και η τυπική απόκλιση κάθε τιμής χαρακτηριστικού σε κάθε κατηγορία υπολογίζονται από τα δεδομένα εκπαίδευσης.

## **1.4.4 Χειρισμός Συνεχών και Κατηγορικών Δεδομένων**

Ο Naive Bayes μπορεί να χειριστεί τόσο συνεχή όσο και κατηγορηματικά χαρακτηριστικά. Για συνεχή χαρακτηριστικά, συχνά διαχωρίζονται σε bins[18](μια διαδικασία που χρησιμοποιείται για την ελαχιστοποίηση των επιπτώσεων των σφαλμάτων παρατήρησης) ή μοντελοποιούνται χρησιμοποιώντας κατανομές πιθανοτήτων. Για τα κατηγορικά χαρακτηριστικά, αντιμετωπίζονται ως διακριτές μεταβλητές και οι πιθανότητές τους υπολογίζονται με βάση τη συχνότητα κάθε τιμής χαρακτηριστικού σε κάθε κατηγορία.

Ο Naive Bayes αναλαμβάνει την ανεξαρτησία μεταξύ των χαρακτηριστικών δεδομένης της μεταβλητής κλάσης, κάτι που δεν ισχύει πάντα σε σενάρια πραγματικού κόσμου. Παρά αυτή την απλοποιητική υπόθεση, ο Naive Bayes συχνά αποδίδει καλά στην πράξη και μπορεί να προσφέρει λογικά αποτελέσματα.

## **1.4.5 Εφαρμογές**

Ο Naive Bayes έχει βρει εφαρμογές σε διάφορους τομείς, ειδικά στην ταξινόμηση κειμένου και το φιλτράρισμα ανεπιθύμητων μηνυμάτων. Μερικές κοινές εφαρμογές περιλαμβάνουν:

### 1. Κατηγοριοποίηση Κειμένου

Το Naive Bayes χρησιμοποιείται ευρέως για εργασίες κατηγοριοποίηση κειμένου, όπως ανάλυση συναισθήματος, κατηγοριοποίηση εγγράφων και φιλτράρισμα ανεπιθύμητων μηνυμάτων. Λαμβάνοντας υπόψη τις συχνότητες των λέξεων ή άλλων χαρακτηριστικών στο κείμενο, ο Naive Bayes μπορεί να ταξινομήσει τα έγγραφα σε διαφορετικές κατηγορίες.

### 2. Φιλτράρισμα ανεπιθύμητης αλληλογραφίας

Το Naive Bayes χρησιμοποιείται συνήθως σε συστήματα φιλτραρίσματος ανεπιθύμητων μηνυμάτων ηλεκτρονικού ταχυδρομείου. Αναλύοντας το περιεχόμενο, το θέμα και άλλα χαρακτηριστικά των email, ο Naive Bayes μπορεί να αναγνωρίσει τα ανεπιθύμητα μηνύματα και να τα διαχωρίσει από τα θεμιτά email.

3. Ιατρική Διάγνωση

Το Naive Bayes χρησιμοποιείται σε συστήματα ιατρικής διάγνωσης, όπου ο στόχος είναι η ταξινόμηση των ασθενών σε διαφορετικές κατηγορίες ασθενειών με βάση τα συμπτώματά τους, το ιατρικό ιστορικό και τα αποτελέσματα των εξετάσεων. Ο Naive Bayes μπορεί να βοηθήσει στην πραγματοποίηση προκαταρκτικών διαγνώσεων και στην πρόταση κατάλληλων θεραπευτικών επιλογών.

4. Ανίχνευση Απάτης

Το Naive Bayes μπορεί να χρησιμοποιηθεί σε συστήματα ανίχνευσης απάτης, όπου ο στόχος είναι να εντοπιστούν δυνητικά δόλιες συναλλαγές ή δραστηριότητες. Αναλύοντας μοτίβα συναλλαγών, συμπεριφορά χρήστη και άλλα σχετικά δεδομένα, ο Naive Bayes μπορεί να επισημάνει ύποπτες περιπτώσεις για περαιτέρω διερεύνηση.



## 2. OpenCV και Tensorflow Lite

---

Η OpenCV και η TensorFlow Lite είναι δύο δημοφιλείς βιβλιοθήκες που χρησιμοποιούνται στον τομέα της όρασης υπολογιστών και της μηχανικής μάθησης. Ενώ το OpenCV παρέχει ένα ολοκληρωμένο σύνολο αλγορίθμων και λειτουργιών όρασης υπολογιστή, το TensorFlow Lite εστιάζει στην ανάπτυξη μοντέλων μηχανικής μάθησης σε συσκευές με περιορισμένους πόρους.

### 2.1 OpenCV

Το OpenCV [3], συντομογραφία του Open Source Computer Vision Library, είναι μια βιβλιοθήκη λογισμικού ανοιχτού κώδικα που παρέχει μια πλούσια συλλογή λειτουργιών και αλγορίθμων για εργασίες όρασης υπολογιστή. Προσφέρει ένα ευρύ φάσμα λειτουργιών, όπως I/O εικόνας και βίντεο, επεξεργασία εικόνας, ανίχνευση χαρακτηριστικών, αναγνώριση αντικειμένων, βαθμονόμηση κάμερας και ενσωμάτωση μηχανικής μάθησης. Το OpenCV είναι γραμμένο σε C++ και υποστηρίζει διεπαφές για πολλές γλώσσες προγραμματισμού, καθιστώντας το προσβάσιμο και ευέλικτο για προγραμματιστές. Χρησιμοποιείται ευρέως σε διάφορους τομείς όπως η ρομποτική, η επαυξημένη πραγματικότητα, η ιατρική απεικόνιση και η παρακολούθηση.

#### 2.1.1 Αρχιτεκτονική και Σχεδίαση

Το OpenCV (Open Source Computer Vision Library) είναι μια βιβλιοθήκη λογισμικού ανοιχτού κώδικα που χρησιμοποιείται ευρέως για εργασίες όρασης υπολογιστή και επεξεργασίας εικόνας. Η αρχιτεκτονική και ο σχεδιασμός του διαδραματίζουν κρίσιμο ρόλο στην παροχή ενός εκτεταμένου συνόλου λειτουργιών και στην αποτελεσματική επεξεργασία των οπτικών δεδομένων. Σε αυτήν την περιγραφή, θα εμβαθύνουμε στην αρχιτεκτονική και τις αρχές σχεδιασμού του OpenCV, εξηγώντας την υποκείμενη δομή, τις ενότητες και τα βασικά στοιχεία του.

- **Modular Design:** Το OpenCV έχει μια αρθρωτή (modular) προσέγγιση στο σχεδιασμό, η οποία προωθεί την ευελιξία, την επεκτασιμότητα και την απλή ενσωμάτωση νέων χαρακτηριστικών. Η συλλογή χωρίζεται σε έναν αριθμό ενότητων, καθεμία από τις οποίες εστιάζει σε διαφορετικό τομέα όρασης υπολογιστή ή επεξεργασίας εικόνας. Το OpenCV περιλαμβάνει μια σειρά από σημαντικές ενότητες, μερικές από τις οποίες είναι core, imgproc, video, features2d, objdetect και ml. Οι προγραμματιστές μπορούν να εντοπίσουν και να χρησιμοποιήσουν την απαραίτητη λειτουργικότητα πιο εύκολα, επειδή κάθε ενότητα περιέχει σχετικούς αλγόριθμους και λειτουργίες.
- **Core Module:** Η βασική ενότητα (core module) του OpenCV, η οποία προσφέρει θεμελιώδεις δομές δεδομένων και βασικές λειτουργίες, είναι το κεντρικό εξάρτημα (component) του. Περιέχει διάφορους τύπους δεδομένων για την αναπαράσταση εικόνων, σημείων, ορθογώνιων και τιμών χρώματος, συμπεριλαμβανομένων των Mat (matrix), Point, Rect και Scalar. Η βασική μονάδα προσφέρει επίσης κρίσιμες

λειτουργίες, όπως διαχείριση μνήμης, picture I/O και βασικές αριθμητικές πράξεις σε πίνακες και εικόνες.

- **Image Processing Module (imgproc):** Η μονάδα (module) imgproc είναι αφιερωμένη σε διάφορες λειτουργίες επεξεργασίας εικόνας. Περιλαμβάνει συναρτήσεις για φιλτράρισμα εικόνας, thresholding(τρόπος διαχωρισμού μιας εικόνας σε foreground και background), μορφολογικές πράξεις(ένα ευρύ σύνολο λειτουργιών επεξεργασίας εικόνας που επεξεργάζονται ψηφιακές εικόνες με βάση τα σχήματά τους.), γεωμετρικούς μετασχηματισμούς και μετατροπές χρωματικού χώρου. Αυτή η ενότητα διαδραματίζει κρίσιμο ρόλο στη βελτίωση, τον χειρισμό και τη μετατροπή εικόνων.
- **Video Processing Module (video):** Οι εργασίες που σχετίζονται με το βίντεο, όπως το video I/O, η ανάλυση βίντεο και η σταθεροποίηση βίντεο, είναι τα κύρια θέματα της ενότητας βίντεο(video module). Προσφέρει δυνατότητες για πρόσβαση σε ροές βίντεο κάμερας(camera video streams), ανάγνωση και εγγραφή αρχείων βίντεο και εξαγωγή καρτέ(frames) βίντεο. Αλγόριθμοι για ανίχνευση κίνησης, εκτίμηση οπτικής ροής και αφαίρεση φόντου περιλαμβάνονται επίσης στη μονάδα(module) βίντεο.
- **Feature Detection and Description Module (features2d):** Η ενότητα (module) features2d περιλαμβάνει αλγόριθμους για τον εντοπισμό και την περιγραφή χαρακτηριστικών εικόνας. Περιλαμβάνει λειτουργίες ανίχνευσης γωνιών φυσαλίδων(blobs), ακμών και άλλων χαρακτηριστικών σημείων κλειδιού(χρησιμοποιείται για τον προσδιορισμό βασικών περιοχών ενός αντικειμένου που χρησιμοποιούνται ως βάση για να ταιριάξουν αργότερα και να το αναγνωρίσουν σε μια νέα εικόνα). Αυτή η ενότητα παρέχει επίσης αλγόριθμους για περιγραφικούς υπολογιστές (π.χ. SIFT, SURF, ORB) που αντιπροσωπεύουν την τοπική εμφάνιση των χαρακτηριστικών. Η ανίχνευση και η περιγραφή χαρακτηριστικών είναι απαραίτητες σε εργασίες όπως η αναγνώριση αντικειμένων, η αντιστοίχιση εικόνας και η καταχώριση εικόνας.
- **Object Detection Module (objdetect):** Η μονάδα (module) objdetect είναι αφιερωμένη σε εργασίες ανίχνευσης και αναγνώρισης αντικειμένων. Περιλαμβάνει προεκπαιδευμένα μοντέλα και αλγόριθμους όπως Haar cascades και HOG (Histogram of Oriented Gradients) [19]. Αυτοί οι αλγόριθμοι επιτρέπουν την ανίχνευση αντικειμένων ενδιαφέροντος(μπορεί να ανιχνεύσει πολλαπλές εμφανίσεις διαφορετικών τύπων αντικειμένων στην ίδια εικόνα και να παρέχει συντεταγμένες στην εικόνα όπου βρίσκονται αυτά τα αντικείμενα) σε εικόνες ή καρτέ βίντεο. Η ενότητα objdetect υποστηρίζει επίσης την εκπαίδευση προσαρμοσμένων μοντέλων ανίχνευσης αντικειμένων χρησιμοποιώντας αλγόριθμους μηχανικής μάθησης.
- **Machine Learning Module (ml):** Η μονάδα (module) ml παρέχει λειτουργίες μηχανικής μάθησης εντός του OpenCV. Περιλαμβάνει αλγόριθμους κατηγοριοποίησης, παλινδρόμησης, ομαδοποίησης και μείωσης διαστάσεων(τεχνική μείωσης του αριθμού των τυχαίων μεταβλητών σε ένα πρόβλημα με τη λήψη ενός συνόλου κύριων μεταβλητών.). Αυτή η ενότητα υποστηρίζει διάφορες τεχνικές μηχανικής μάθησης, όπως δέντρα αποφάσεων, μηχανές διανυσμάτων υποστήριξης (SVM), k-πλησιέστερους γείτονες (KNN) και νευρωνικά δίκτυα.

- Επιτρέπει στους προγραμματιστές να εκπαιδεύουν και να εφαρμόζουν μοντέλα μηχανικής μάθησης σε εικόνες και άλλα δεδομένα.
- **HighGUI and GUI Tools:** Το OpenCV περιλαμβάνει μια μονάδα (module) HighGUI που παρέχει λειτουργίες γραφικού περιβάλλοντος χρήστη (GUI) για οπτικοποίηση εικόνων, βίντεο και γραφικών στοιχείων. Επιτρέπει στους προγραμματιστές να δημιουργούν παράθυρα, να εμφανίζουν εικόνες, να χειρίζονται συμβάντα με το ποντίκι και το πληκτρολόγιο και να δημιουργούν διαδραστικές εφαρμογές. Το HighGUI επιτρέπει την απεικόνιση αλγορίθμων επεξεργασίας εικόνας σε πραγματικό χρόνο και διευκολύνει την αλληλεπίδραση με τον χρήστη.
  - **Integration with External Libraries:** Το OpenCV ενσωματώνεται απρόσκοπτα με εξωτερικές βιβλιοθήκες, επεκτείνοντας τις δυνατότητές του και καθιστώντας το συμβατό με διάφορα εργαλεία και frameworks. Για παράδειγμα, το OpenCV μπορεί να λειτουργήσει σε συνδυασμό με βιβλιοθήκες όπως οι NumPy και SciPy για προηγμένες αριθμητικές λειτουργίες και επιστημονικούς υπολογισμούς. Επιπλέον, το OpenCV ενσωματώνεται με frameworks βαθιάς μάθησης όπως το TensorFlow και το PyTorch, επιτρέποντας τη χρήση προεκπαιδευμένων μοντέλων βαθιάς μάθησης για εργασίες όρασης υπολογιστή.
  - **Cross-platform Support:** Το OpenCV έχει σχεδιαστεί για να είναι cross-platform, υποστηρίζοντας διαφορετικά λειτουργικά συστήματα όπως Windows, macOS, Linux, Android και iOS. Παρέχει συνεπή API και λειτουργίες σε όλες τις πλατφόρμες, επιτρέποντας στους προγραμματιστές να γράφουν κώδικα που είναι φορητός και συμβατός με διαφορετικά περιβάλλοντα.
  - **Performance Optimization:** Η βελτιστοποίηση της απόδοσης αποτελεί σημαντικό παράγοντα στο σχεδιασμό του OpenCV. Η βιβλιοθήκη χρησιμοποιεί τεχνικές όπως multi-threading, SIMD (Single Instruction, Multiple Data) και GPU acceleration για τη βελτίωση της ταχύτητας και της αποτελεσματικότητας επεξεργασίας. Το OpenCV εκμεταλλεύεται τις δυνατότητες parallel computing και hardware acceleration για να αξιοποιήσει πλήρως τις δυνατότητες των σύγχρονων επεξεργαστών και GPU.
  - **Documentation and Community:** Το OpenCV παρέχει εκτενή documentation, συμπεριλαμβανομένων λεπτομερών αναφορών API, tutorials, guides και παραδειγμάτων. Η τεκμηρίωση καλύπτει τη χρήση διαφορετικών μονάδων(modules), συναρτήσεων και αλγορίθμων, επιτρέποντας στους προγραμματιστές να μάθουν και να χρησιμοποιούν αποτελεσματικά το OpenCV. Επιπλέον, το OpenCV έχει μια μεγάλη και ενεργή κοινότητα χρηστών και συντελεστών(contributors) που παρέχουν υποστήριξη, μοιράζονται γνώσεις και συμβάλλουν στην ανάπτυξη της βιβλιοθήκης.

Συμπερασματικά, η αρχιτεκτονική και ο σχεδιασμός του OpenCV ακολουθούν μια αρθρωτή (modular) προσέγγιση, επιτρέποντας ευελιξία, επεκτασιμότητα και ευκολία ενσωμάτωσης. Οι ενότητες του καλύπτουν διάφορες πτυχές της όρασης υπολογιστή και της επεξεργασίας εικόνας, συμπεριλαμβανομένης της βασικής λειτουργικότητας(core module), της επεξεργασίας εικόνας και βίντεο, ανίχνευσης χαρακτηριστικών, ανίχνευσης αντικειμένων, μηχανικής μάθησης, ανάπτυξης GUI και βελτιστοποίησης απόδοσης. Ο σχεδιασμός του OpenCV επιτρέπει τον αποτελεσματικό χειρισμό και την ανάλυση των οπτικών δεδομένων, καθιστώντας το μια ευέλικτη και ευρέως χρησιμοποιούμενη βιβλιοθήκη στον τομέα της όρασης υπολογιστών.



## 2.1.2 APIs και Χρήση

Το OpenCV (Open Source Computer Vision Library) παρέχει ένα πλούσιο σύνολο διεπαφών προγραμματισμού εφαρμογών (API) που επιτρέπουν στους προγραμματιστές να αλληλεπιδρούν με τη βιβλιοθήκη και να χρησιμοποιούν τις λειτουργίες της για εργασίες όρασης υπολογιστή και επεξεργασίας εικόνας. Σε αυτήν την περιγραφή, θα εξερευνήσουμε τα API και τη χρήση τους στο OpenCV, καλύπτοντας τα κύρια στοιχεία, τις δομές δεδομένων και τις λειτουργίες που είναι διαθέσιμες για τους προγραμματιστές.

- **Βασικές Δομές Δεδομένων:** Οι βασικές δομές δεδομένων στο OpenCV είναι απαραίτητες για την αναπαράσταση και το χειρισμό εικόνων και πινάκων. Η κεντρική δομή δεδομένων είναι η `cv::Mat`, η οποία είναι μια ευέλικτη κλάση matrix που μπορεί να αποθηκεύσει εικόνες, πολυδιάστατους πίνακες και άλλους τύπους δεδομένων. Οι προγραμματιστές μπορούν να δημιουργήσουν και να χειριστούν αντικείμενα `cv::Mat` για ανάγνωση, εγγραφή και επεξεργασία εικόνων αποτελεσματικά.
- **Image I/O:** Το OpenCV παρέχει λειτουργίες για ανάγνωση και εγγραφή εικόνων σε διάφορες μορφές. Η συνάρτηση `cv::imread()` επιτρέπει στους προγραμματιστές να φορτώνουν εικόνες από αρχεία ή ροές, υποστηρίζοντας κοινές μορφές εικόνας όπως JPEG, PNG, BMP και TIFF. Η συνάρτηση `cv::imwrite()` επιτρέπει την αποθήκευση εικόνων στο δίσκο σε διαφορετικές μορφές. Αυτές οι λειτουργίες παρέχουν επιλογές για τον έλεγχο των ιδιοτήτων της εικόνας, όπως η σειρά των καναλιών χρώματος, το βάθος bit και οι παράμετροι συμπίεσης.
- **Λειτουργίες Επεξεργασίας Εικόνας:** Το OpenCV προσφέρει ένα ευρύ φάσμα λειτουργιών επεξεργασίας εικόνας που μπορούν να εφαρμοστούν σε εικόνες και πίνακες. Αυτές οι συναρτήσεις περιλαμβάνουν λειτουργίες φιλτραρίσματος, όπως θόλωση, ευκρίνεια και αφαίρεση θορύβου, που εκτελούνται με συναρτήσεις όπως `cv::blur()`, `cv::GaussianBlur()` και `cv::medianBlur()`. Άλλες λειτουργίες περιλαμβάνουν thresholding, ανίχνευση ακμών, μορφολογικές πράξεις, μετατροπές χρωματικού χώρου και γεωμετρικούς μετασχηματισμούς. Οι προγραμματιστές μπορούν να επιλέξουν από μια ποικιλία λειτουργιών για να χειριστούν και να βελτιώσουν τις εικόνες με βάση τις συγκεκριμένες απαιτήσεις τους.
- **Ανίχνευση και Περιγραφή Χαρακτηριστικών:** Το OpenCV παρέχει API για τον εντοπισμό και την περιγραφή χαρακτηριστικών σε εικόνες. Η κλάση `cv::FeatureDetector` και οι παραγόμενες κλάσεις της, όπως οι `cv::FastFeatureDetector` και `cv::ORB`, επιτρέπουν στους προγραμματιστές να εντοπίζουν σημεία-κλειδιά ή σημεία ενδιαφέροντος σε εικόνες. Ομοίως, η κλάση `cv::DescriptorExtractor` και οι παραγόμενες κλάσεις της, όπως `cv::SIFT` και `cv::SURF`, επιτρέπουν τον υπολογισμό των περιγραφικών χαρακτηριστικών που αντιπροσωπεύουν την τοπική εμφάνιση των σημείων κλειδιών. Αυτές οι δυνατότητες είναι απαραίτητες για εργασίες όπως η αναγνώριση αντικειμένων, η αντιστοίχιση εικόνας και η καταχώριση εικόνας.
- **Object Detection:** Το OpenCV περιλαμβάνει API για ανίχνευση αντικειμένων, επιτρέποντας στους προγραμματιστές να ανιχνεύουν αντικείμενα ενδιαφέροντος σε εικόνες ή ροές βίντεο. Η βιβλιοθήκη παρέχει προεκπαιδευμένα μοντέλα και Haar cascades (ένας αλγόριθμος που μπορεί να ανιχνεύσει αντικείμενα σε εικόνες,

ανεξάρτητα από την κλίμακα τους στην εικόνα και τη θέση τους) που μπορούν να χρησιμοποιηθούν με λειτουργίες όπως το `cv::CascadeClassifier` για τον εντοπισμό προσώπων, ματιών και άλλων αντικειμένων. Το OpenCV υποστηρίζει επίσης πιο προηγμένες τεχνικές ανίχνευσης αντικειμένων, όπως το HOG (Histogram of Oriented Gradients), το οποίο μπορεί να χρησιμοποιηθεί με την κλάση `cv::HOGDescriptor`

- **Επεξεργασία Βίντεο:** Το OpenCV προσφέρει API για εργασίες επεξεργασίας βίντεο, συμπεριλαμβανομένων `video I/O`, ανάλυσης και χειρισμού. Οι προγραμματιστές μπορούν να χρησιμοποιούν συναρτήσεις όπως `cv::VideoCapture` για πρόσβαση σε ροές βίντεο από κάμερες ή αρχεία και `cv::VideoWriter` για αποθήκευση επεξεργασμένων βίντεο. Το OpenCV παρέχει συναρτήσεις για ανάγνωση και εγγραφή μεμονωμένων καρέ βίντεο, εκτέλεση ανάλυσης καρέ-καρέ και εξαγωγή πληροφοριών κίνησης, όπως οπτική ροή, χρησιμοποιώντας αλγόριθμους όπως `cv::calcOpticalFlowPyrLK()`.
- **Ενσωμάτωση Μηχανικής Μάθησης:** Το OpenCV ενσωματώνει τις δυνατότητες μηχανικής μάθησης μέσω της μονάδας(module) `cv::ml`. Οι προγραμματιστές μπορούν να χρησιμοποιήσουν διάφορους αλγόριθμους μηχανικής μάθησης, συμπεριλαμβανομένων δέντρων αποφάσεων, μηχανών υποστήριξης διανυσμάτων (SVM), k-πλησιέστερων γειτόνων (KNN) και νευρωνικών δικτύων, για εργασίες όπως ταξινόμηση, παλινδρόμηση, ομαδοποίηση και μείωση διαστάσεων. Το module παρέχει μαθήματα για εκπαίδευση και χρήση μοντέλων μηχανικής μάθησης σε εικόνες και αριθμητικά δεδομένα.
- **Συναρτήσεις Γραφικού Περιβάλλοντος Χρήστη (GUI):** Το OpenCV περιλαμβάνει συναρτήσεις για τη δημιουργία γραφικών διεπαφών χρήστη (GUI) για την οπτικοποίηση εικόνων και την αλληλεπίδραση με τους χρήστες. Η συνάρτηση `cv::imshow()` εμφανίζει εικόνες στα παράθυρα και οι προγραμματιστές μπορούν να δημιουργήσουν πολλά παράθυρα και να διαχειριστούν τις ιδιότητές τους χρησιμοποιώντας συναρτήσεις όπως `cv::namedWindow()` και `cv::destroyWindow()`. Το OpenCV παρέχει επίσης συναρτήσεις χειρισμού συμβάντων για την καταγραφή συμβάντων με το ποντίκι και το πληκτρολόγιο και την απόκριση στις αλληλεπιδράσεις των χρηστών με εικόνες ή στοιχεία GUI.
- **Performance Optimization:** Το OpenCV προσφέρει λειτουργίες και τεχνικές για βελτιστοποίηση απόδοσης. Οι προγραμματιστές μπορούν να χρησιμοποιήσουν τις δυνατότητες πολλαπλών νημάτων για να παραλληλίσουν τις εργασίες επεξεργασίας εικόνας και να βελτιώσουν την απόδοση σε επεξεργαστές πολλαπλών πυρήνων. Το OpenCV υποστηρίζει επίσης Single Instruction, Multiple Data (SIMD) για να αξιοποιήσει το hardware acceleration για βελτιστοποιημένο υπολογισμό. Επιπλέον, το OpenCV παρέχει συναρτήσεις GPU-accelerated που εκμεταλλεύονται τις Μονάδες Επεξεργασίας Γραφικών (GPUs) για την επιτάχυνση ορισμένων λειτουργιών.

Εν κατακλείδι, το OpenCV προσφέρει ένα ολοκληρωμένο σύνολο API που δίνουν την δυνατότητα στους προγραμματιστές να εκτελούν ένα ευρύ φάσμα εργασιών όρασης υπολογιστή και επεξεργασίας εικόνας. Από I/O εικόνας μέχρι ανίχνευση χαρακτηριστικών, ανίχνευση αντικειμένων, επεξεργασία βίντεο και ενσωμάτωση μηχανικής μάθησης, το OpenCV παρέχει μια ευέλικτη εργαλειοθήκη για εργασία με οπτικά δεδομένα. Το εκτεταμένο

documentation και η ενεργή κοινότητα της βιβλιοθήκης ενισχύουν περαιτέρω τη χρηστικότητα και την υποστήριξή της για προγραμματιστές.

### 2.1.3 Εφαρμογές στο Πραγματικό Κόσμο

Το OpenCV (Open Source Computer Vision Library) έχει βρει πολλές εφαρμογές σε διάφορους κλάδους και τομείς. Το εκτεταμένο σύνολο λειτουργιών όρασης υπολογιστή και επεξεργασίας εικόνας το καθιστά ένα ευέλικτο εργαλείο για την επίλυση σύνθετων οπτικών προβλημάτων. Ακολουθούν ορισμένες αξιόλογες εφαρμογές του OpenCV στον πραγματικό κόσμο:

1. **Αναγνώριση και Παρακολούθηση Αντικειμένων**  
Το OpenCV χρησιμοποιείται ευρέως για την αναγνώριση και παρακολούθηση αντικειμένων σε εφαρμογές όπως συστήματα επιτήρησης, ρομποτική και επαυξημένη πραγματικότητα. Επιτρέπει τον εντοπισμό και την παρακολούθηση αντικειμένων ενδιαφέροντος σε ροές βίντεο σε πραγματικό χρόνο ή στατικές εικόνες. Αυτή η λειτουργία είναι απαραίτητη σε συστήματα ασφαλείας, αυτόνομα οχήματα και διαδραστικά παιχνίδια.
2. **Αναγνώριση Προσώπου και Βιομετρία:**  
Οι δυνατότητες αναγνώρισης προσώπου του OpenCV έχουν βρει εφαρμογές σε βιομετρικά συστήματα, επαλήθευση ταυτότητας και έλεγχο πρόσβασης. Επιτρέπει τον εντοπισμό και την ανάλυση των χαρακτηριστικών του προσώπου, την αναγνώριση ατόμων και την αντιστοίχιση προσώπων με βάσεις δεδομένων. Η αναγνώριση προσώπου που υποστηρίζεται από το OpenCV χρησιμοποιείται σε υπηρεσίες επιβολής του νόμου, συστήματα ελέγχου ταυτότητας και εξατομικευμένες εμπειρίες χρηστών (personalized user experiences).
3. **Ιατρική Ανάλυση Εικόνας:**  
Το OpenCV χρησιμοποιείται σε εφαρμογές ιατρικής απεικόνισης για εργασίες όπως τμηματοποίηση εικόνας, ανίχνευση όγκου και διάγνωση. Βοηθά στην εξαγωγή σημαντικών πληροφοριών από ιατρικές εικόνες, όπως ακτινογραφίες, μαγνητικές τομογραφίες και αξονικές τομογραφίες, διευκολύνοντας την ανίχνευση ασθενειών, την καταγραφή εικόνων και την ποσοτική ανάλυση. Το OpenCV διαδραματίζει σημαντικό ρόλο στην ιατρική έρευνα, τη διάγνωση και τον σχεδιασμό θεραπειών.
4. **Αυτόνομα Οχήματα και Ρομποτική:**  
Το OpenCV χρησιμοποιείται σε αυτόνομα οχήματα και ρομποτική για διάφορες εργασίες που βασίζονται στην όραση. Βοηθά στον εντοπισμό και την παρακολούθηση λωρίδας, τον εντοπισμό πεζών, την αναγνώριση σημάτων κυκλοφορίας και την αποφυγή εμποδίων. Οι αλγόριθμοι υπολογιστικής όρασης του OpenCV συμβάλλουν στην αντίληψη και στις δυνατότητες λήψης αποφάσεων των αυτόνομων οχημάτων και των ρομπότ, επιτρέποντάς τους να πλοηγούνται και να αλληλεπιδρούν με το περιβάλλον.
5. **Επαυξημένη Πραγματικότητα (AR) και Εικονική Πραγματικότητα (VR):**

Το OpenCV χρησιμοποιείται σε εφαρμογές AR και VR για την επικάλυψη εικονικών αντικειμένων σε σκηνές πραγματικού κόσμου. Διευκολύνει την παρακολούθηση marker(μπορούν να ανιχνευθούν μόνο από μια κάμερα)/markerless (μπορούν να ανιχνευθούν από μια κάμερα και ένα σύστημα όρασης υπολογιστή), την αναγνώριση χειρονομιών και την ανακατασκευή αντικειμένων 3D. Οι δυνατότητες επεξεργασίας εικόνας σε πραγματικό χρόνο του OpenCV επιτρέπουν καθηλωτικές εμπειρίες AR και ρεαλιστικά εικονικά περιβάλλοντα.

6. Βιομηχανικός Αυτοματισμός και Έλεγχος Ποιότητας:

Το OpenCV διαδραματίζει σημαντικό ρόλο στις διαδικασίες βιομηχανικού αυτοματισμού και ποιοτικού ελέγχου. Χρησιμοποιείται για εργασίες όπως ανίχνευση ελαττωμάτων, αναγνώριση αντικειμένων, barcode και επιθεώρηση προϊόντος. Το OpenCV βοηθά στην αυτοματοποίηση των διαδικασιών παραγωγής, στη διασφάλιση της ποιότητας των προϊόντων και στη βελτίωση της αποτελεσματικότητας σε κλάδους όπως η αυτοκινητοβιομηχανία, τα ηλεκτρονικά και τα φαρμακευτικά προϊόντα.

7. Sports Analytics and Tracking:

Το OpenCV χρησιμοποιείται στα Sports Analytics για την παρακολούθηση των κινήσεων των παικτών, την ανάλυση του παιχνιδιού και την εξαγωγή στατιστικών. Επιτρέπει την ανίχνευση παικτών, την παρακολούθηση της μπάλας και την πρόβλεψη τροχιάς. Τα συστήματα που βασίζονται στο OpenCV παρέχουν πολύτιμες πληροφορίες για την προπόνηση, την εκπαίδευση και τη μετάδοση σε αθλήματα όπως το ποδόσφαιρο, το μπάσκετ και το τένις.

8. Επεξεργασία Εικόνας και Βίντεο:

Οι δυνατότητες επεξεργασίας εικόνας του OpenCV αξιοποιούνται στο λογισμικό επεξεργασίας εικόνας και βίντεο. Επιτρέπει εργασίες όπως βελτίωση εικόνας, φιλτράρισμα, μορφοποίηση και σταθεροποίηση βίντεο. Τα εργαλεία που βασίζονται στο OpenCV προσφέρουν ένα ευρύ φάσμα δημιουργικών επιλογών για φωτογράφους, βιντεογράφους και καλλιτέχνες οπτικών εφέ.

Αυτά είναι μόνο μερικά παραδείγματα των ευρέος φάσματος εφαρμογών του OpenCV. Η ευελιξία και η εκτεταμένη λειτουργικότητά του το καθιστούν δημοφιλή επιλογή σε έργα όρασης υπολογιστών και επεξεργασίας εικόνας σε διάφορους κλάδους και τομείς.

## 2.2 TensorFlow Lite

Το TensorFlow Lite [4] είναι μια ελαφριά έκδοση της βιβλιοθήκης TensorFlow ειδικά σχεδιασμένη για την ανάπτυξη μοντέλων μηχανικής μάθησης σε συσκευές με περιορισμούς πόρων. Είναι βελτιστοποιημένο για απόδοση και αποδοτικότητα σε κινητές συσκευές, ενσωματωμένες και IoT συσκευές. Το TensorFlow Lite παρέχει εργαλεία για βελτιστοποίηση μοντέλων, συμπεριλαμβανομένων τεχνικών όπως η κβαντοποίηση (όρος που καλύπτει πολλές διαφορετικές τεχνικές για τη μετατροπή των τιμών εισόδου από ένα μεγάλο σύνολο σε τιμές εξόδου σε ένα μικρότερο σύνολο) και η συμπίεση, για τη μείωση του μεγέθους του μοντέλου και τη βελτίωση της ταχύτητας συμπερασμάτων. Υποστηρίζει πολλαπλές πλατφόρμες και επιτρέπει στους προγραμματιστές να εκτελούν μοντέλα μηχανικής μάθησης τοπικά σε συσκευές χωρίς να απαιτείται συνεχής σύνδεση στο διαδίκτυο. Το TensorFlow Lite χρησιμοποιείται συνήθως σε κινητές εφαρμογές, συσκευές IoT και ενσωματωμένα συστήματα για εργασίες όπως η αναγνώριση εικόνας, η επεξεργασία φυσικής γλώσσας και η προγνωστική συντήρηση [20] (σου επιτρέπει να εκτελείς αυτοματοποιημένη επεξεργασία δεδομένων σε ένα παράδειγμα συνόλου δεδομένων ή στο δικό σου σύνολο δεδομένων).

### 2.2.1 Αρχιτεκτονική και Σχεδίαση

Το TensorFlow Lite είναι ένα ελαφρύ και αποτελεσματικό πλαίσιο βαθιάς μάθησης (Deep Learning) ειδικά σχεδιασμένο για κινητές συσκευές και ενσωματωμένες συσκευές. Επιτρέπει στους προγραμματιστές να αναπτύσσουν και να εκτελούν εκπαιδευμένα μοντέλα μηχανικής μάθησης σε πλατφόρμες με περιορισμένους πόρους, επιτρέποντας την εξαγωγή συμπερασμάτων στη συσκευή χωρίς να βασίζονται σε διακομιστές που βασίζονται σε σύννεφο (cloud). Σε αυτήν την περιγραφή, θα διερευνήσουμε τις αρχές αρχιτεκτονικής και σχεδίασης του TensorFlow Lite, επισημαίνοντας τα βασικά του στοιχεία και πώς διευκολύνει την αποτελεσματική ανάπτυξη μοντέλων βαθιάς μάθησης.

- **Model Conversion:** Η διαδικασία ανάπτυξης μοντέλων βαθιάς μάθησης με το TensorFlow Lite ξεκινά με τη μετατροπή μοντέλων. Το TensorFlow Lite υποστηρίζει τη μετατροπή μοντέλων TensorFlow που είναι αποθηκευμένα στην τυπική μορφή SavedModel ή στην παλαιότερη μορφή παγωμένου γραφήματος [21] (χρησιμοποιούνται συνήθως για εξαγωγή συμπερασμάτων στο TensorFlow και αποτελούν σκαλοπάτι για εξαγωγή συμπερασμάτων για άλλα frameworks). Η διαδικασία μετατροπής περιλαμβάνει βελτιστοποιήσεις όπως η κβαντοποίηση, η οποία μειώνει την ακρίβεια των βαρών και τις ενεργοποιήσεις για την ελαχιστοποίηση του μεγέθους του μοντέλου και τη βελτίωση της απόδοσης σε συσκευή χαμηλής κατανάλωσης.
- **FlatBuffer Format:** Το TensorFlow Lite χρησιμοποιεί το format FlatBuffer για να αναπαραστήσει τα μοντέλα που έχουν μετατραπεί. Τα FlatBuffer είναι μια δυαδική μορφή σειριοποίησης που επιτρέπει την αποτελεσματική αποθήκευση και ανάκτηση δεδομένων με ελάχιστο αποτύπωμα μνήμης. Η χρήση των FlatBuffers συμβάλλει στη μείωση της επιβάρυνσης της μνήμης και του χρόνου φόρτωσης, κάτι που είναι ζωτικής σημασίας για συσκευές με περιορισμένους πόρους.

- **Interpreter:** Στον πυρήνα του TensorFlow Lite βρίσκεται ο διερμηνέας, ο οποίος παρέχει το περιβάλλον χρόνου εκτέλεσης (runtime environment) για την εκτέλεση των μοντέλων που έχουν μετατραπεί στη συσκευή προορισμού. Ο διερμηνέας είναι υπεύθυνος για τη φόρτωση του μοντέλου από το FlatBuffer format, τη διαχείριση της κατάστασης του μοντέλου και την εκτέλεση συμπερασμάτων. Βελτιστοποιεί την εκτέλεση λειτουργιών αξιοποιώντας hardware acceleration, όπως οδηγίες SIMD (Single Instruction, Multiple Data) και GPU acceleration.
- **Tensor Memory Allocation:** Το TensorFlow Lite χρησιμοποιεί έναν εκχωρητή μνήμης (allocator) ειδικά σχεδιασμένο για αποτελεσματική διαχείριση tensor. Οι tensors, οι οποίοι είναι πολυδιάστατοι πίνακες που αντιπροσωπεύουν δεδομένα σε μοντέλα βαθιάς μάθησης, κατανέμονται σε buffer μνήμης που επαναχρησιμοποιούνται κατά τη διάρκεια της εξαγωγής συμπερασμάτων για την ελαχιστοποίηση του κατακερματισμού της μνήμης. Ο εκχωρητής μνήμης βελτιστοποιεί τη χρήση της μνήμης και μειώνει τα έξοδα με την επαναχρησιμοποίηση των buffer μνήμης καθώς οι tensors εκχωρούνται(allocated) και εκχωρούνται(deallocated).
- **Operator Fusion:** Για τη βελτίωση της απόδοσης συμπερασμάτων, το TensorFlow Lite εφαρμόζει τεχνικές operator fusion. Το Operator fusion συνδυάζει πολλαπλές διαδοχικές λειτουργίες σε μια ενιαία βελτιστοποιημένη λειτουργία, μειώνοντας τις μεταφορές μνήμης και τα ενδιάμεσα buffers. Το Fusion εξαλείφει τον περιττό υπολογισμό και την επιβάρυνση της μνήμης, με αποτέλεσμα την ταχύτερη εκτέλεση συμπερασμάτων σε συσκευές με περιορισμούς πόρων.
- **Hardware Acceleration:** Το TensorFlow Lite αξιοποιεί την επιτάχυνση υλικού για να μεγιστοποιήσει την απόδοση της εκτέλεσης του μοντέλου. Υποστηρίζει βιβλιοθήκες και API επιτάχυνσης υλικού, όπως το Android Neural Networks API (NNAPI) και το GPU Compute API, για τη μεταφόρτωση υπολογισμών σε εξειδικευμένο υλικό. Αυτό επιτρέπει την αποτελεσματική χρήση των πόρων υλικού, παρέχοντας ταχύτερα συμπεράσματα και μειωμένη κατανάλωση ενέργειας.
- **Platform-specific Implementations:** Το TensorFlow Lite παρέχει υλοποιήσεις για συγκεκριμένα πλατφόρμα για να διασφαλίσει τη βέλτιστη απόδοση σε διαφορετικές αρχιτεκτονικές συσκευών. Αυτές οι υλοποιήσεις επωφελούνται από χαρακτηριστικά ειδικά για τη συσκευή, όπως οδηγίες διανυσματοποίησης(ένα ειδικό σύνολο εγγενών συναρτήσεων που εκμεταλλεύονται τον παραλληλισμό δεδομένων) και επιταχυντές υλικού, για να επιταχύνουν την εκτέλεση του μοντέλου. Το TensorFlow Lite υποστηρίζει διάφορες πλατφόρμες, όπως Android, iOS, Raspberry Pi και μικροελεγκτές.
- **Flexibility and Extensibility:** Το TensorFlow Lite προσφέρει έναν ευέλικτο και επεκτάσιμο σχεδιασμό για να φιλοξενήσει ένα ευρύ φάσμα περιπτώσεων χρήσης. Παρέχει API για ενοποίηση με γλώσσες προγραμματισμού, όπως C++, Java και Python, επιτρέποντας στους προγραμματιστές να ενσωματώνουν απρόσκοπτα το TensorFlow Lite στις εφαρμογές τους. Το TensorFlow Lite υποστηρίζει επίσης προσαρμοσμένους χειριστές, επιτρέποντας την υλοποίηση εξειδικευμένων λειτουργιών και την ενσωμάτωση προσαρμοσμένων επιταχυντών υλικού(hardware accelerators).

- **Model Optimization Toolkit:** Για να βελτιώσει περαιτέρω την αποτελεσματικότητα και την απόδοση των μοντέλων TensorFlow Lite, το TensorFlow παρέχει ένα Kit εργαλείων βελτιστοποίησης μοντέλου. Αυτή η εργαλειοθήκη περιλαμβάνει εργαλεία για εκπαίδευση με επίγνωση κβαντισμού, κβαντοποίηση μετά την εκπαίδευση και τεχνικές βελτιστοποίησης μοντέλων, όπως το κλάδεμα και η αραιότητα (ένας τύπος δεδομένων που δεν περιέχει τις πραγματικές τιμές των χαρακτηριστικών). Αυτές οι τεχνικές βελτιστοποίησης συμβάλλουν στη μείωση του μεγέθους του μοντέλου, στη βελτίωση της ταχύτητας συμπερασμάτων και στην ελαχιστοποίηση της χρήσης μνήμης.

Συμπερασματικά, το TensorFlow Lite είναι ένα ισχυρό framework που επιτρέπει την αποτελεσματική ανάπτυξη μοντέλων βαθιάς μάθησης σε συσκευές με περιορισμένους πόρους. Οι αρχές της αρχιτεκτονικής και του σχεδιασμού του επικεντρώνονται στη μετατροπή μοντέλων, τη διαχείριση μνήμης, τη συγχώνευση χειριστή, την επιτάχυνση υλικού, τις εφαρμογές για συγκεκριμένες πλατφόρμες, την ευελιξία και την επεκτασιμότητα. Αξιοποιώντας το TensorFlow Lite, οι προγραμματιστές μπορούν να φέρουν τη δύναμη της βαθιάς μάθησης σε κινητές συσκευές και ενσωματωμένες συσκευές, ανοίγοντας νέες δυνατότητες για εξαγωγή συμπερασμάτων στη συσκευή και στις έξυπνες εφαρμογές.

### 2.2.2 APIs και Χρήση

Το TensorFlow Lite παρέχει ένα σύνολο API και εργαλείων που επιτρέπουν στους προγραμματιστές να ενσωματώνουν και να αναπτύσσουν μοντέλα βαθιάς μάθησης σε συσκευές με περιορισμένους πόρους. Αυτά τα API προσφέρουν διάφορες λειτουργίες για τη φόρτωση μοντέλων, την εξαγωγή συμπερασμάτων, τον χειρισμό εισόδων/εξόδων και τη βελτιστοποίηση της απόδοσης. Σε αυτήν την περιγραφή, θα γίνει αναφορά στα API TensorFlow Lite και τις χρήσεις τους, μαζί με πρακτικά παραδείγματα και εφαρμογές οι προγραμματιστές μπορούν να φέρουν τη δύναμη της βαθιάς μάθησης σε κινητές συσκευές και ενσωματωμένες συσκευές, ανοίγοντας νέες δυνατότητες για εξαγωγή συμπερασμάτων στη συσκευή και στις έξυπνες εφαρμογές.

- **Model Loading and Initialization:** Το TensorFlow Lite API περιλαμβάνει συναρτήσεις για τη φόρτωση και την προετοιμασία των μοντέλων που έχουν μετατραπεί. Η κλάση Interpreter είναι το κεντρικό στοιχείο για τη φόρτωση του μοντέλου και το συμπέρασμα. Επιτρέπει στους προγραμματιστές να φορτώσουν ένα μοντέλο από ένα αρχείο FlatBuffer, να ορίσουν tensors εισόδου και εξόδου και να διαμορφώσουν επιλογές διεργασίας.
- **Input and Output Handling:** Το TensorFlow Lite παρέχει API για τον ορισμό δεδομένων εισόδου και την ανάκτηση δεδομένων εξόδου από το μοντέλο. Αυτά τα API επιτρέπουν στους προγραμματιστές να τροφοδοτούν δεδομένα εισόδου στο μοντέλο, να εκτελούν συμπεράσματα και να ανακτούν τις προβλέψεις του μοντέλου.
- **Quantization and Dequantization:** Το TensorFlow Lite υποστηρίζει τεχνικές κβαντοποίησης για τη μείωση του μεγέθους του μοντέλου και τη βελτίωση της

ταχύτητας συμπερασμάτων σε συσκευές χαμηλής κατανάλωσης. Το API παρέχει λειτουργίες για την κβαντοποίηση και την αποκβαντοποίηση tensors.

- **Performance Optimization:** Το TensorFlow Lite προσφέρει API για βελτιστοποίηση απόδοσης σε διαφορετικές πλατφόρμες. Για παράδειγμα, σε συσκευές Android, το AndroidNDDelegate API επιτρέπει στους προγραμματιστές να εκφορτώνουν υπολογισμούς στο Android Neural Networks API (NNAPI) για επιτάχυνση υλικού. Αυτό βελτιώνει την ταχύτητα συμπερασμάτων και μειώνει την κατανάλωση ενέργειας. Άλλες πλατφόρμες ενδέχεται να έχουν παρόμοιες βελτιστοποιήσεις ειδικά για τις δυνατότητες υλικού τους.
- **Multithreading and Asynchronous Inference:** Το TensorFlow Lite υποστηρίζει πολλαπλά νήματα (multithreading) και ασύγχρονα συμπεράσματα, επιτρέποντας στους προγραμματιστές να επωφεληθούν από την παράλληλη επεξεργασία και να βελτιώσουν την ανταπόκριση της εφαρμογής. Το API παρέχει λειτουργίες για τον έλεγχο του αριθμού των νημάτων που χρησιμοποιούνται για εξαγωγή συμπερασμάτων και για την ασύγχρονη εκτέλεση των συμπερασμάτων.
- **Custom Operators:** Το TensorFlow Lite επιτρέπει την εφαρμογή προσαρμοσμένων τελεστών (custom operators) για την επέκταση της λειτουργικότητάς του. Οι προσαρμοσμένοι τελεστές μπορούν να υλοποιηθούν σε C++ ή χρησιμοποιώντας το Operator API του TensorFlow, παρέχοντας ευελιξία για την ενσωμάτωση λειτουργιών σε συγκεκριμένο τομέα ή την αξιοποίηση προσαρμοσμένων επιταχυντών υλικού. Οι προσαρμοσμένοι χειριστές επιτρέπουν στους προγραμματιστές να βελτιστοποιούν την απόδοση για συγκεκριμένες εργασίες ή να χρησιμοποιούν ιδιόκτητους αλγόριθμους μηχανικής μάθησης.
- **Model Optimization Toolkit Integration:** Το TensorFlow Lite ενσωματώνεται με το TensorFlow Model Optimization Toolkit, παρέχοντας εργαλεία για συμπίεση, κλάδεμα και κβαντοποίηση μοντέλων. Αυτά τα εργαλεία επιτρέπουν στους προγραμματιστές να βελτιστοποιούν τα μοντέλα πριν από τη μετατροπή σε μορφή TensorFlow Lite, με αποτέλεσμα μικρότερα μεγέθη μοντέλων, ταχύτερη εξαγωγή συμπερασμάτων και μειωμένη χρήση μνήμης.

### 2.2.3 Εφαρμογές στο Πραγματικό Κόσμο

Το TensorFlow Lite, ως ένα ελαφρύ framework βαθιάς μάθησης που έχει σχεδιαστεί για κινητές συσκευές και ενσωματωμένες συσκευές, έχει βρει πολλές εφαρμογές πραγματικού κόσμου σε διάφορους τομείς. Ακολουθούν μερικά αξιοσημείωτα παραδείγματα:

#### 1. Mobile Vision Applications

Το TensorFlow Lite επιτρέπει την κατηγοριοποίηση εικόνων στη συσκευή, τον εντοπισμό αντικειμένων και την αναγνώριση προσώπου σε εφαρμογές για κινητές συσκευές. Ενισχύει εφαρμογές που βασίζονται σε κάμερα σε πραγματικό χρόνο που μπορούν να αναγνωρίσουν αντικείμενα, να αναγνωρίσουν πρόσωπα και να πραγματοποιήσουν αναζητήσεις βάσει εικόνων. Αυτές οι εφαρμογές περιλαμβάνουν φίλτρα επαυξημένης πραγματικότητας (AR), οπτική αναζήτηση, αναγνώριση εικόνας και παιχνίδια για κινητά.



## 2. Natural Language Processing

Το TensorFlow Lite ενεργοποιεί εφαρμογές επεξεργασίας φυσικής γλώσσας (NLP) στη συσκευή, όπως βοηθός φωνής, chatbot και εργαλεία ανάλυσης κειμένου. Με το TensorFlow Lite, αυτές οι εφαρμογές μπορούν να εκτελούν εργασίες όπως αναγνώριση ομιλίας, μετάφραση γλώσσας, ανάλυση συναισθήματος και ταξινόμηση κειμένου απευθείας στη συσκευή χωρίς να βασίζονται σε υπηρεσίες που βασίζονται στο cloud.

## 3. Health Monitoring and Biometrics

Το TensorFlow Lite χρησιμοποιείται σε εφαρμογές υγειονομικής περίθαλψης για την ανάλυση ιατρικών εικόνων, την ανίχνευση ασθενειών και την παρακολούθηση ζωτικών σημείων. Επιτρέπει την ανάπτυξη φορητών συσκευών για την παρακολούθηση του καρδιακού ρυθμού, της αρτηριακής πίεσης και άλλων βιομετρικών στοιχείων. Το TensorFlow Lite διευκολύνει επίσης την ανάλυση ιατρικών εικόνων, συμπεριλαμβανομένων ακτινογραφιών, μαγνητικής τομογραφίας και αξονικής τομογραφίας, βοηθώντας στη διάγνωση και τον σχεδιασμό της θεραπείας.

## 4. Autonomous Robotics and Drones

Το TensorFlow Lite επιτρέπει την ανίχνευση αντικειμένων σε πραγματικό χρόνο, την αναγνώριση εικόνας και την παρακολούθηση κίνησης σε αυτόνομα ρομπότ και drones. Επιτρέπει σε αυτές τις συσκευές να αντιλαμβάνονται και να κατανοούν το περιβάλλον τους, να πλοηγούνται σε εμπόδια και να εκτελούν πολύπλοκες εργασίες. Οι εφαρμογές περιλαμβάνουν αυτόνομη πλοήγηση, παρακολούθηση αντικειμένων και εναέριες επιθεωρήσεις.

## 5. Retail and E-commerce

Το TensorFlow Lite ενισχύει τα συστήματα οπτικής αναζήτησης και συστάσεων σε εφαρμογές λιανικής και ηλεκτρονικού εμπορίου. Με την ανάπτυξη μοντέλων απευθείας σε κινητές συσκευές, οι χρήστες μπορούν να τραβήξουν εικόνες προϊόντων και να λαμβάνουν άμεσα αποτελέσματα αναζήτησης, προτάσεις και πληροφορίες τιμολόγησης. Το TensorFlow Lite βελτιώνει την εμπειρία του χρήστη και επιτρέπει την απρόσκοπτη ενσωμάτωση της τεχνητής νοημοσύνης σε εφαρμογές αγορών.

## 6. Automotive Industry

Το TensorFlow Lite χρησιμοποιείται σε εφαρμογές αυτοκινήτου για εργασίες όπως η βοήθεια οδηγού, ο εντοπισμός αντικειμένων και η αναγνώριση χειρονομιών. Επιτρέπει στα προηγμένα συστήματα υποβοήθησης οδηγού (ADAS) να αναλύουν δεδομένα αισθητήρων σε πραγματικό χρόνο και να λαμβάνουν κρίσιμες αποφάσεις. Το TensorFlow Lite επιτρέπει επίσης συστήματα ψυχαγωγίας στο αυτοκίνητο με αναγνώριση φωνής και δυνατότητες κατανόησης φυσικής γλώσσας.

Αυτά είναι μερικά μόνο παραδείγματα από το ευρύ φάσμα εφαρμογών όπου χρησιμοποιείται το TensorFlow Lite. Ο ελαφρύς και αποτελεσματικός χαρακτήρας του, σε συνδυασμό με τις δυνατότητες εξαγωγής συμπερασμάτων στη συσκευή, το καθιστούν κατάλληλο για διάφορα σενάρια πραγματικού κόσμου, δίνοντας τη δυνατότητα στους προγραμματιστές να φέρνουν τα μοντέλα βαθιάς μάθησης σε πλεονεκτική θέση.



## 3. ML Kit

---

Το ML Kit [5, 6] είναι ένα framework μηχανικής μάθησης που αναπτύχθηκε από την Google για εφαρμογές Android. Παρέχει έτοιμα προς χρήση API και προεκπαιδευμένα μοντέλα που επιτρέπουν στους προγραμματιστές να ενσωματώνουν εύκολα τις δυνατότητες μηχανικής μάθησης στις εφαρμογές τους Android χωρίς να απαιτούν εκτεταμένη γνώση αλγορίθμων μηχανικής μάθησης ή συνόλων δεδομένων εκπαίδευσης. Το ML Kit προσφέρει ένα ευρύ φάσμα λειτουργιών, όπως αναγνώριση κειμένου, επισήμανση εικόνων, ανίχνευση προσώπου, σάρωση barcode και παρακολούθηση αντικειμένων. Με το ML Kit, οι προγραμματιστές μπορούν να αξιοποιήσουν τη δύναμη της μηχανικής μάθησης για να βελτιώσουν τις εφαρμογές τους με έξυπνες λειτουργίες, όπως μετάφραση σε πραγματικό χρόνο, έξυπνη αναγνώριση εικόνας και αυτοματοποιημένη εξαγωγή δεδομένων. Το ML Kit απλοποιεί την ενσωμάτωση της μηχανικής μάθησης σε εφαρμογές Android, καθιστώντας το προσβάσιμο σε προγραμματιστές όλων των επιπέδων δεξιοτήτων και δίνοντάς τους τη δυνατότητα να δημιουργήσουν πιο ελκυστικές και καινοτόμες εμπειρίες για τους χρήστες.

### 3.1 Αρχιτεκτονική και Σχεδίαση

Το ML Kit περιλαμβάνει μια προηγμένη αρχιτεκτονική και αποτελεί ένα σύστημα σχεδίασης για την εφαρμογή της μηχανικής μάθησης σε κινητές συσκευές. Η σχεδίαση του ML Kit έχει επιτρέψει την εύκολη ενσωμάτωση και χρήση των λειτουργιών μηχανικής μάθησης στις εφαρμογές Android και iOS. Η αρχιτεκτονική του έχει σχεδιαστεί με στόχο την απλότητα και την αποδοτικότητα. Έχει βασιστεί σε προηγμένες τεχνικές μηχανικής μάθησης, όπως η μεταφορά μάθησης και η εκπαίδευση με μεγάλα σύνολα δεδομένων, για να προσφέρει αξιόπιστες και αποτελεσματικές λειτουργίες.

Βασική αρχή του ML Kit αποτελεί η απλότητα της χρήσης. Η Google έχει δημιουργήσει προετοιμασμένες βιβλιοθήκες με ετοιμοπαράδοτες λειτουργίες μηχανικής μάθησης που μπορούν να ενσωματωθούν εύκολα στις εφαρμογές μας. Αυτές οι βιβλιοθήκες παρέχουν προηγμένες λειτουργίες, όπως ανίχνευση αντικειμένων, αναγνώριση προσώπου και αναγνώριση κειμένου, με απλό API για να τις χρησιμοποιήσουμε. Η σχεδίαση του επιτρέπει επίσης την ευελιξία και την προσαρμοστικότητα. Έχουμε την δυνατότητα να προσθέσουμε τα δικά μας μοντέλα μηχανικής μάθησης και να τα εκπαιδύσουμε για να ανταποκρίνονται στις ανάγκες μας. Αυτό μας επιτρέπει να δημιουργήσουμε προσαρμοσμένες λύσεις που ταιριάζουν ακριβώς στην εφαρμογή μας και μπορούν να παρέχουν ακόμη καλύτερες αποδόσεις.

Επιπλέον, όσον αφορά τη σχεδίαση, το ML Kit προσφέρει μια ολοκληρωμένη εμπειρία χρήστη με φιλικό προς τον χρήστη περιβάλλον ανάπτυξης (IDE). Υποστηρίζει τόσο το Android Studio όσο και το Xcode για την ανάπτυξη εφαρμογών Android και iOS αντίστοιχα. Μέσω του IDE, μπορούμε να προσθέσουμε το ML Kit στο έργο μας, να ενσωματώσουμε τις λειτουργίες του και να δοκιμάσουμε την απόδοσή τους. Μια ακόμη σημαντική πτυχή της αρχιτεκτονικής του ML Kit είναι η επεξεργασία δεδομένων σε πραγματικό χρόνο. Είναι σχεδιασμένο έτσι ώστε να μπορεί να ανταποκρίνεται γρήγορα στα δεδομένα που παρέχονται από την συσκευή,

επιτρέποντας την πραγματοποίηση λειτουργιών μηχανικής μάθησης σε πραγματικό χρόνο. Αυτό είναι ιδιαίτερα χρήσιμο για εφαρμογές που απαιτούν άμεση ανταπόκριση, όπως οι εφαρμογές πραγματικού χρόνου και οι εφαρμογές αυξημένης πραγματικότητας.

Παράλληλα, η σχεδίαση του ML Kit δίνει έμφαση στη φορητότητα και την αποτελεσματική χρήση των πόρων της συσκευής. Οι αλγόριθμοι μηχανικής μάθησης του έχουν βελτιστοποιηθεί για να λειτουργούν αποδοτικά σε περιβάλλοντα με περιορισμένους πόρους, όπως οι κινητές συσκευές. Αυτό εξασφαλίζει την ομαλή λειτουργία των εφαρμογών και την αποφυγή υπερβολικής χρήσης μπαταρίας ή υπερθέρμανση της συσκευής. Επιπρόσθετα, η σχεδίαση του ML Kit προσφέρει δυνατότητες για την προστασία της ιδιωτικότητας των χρηστών, παρέχοντας επιλογές για την επεξεργασία δεδομένων στην συσκευή χωρίς την ανάγκη αποστολής στοιχείων στους εξυπηρετητές της Google. Έτσι επιτυγχάνεται υψηλό επίπεδο απορρήτου και ασφάλειας για τους χρήστες των εφαρμογών.

Συνολικά, η αρχιτεκτονική και η σχεδίαση του ML Kit έχουν δημιουργηθεί με γνώμονα την ευκολία χρήσης, την απόδοση, τη φορητότητα και την ιδιωτικότητα. Αυτές οι σχεδιαστικές αρχές καθιστούν το ML Kit ένα ισχυρό εργαλείο για την εφαρμογή της μηχανικής μάθησης στις κινητές συσκευές, προσφέροντας ευκολία, απόδοση και αξιοπιστία στις εφαρμογές μας.

### 3.2 APIs και Χρήση

Τα APIs (Application Programming Interfaces) που σχετίζονται με το ML Kit προσφέρουν διεπαφές προγραμματισμού για την εύκολη ενσωμάτωση και χρήση των λειτουργιών μηχανικής μάθησης στις εφαρμογές σας. Εδώ θα αναλύσουμε τα APIs που παρέχονται από το ML Kit:

#### 1. Image Labeling API

Με τα API ετικετών εικόνων μπορούμε να εντοπίσουμε και να εξάγουμε πληροφορίες σχετικά με οντότητες σε μια εικόνα σε μια ευρεία ομάδα κατηγοριών. Το προεπιλεγμένο μοντέλο ετικετών εικόνων μπορεί να αναγνωρίσει γενικά αντικείμενα, μέρη, δραστηριότητες, είδη ζώων, προϊόντα και άλλα. Αποτελεί ένα ισχυρό βασικό ταξινομητή γενικής χρήσης, καθώς αναγνωρίζει περισσότερες από 400 κατηγορίες που περιγράφουν τα αντικείμενα που βρίσκονται πιο συχνά στις φωτογραφίες. Επιπλέον, δίνει την δυνατότητα προσαρμογής στη θήκη χρήσης μας με προσαρμοσμένα μοντέλα από το TensorFlow Hub ή το δικό μας προσαρμοσμένο μοντέλο που έχει εκπαιδευτεί με τον κατασκευαστή TensorFlow, AutoML Vision Edge ή TensorFlow Lite Model. Τέλος, δεν χρειάζεται να ασχοληθούμε με την είσοδο / έξοδο μοντέλων χαμηλού επιπέδου, την προ και μετά την επεξεργασία εικόνας ή την δημιουργία διοχέτευσης επεξεργασίας. Το ML Kit εξάγει τις ετικέτες από το μοντέλο TensorFlow Lite και τις παρέχει ως περιγραφή κειμένου.

#### 2. Text Recognition API

Το API ML Kit Text Recognition v2 μπορεί να αναγνωρίσει κείμενο σε διάφορα σενάρια και γλώσσες, υποστηρίζει την αναγνώριση κειμένου σε κινέζικα, ντεβαναγκάρι, ιαπωνικά, κορεατικά και λατινικά σενάρια. Αναλύει τη δομή του κειμένου μέσω του εντοπισμού συμβόλων, στοιχείων, γραμμών και παραγράφων. Ακόμα, προσδιορίζει τη γλώσσα του αναγνωρισμένου κειμένου, ενώ μπορεί να αναγνωρίσει κείμενο σε πραγματικό χρόνο σε ένα

ευρύ φάσμα συσκευών. Το API μπορεί επίσης να χρησιμοποιηθεί για την αυτοματοποίηση εργασιών εισαγωγής δεδομένων, όπως η επεξεργασία πιστωτικών καρτών, αποδείξεων και επαγγελματικών καρτών.

### 3. Face Detection API

Με το API ανίχνευσης προσώπου μπορούμε να ανιχνεύσουμε πρόσωπα σε μια εικόνα, να αναγνωρίσουμε βασικά χαρακτηριστικά του προσώπου και να λάβουμε τα περιγράμματα των προσώπων που εντοπίστηκαν. Μπορούμε να αναγνωρίσουμε τις εκφράσεις του προσώπου, να παρακολουθήσουμε πρόσωπα σε καρέ βίντεο και να το επεξεργαστούμε σε πραγματικό χρόνο. Με την ανίχνευση προσώπου μπορούμε, επιπλέον, να λάβουμε τις πληροφορίες που χρειαζόμαστε για να εκτελέσουμε εργασίες όπως να στολίσουμε selfie και πορτρέτα ή να δημιουργήσουμε avatar από τη φωτογραφία ενός χρήστη. Επειδή το ML Kit μπορεί να εκτελέσει ανίχνευση προσώπου σε πραγματικό χρόνο, μπορούμε να το χρησιμοποιήσουμε σε εφαρμογές όπως συνομιλία μέσω βίντεο ή παιχνίδια που ανταποκρίνονται στις εκφράσεις του παίκτη.

### 4. Barcode Scanning API

Με το API σάρωσης γραμμωτού κώδικα μπορούμε να διαβάσουμε δεδομένα που κωδικοποιούνται χρησιμοποιώντας τις περισσότερες τυπικές μορφές γραμμικού κώδικα. Η σάρωση του γραμμικού κώδικα πραγματοποιείται στη συσκευή και δεν απαιτεί σύνδεση δικτύου. Οι γραμμωτοί κώδικες είναι ένα βολικός τρόπος για να μεταβιβάσουμε πληροφορίες από τον πραγματικό κόσμο στην εφαρμογή μας. Ειδικότερα, όταν χρησιμοποιούμε μορφές 2D, όπως ο κώδικας QR, μπορούμε να κωδικοποιήσουμε δομημένα δεδομένα, όπως στοιχεία επικοινωνίας ή διαπιστευτήρια δικτύου WiFi. Επειδή το ML Kit μπορεί αυτόματα να αναγνωρίσει και να αναλύσει αυτά τα δεδομένα, η εφαρμογή μας μπορεί να ανταποκρίνεται έξθπνα όταν ένας χρήστης σαρώνει ένα γραμμωτό κώδικα.

### 5. Smart Reply API

Με το API έξυπνης απάντησης μπορούμε να δημιουργήσουμε αυτόματα σχετικές απαντήσεις σε μηνύματα. Η έξυπνη απάντηση βοηθά τους χρήστες να ανταποκρίνονται γρήγορα στα μηνύματα και διευκολύνει την απάντηση σε μηνύματα σε συσκευές με περιορισμένες δυνατότητες εισαγωγής. Το μοντέλο έξυπνης απάντησης δημιουργεί προτάσεις απάντησης με βάση το πλήρες πλαίσιο μιας συνομιλίας και όχι μόνο ένα μήνυμα. Το μοντέλο της συσκευής δημιουργεί απαντήσεις γρήγορα και δεν απαιτεί από εμάς να στείλουμε τα μηνύματα των χρηστών σε έναν απομακρυσμένο διακομιστή.

### 6. Language Identification API

Με το API αναγνώρισης γλώσσας στη συσκευή, μπορούμε να προσδιορίσουμε τη γλώσσα μιας συμβολοσειράς κειμένου. Η αναγνώριση γλώσσας μπορεί να είναι χρήσιμη όταν εργαζόμαστε με κείμενο που παρέχεται από τον χρήστη, το οποίο συχνά δεν συνοδεύεται από πληροφορίες γλώσσας. Προσδιορίζει πάνω από εκατό διαφορετικές γλώσσες ενώ ταυτόχρονα υποστηρίζει αραβικά, βουλγαρικά, ελληνικά, χίντι, ιαπωνικά, ρωσικά και κινέζικα κείμενα τόσο σε εγγενή όσο και σε ρομανική γραφή.

## 3.2 Εφαρμογές στον Πραγματικό Κόσμο

Οι δυνατότητες του ML Kit έχουν βρει εφαρμογές σε διάφορους τομείς λόγω της ευκολίας ενσωμάτωσης και της επεξεργασίας στη συσκευή. Ακολουθούν ορισμένες πραγματικές εφαρμογές όπου έχει χρησιμοποιηθεί το ML Kit:

1. Αναγνώριση Αντικειμένων:

Με το Image Labeling API του ML Kit, μπορούμε να δημιουργήσουμε εφαρμογές που αναγνωρίζουν αντικείμενα σε εικόνες. Αυτό μπορεί να χρησιμοποιηθεί για αναγνώριση προϊόντων, ανίχνευση αντικειμένων σε αυτόματα συστήματα, αναγνώριση συμβόλων και άλλα.

2. Αναγνώριση Προσώπων:

Με το Face Detection API του ML Kit, μπορούμε να αναγνωρίσουμε πρόσωπα σε εικόνες και βίντεο. Αυτό μπορεί να χρησιμοποιηθεί για εφαρμογές αναγνώρισης προσώπων, αυθεντικοποίηση πρόσβασης, δημιουργία προφίλ χρηστών και άλλα.

3. Σάρωση Γραφικών Κωδίκων:

Με το Barcode Scanning API του ML Kit, μπορούμε να σκανάρουμε και να αναγνωρίζουμε γραφικούς κώδικες, όπως QR codes και EAN-13 barcodes. Αυτό μπορεί να χρησιμοποιηθεί για εφαρμογές όπως η αναγνώριση προϊόντων, η αυτόματη σάρωση εισιτηρίων, η ανίχνευση προσωπικών στοιχείων και άλλα.

4. Εφαρμογές Κειμένου:

Αναγνώριση Κειμένου: Χρησιμοποιώντας το Text Recognition API του ML Kit, μπορούμε να αναγνωρίζουμε και να εξάγουμε κείμενο από εικόνες και από τη ζωντανή κάμερα. Αυτό μπορεί να χρησιμοποιηθεί για εφαρμογές όπως η μετάφραση κειμένου, η αναγνώριση κειμένου σε εικόνες που περιέχουν σημειώσεις, ετικέτες, πινακίδες και άλλα.

5. Σάρωση Καρτών Επισκεπτών:

Χρησιμοποιώντας το Text Recognition API, μπορούμε να αναγνωρίζουμε κείμενο από κάρτες επισκεπτών. Αυτό μπορεί να χρησιμοποιηθεί για εφαρμογές που επιτρέπουν τη σάρωση και αυτόματη εισαγωγή προσωπικών στοιχείων από κάρτες επισκεπτών, όπως ονόματα, διευθύνσεις, τηλέφωνα και ηλεκτρονικές διευθύνσεις.

6. Αναγνώριση Χειρόγραφου:

Χρησιμοποιώντας το Text Recognition API, μπορούμε να αναγνωρίζουμε και να μετατρέπουμε χειρόγραφο κείμενο σε ψηφιακή μορφή. Αυτό μπορεί να χρησιμοποιηθεί για εφαρμογές που επιτρέπουν την εισαγωγή χειρόγραφου κειμένου, όπως σημειώσεις, λίστες, γράμματα και άλλα.

7. Αναγνώριση Γλώσσας:

Με το Language Identification API του ML Kit, μπορούμε να αναγνωρίζουμε τη γλώσσα ενός κειμένου. Αυτό μπορεί να χρησιμοποιηθεί για εφαρμογές που αυτόματα αναγνωρίζουν τη γλώσσα του χρήστη και προσαρμόζουν την εμπειρία χρήσης ανάλογα, όπως οι ρυθμίσεις γλώσσας και οι προτεινόμενες περιεχόμενο.

8. Έξυπνες Απαντήσεις:

Με το Smart Reply API του ML Kit, μπορούμε να δημιουργήσουμε εφαρμογές που παράγουν έξυπνες απαντήσεις και προτάσεις από κείμενο. Αυτό μπορεί να χρησιμοποιηθεί για εφαρμογές συνομιλίας, ηλεκτρονικού ταχυδρομείου, κοινωνικών δικτύων και άλλα.

9. Αναγνώριση Προσαρμοσμένων Αντικειμένων:

Χρησιμοποιώντας το Custom Model API του ML Kit, μπορούμε να ενσωματώσουμε προσαρμοσμένα μοντέλα μηχανικής μάθησης για την αναγνώριση προσαρμοσμένων αντικειμένων. Αυτό μπορεί να χρησιμοποιηθεί σε εφαρμογές που απαιτούν την αναγνώριση συγκεκριμένων αντικειμένων ή αντικειμένων σε ειδικά πεδία, όπως αναγνώριση ασθενειών σε ιατρικές εικόνες, εντοπισμός ελαττωματικών προϊόντων σε γραμμές παραγωγής και άλλα.

10. Αυτόνομη Οδήγηση Οχημάτων:

Χρησιμοποιώντας προσαρμοσμένα μοντέλα μηχανικής μάθησης και το Custom Model API του ML Kit, μπορούμε να δημιουργήσουμε εφαρμογές που υποστηρίζουν την αυτόνομη οδήγηση οχημάτων. Αυτό μπορεί να χρησιμοποιηθεί για εφαρμογές σε αυτότον τομέα της αυτοκινητοβιομηχανίας και της μεταφοράς, που εκμεταλλεύονται την τεχνολογία της μηχανικής μάθησης για την ανίχνευση και την αντίδραση σε περιβαλλοντικές συνθήκες, την αυτόματη παρακολούθηση της οδήγησης και την πρόβλεψη της συμπεριφοράς άλλων οδηγών και πεζών.

Αυτές είναι μερικές από τις εφαρμογές του ML Kit στον πραγματικό κόσμο. Ωστόσο, οι δυνατότητες του ML Kit είναι πολλές και οι εφαρμογές μπορούν να προσαρμοστούν σε μια ευρεία γκάμα τομέων και αναγκών. Οι παραπάνω εφαρμογές αποτελούν μόνο μια εικόνα των δυνατοτήτων του ML Kit και υπάρχουν πολλές άλλες δυνατότητες που μπορούν να εξερευνηθούν.





## 4. Υλοποίηση Εφαρμογής

---

### 4.1 Λειτουργικό Σύστημα Android

Το λειτουργικό σύστημα Android [7] είναι ένα λειτουργικό σύστημα για κινητά που αναπτύχθηκε από την Google. Βασίζεται στον πυρήνα Linux και έχει σχεδιαστεί κυρίως για συσκευές με οθόνη αφής όπως smartphone και tablet. Το Android έχει γίνει το πιο ευρέως χρησιμοποιούμενο λειτουργικό σύστημα για κινητά παγκοσμίως.

#### 4.1.1 Βασικά Χαρακτηριστικά

1. Customization

Το Android προσφέρει υψηλό βαθμό προσαρμογής(customization), επιτρέποντας στους χρήστες να εξατομικεύουν την εμφάνιση και τη λειτουργικότητα των συσκευών τους. Οι χρήστες μπορούν να προσαρμόσουν τη διάταξη της αρχικής οθόνης, να προσθέσουν γραφικά στοιχεία, να αλλάξουν θέματα και να εγκαταστήσουν διάφορες εφαρμογές τρίτων για να προσαρμόσουν τη συσκευή τους στις προτιμήσεις τους.

2. Application Ecosystem

Το Android διαθέτει ένα τεράστιο και ποικίλο οικοσύστημα εφαρμογών που διατίθεται μέσω του Google Play Store. Οι χρήστες μπορούν να κάνουν λήψη και εγκατάσταση εφαρμογών για διάφορους σκοπούς, όπως παραγωγικότητα, ψυχαγωγία, κοινωνική δικτύωση, παιχνίδια και άλλα. Η ανοιχτή φύση του Android επιτρέπει στους προγραμματιστές να δημιουργούν και να διανέμουν εφαρμογές ανεξάρτητα.

3. Google Integration

Το Android ενσωματώνεται στενά με τις υπηρεσίες της Google, όπως η Αναζήτηση Google, οι Χάρτες Google, το Gmail, το Google Drive και άλλα. Αυτή η ενσωμάτωση παρέχει απρόσκοπτη πρόσβαση στο οικοσύστημα της Google, συμπεριλαμβανομένου του χώρου αποθήκευσης cloud, του email, της πλοήγησης και των διαδικτυακών υπηρεσιών.

4. Multitasking

Το Android υποστηρίζει πολλαπλές εργασίες, επιτρέποντας στους χρήστες να εκτελούν πολλές εφαρμογές ταυτόχρονα. Οι χρήστες μπορούν να κάνουν εναλλαγή μεταξύ εφαρμογών άνετα να χρησιμοποιούν τη λειτουργία διαχωρισμού οθόνης για να προβάλλουν πολλές εφαρμογές δίπλα-δίπλα και να λαμβάνουν ειδοποιήσεις σε πραγματικό χρόνο.

5. Google Play Services

Το Android χρησιμοποιεί τις Υπηρεσίες Google Play, ένα σύνολο API και υπηρεσιών που παρέχουν βελτιωμένη λειτουργικότητα σε συσκευές Android. Οι Υπηρεσίες Google Play προσφέρουν λειτουργίες όπως υπηρεσίες που βασίζονται στην τοποθεσία, ανταλλαγή

μηνυμάτων στο cloud, έλεγχος ταυτότητας και άλλα, δίνοντας τη δυνατότητα στους προγραμματιστές να αξιοποιήσουν ισχυρές δυνατότητες στις εφαρμογές τους.

#### 6. Device Compatibility

Το Android είναι συμβατό με ένα ευρύ φάσμα συσκευών, από smartphone φιλικά προς την τιμή έως κορυφαίες συσκευές. Υποστηρίζει διαφορετικά μεγέθη οθόνης, αναλύσεις και διαμορφώσεις υλικού, διασφαλίζοντας τη συμβατότητα με διάφορες συσκευές και κατασκευαστές.

#### 7. Regular Updates

Η Google κυκλοφορεί τακτικές ενημερώσεις στο λειτουργικό σύστημα Android, εισάγοντας νέες δυνατότητες, διορθώσεις σφαλμάτων, βελτιώσεις ασφαλείας και βελτιώσεις απόδοσης. Αυτές οι ενημερώσεις παραδίδονται over-the-air (OTA) και συχνά συνοδεύονται από ενημερώσεις στις Υπηρεσίες Google Play.

#### 8. Open Source

Το Android είναι μια πλατφόρμα ανοιχτού κώδικα, που σημαίνει ότι ο πηγαίος κώδικας του είναι δωρεάν διαθέσιμος στους προγραμματιστές. Αυτό το άνοιγμα έχει ενθαρρύνει μια μεγάλη και ενεργή κοινότητα προγραμματιστών που συμβάλλουν στην ανάπτυξή της και δημιουργούν καινοτόμες εφαρμογές

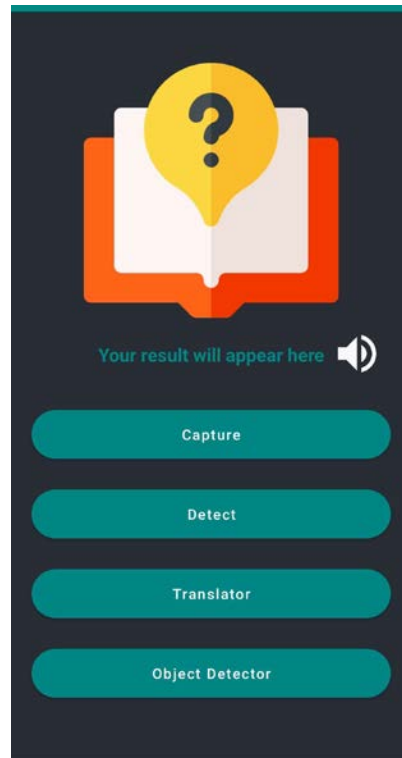
Εν κατακλείδι το Android έχει φέρει επανάσταση στη βιομηχανία των κινητών, ενδυναμώνοντας τους χρήστες με ένα ευέλικτο και προσαρμόσιμο λειτουργικό σύστημα. Ο ανοιχτός χαρακτήρας του και το εκτεταμένο οικοσύστημα εφαρμογών του συνέβαλαν στην ευρεία υιοθέτησή του και τη συνεχή ανάπτυξή του ως κορυφαία πλατφόρμα για κινητά.

## 4.2 Υλοποίηση Εφαρμογής

Στα πλαίσια της παρούσας εργασίας δημιουργήθηκε μια εφαρμογή Android που εκμεταλεύεται το Text Recognition API της βιβλιοθήκη της Google ML Kit καθώς επίσης το Tensorflow Lite και το OpenCV για την λειτουργία Object Detection.

## 4.2.1 UI

Στην κεντρική οθόνη υπάρχουν 4 κουμπιά ελέγχου το καθένα με διαφορετική λειτουργία τα οποία συνιστούν το κύριο μενού της εφαρμογής. Παρακάτω αναλύεται η λειτουργία κάθε κουμπιού ελέγχου.



Εικόνα 1: Main Activity

## 4.2.2 Capture

Αυτό το κουμπί έχει σκοπό να τραβήξει μια εικόνα χρησιμοποιώντας την κάμερα της συσκευής, ελέγχει εάν υπάρχει διαθέσιμη εφαρμογή κάμερας και, στη συνέχεια, ξεκινά την εφαρμογή κάμερας για τη λήψη της εικόνας. Σε κώδικα μεταφράζεται ως:

```
private void captureImage(){
    Intent takePicture = new Intent(MediaStore.ACTION_IMAGE_CAPTURE);
    if(takePicture.resolveActivity(getPackageManager()) != null){
        startActivityForResult(takePicture, REQUEST_IMAGE_CAPTURE);
    }
}
```

Εικόνα 2: Capture code

Ο κώδικας προετοιμάζει ένα αντικείμενο Intent που ονομάζεται "takePicture" με την ενέργεια MediaStore.ACTION\_IMAGE\_CAPTURE. Αυτή η ενέργεια υποδεικνύει ότι η πρόθεση προορίζεται για λήψη μιας εικόνας χρησιμοποιώντας την κάμερα της συσκευής. Η δήλωση if ελέγχει εάν υπάρχει διαθέσιμη δραστηριότητα για τη διαχείριση της πρόθεσης "takePicture".

Με άλλα λόγια, ελέγχει εάν υπάρχει εγκατεστημένη εφαρμογή κάμερας στη συσκευή που μπορεί να χειριστεί τη δράση λήψης εικόνας. Εάν υπάρχει διαθέσιμη δραστηριότητα, καλείται η μέθοδος `startActivityForResult()` με την πρόθεση "takePicture" και έναν κωδικό αιτήματος που ονομάζεται "REQUEST\_IMAGE\_CAPTURE". Αυτή η μέθοδος ξεκινά την εφαρμογή κάμερας, επιτρέποντας στον χρήστη να τραβήξει μια εικόνα. Αφού ολοκληρωθεί η δραστηριότητα λήψης εικόνων, το αποτέλεσμα θα επιστρέψει στη δραστηριότητα κλήσης (αυτή που ξεκίνησε τη λήψη εικόνας). Το αποτέλεσμα θα γίνει με τη μέθοδο `onActivityResult()`, όπου μπορεί να γίνει επεξεργασία της εικόνας που τραβήχτηκε ή να εκτελεστούν οποιεσδήποτε άλλες απαραίτητες ενέργειες με βάση το αποτέλεσμα.

### 4.2.3 Detect

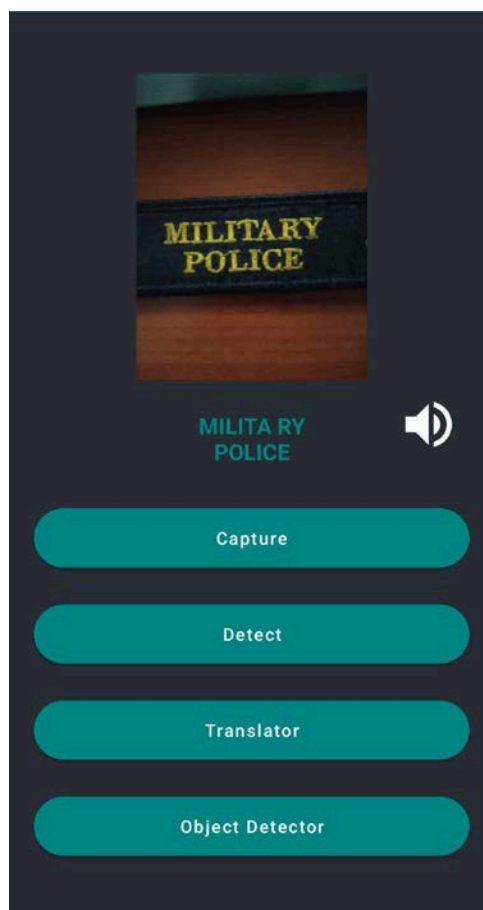
Αυτό το κουμπί εκτελεί αναγνώριση κειμένου σε μια εικόνα εισόδου χρησιμοποιώντας το `TextRecognizer` του `ML Kit`. Επεξεργάζεται την εικόνα, εξαγάγει το αναγνωρισμένο κείμενο σε διάφορα επίπεδα (μπλοκ, γραμμές και στοιχεία) και εμφανίζει το εξαγόμενο κείμενο σε ένα `TextView`. Διαχειρίζεται επίσης περιπτώσεις αποτυχίας και εμφανίζει ένα μήνυμα σφάλματος εάν αποτύχει η αναγνώριση κειμένου. Σε κώδικα μεταφράζεται ως:

```
private void detectText(){
    InputImage image = InputImage.fromBitmap(imgBitmap, rotationDegrees: 0);
    TextRecognizer recognizer = TextRecognition.getClient(TextRecognizerOptions.DEFAULT_OPTIONS);
    Task<Text> result = recognizer.process(image).addOnSuccessListener(new OnSuccessListener<Text>() {
        @Override
        public void onSuccess(Text text) {
            StringBuilder result = new StringBuilder();
            for(Text.TextBlock block: text.getTextBlocks()){
                String blockText = block.getText();
                Point[] blockCornerPoint = block.getCornerPoints();
                Rect blockFrame = block.getBoundingBox();
                for(Text.Line line : block.getLines()){
                    String lineText = line.getText();
                    Point[] lineCornerPoint = line.getCornerPoints();
                    Rect lineRect = line.getBoundingBox();
                    for (Text.Element element: line.getElements()){
                        String elementText = element.getText();
                        result.append(elementText);
                    }
                    tvResult.setText(blockText);
                }
            }
        }
    }).addOnFailureListener(new OnFailureListener() {
        @Override
        public void onFailure(@NonNull Exception e) {
            Toast.makeText(context: ScannerActivity.this, text: "Fail to detect text"+e.getMessage(), Toast.LENGTH_SHORT).show();
        }
    });
}
```

Εικόνα 3: Detect code

Ο κώδικας ξεκινά δημιουργώντας ένα αντικείμενο `InputImage` με το όνομα "image" χρησιμοποιώντας τη μέθοδο `fromBitmap()`. Αυτό το αντικείμενο αντιπροσωπεύει την εικόνα εισόδου για την αναγνώριση κειμένου και λαμβάνει το bitmap `imgBitmap` και μια τιμή περιστροφής (σε αυτήν την περίπτωση, 0). Στη συνέχεια, δημιουργεί ένα αντικείμενο `TextRecognizer` με το όνομα "recognizer" χρησιμοποιώντας το `TextRecognition.getClient()` με το `TextRecognizerOptions.DEFAULT_OPTIONS`. Αυτό το αντικείμενο είναι υπεύθυνο για την

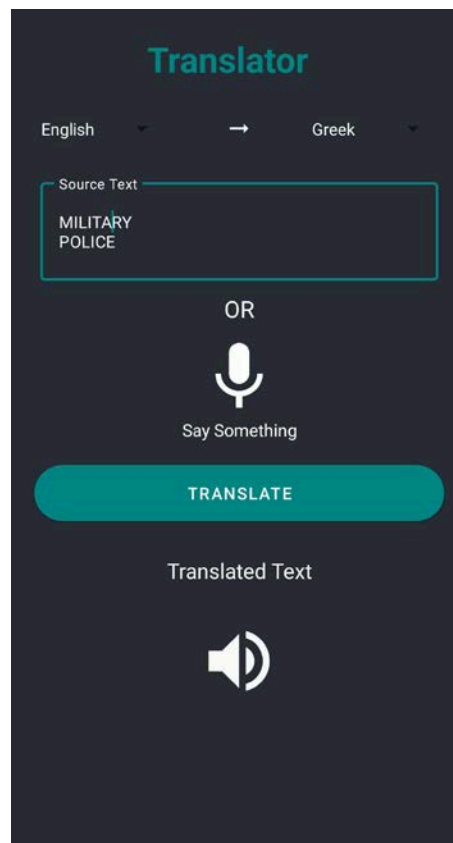
αναγνώριση κειμένου στην εικόνα εισόδου. Η μέθοδος `process()` καλείται στο αντικείμενο `"recogniser"`, περνώντας στην `"image"`. Αυτό ξεκινά τη διαδικασία αναγνώρισης κειμένου και επιστρέφει ένα αντικείμενο `Task<Text>` με το όνομα `"result"`. Ο κώδικας προσθέτει ένα `OnSuccessListener` στην εργασία `"result"`. Όταν η αναγνώριση κειμένου είναι επιτυχής, ενεργοποιείται η μέθοδος `onSuccess()`. Μέσα σε αυτήν τη μέθοδο, ο κώδικας επαναλαμβάνεται μέσω των αναγνωρισμένων μπλοκ κειμένου, γραμμών και στοιχείων για την εξαγωγή του κειμένου και την εμφάνιση του. Για κάθε μπλοκ κειμένου, ο κώδικας ανακτά το κείμενο, τα γωνιακά σημεία και το πλαίσιο οριοθέτησης του μπλοκ. Για κάθε γραμμή εντός του μπλοκ, ανακτά το κείμενο της γραμμής, τα σημεία γωνίας και το πλαίσιο οριοθέτησης. Τέλος, για κάθε στοιχείο εντός της γραμμής, ανακτά το κείμενο του στοιχείου και το προσαρτά στο αποτέλεσμα `StringBuilder`. Το εξαγόμενο κείμενο ορίζεται στη συνέχεια σε ένα `TextView` με το όνομα `"tvResult"` χρησιμοποιώντας τη μέθοδο `setText()`. Εάν αποτύχει η αναγνώριση κειμένου, ο κώδικας προσθέτει ένα `OnFailureListener` στην εργασία `"result"`. Η μέθοδος `onFailure()` ενεργοποιείται όταν προκύπτει μια εξαίρεση κατά τη διαδικασία αναγνώρισης κειμένου. Σε αυτήν την περίπτωση, εμφανίζει ένα μήνυμα `Toast` που υποδεικνύει την αποτυχία και το σχετικό μήνυμα σφάλματος.



Εικόνα 4: Το αποτέλεσμα του button Detect

#### 4.2.4 Translator

Με το πάτημα του button Translate ανοίγει η ομόνυμη δραστηριότητα



Εικόνα 5:Translator Activity

Στο πάνω μέρος της οθόνης υπάρχουν 2 spinner όπου περιέχουν τις γλώσσες εισαγωγής και τις γλώσσες προορισμού. Αυτές είναι : Ελληνικά, Αγγλικά, Γερμανικά, Ιταλικά, Ισπανικά, Γαλλικά, Ρώσικα, Κορεάτικα και Ιαπωνικά.



Εικόνα 6: Επιλογή Γλώσσας

Το κουμπί Translate εμφανίζει τη διαδικασία μετάφρασης χρησιμοποιώντας το API Natural Language του Firebase ML Kit. Κατεβάζει το απαιτούμενο μοντέλο μετάφρασης εάν χρειάζεται, μεταφράζει το κείμενο πηγής και εμφανίζει το μεταφρασμένο αποτέλεσμα στον χρήστη. Διαχειρίζεται επίσης περιπτώσεις αποτυχίας και εμφανίζει κατάλληλα μηνύματα σφάλματος όταν η μετάφραση ή η λήψη μοντέλου αποτυγχάνει. Σε κώδικα μεταφράζεται ως:

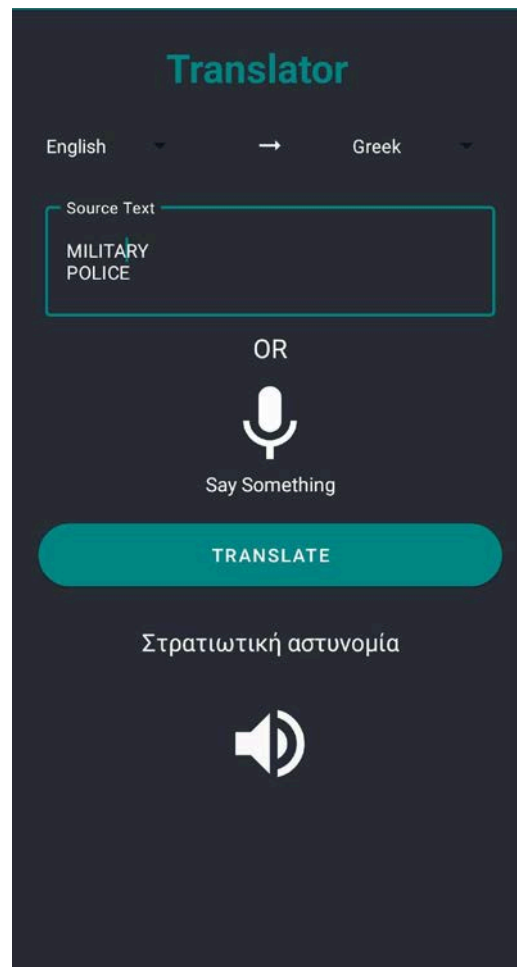
```
private void translateText(int fromLanguageCode, int toLanguageCode, String sourceText){
    tvTransRes.setText("Downloading Model");
    FirebaseTranslatorOptions translatorOptions = new FirebaseTranslatorOptions.Builder()
        .setSourceLanguage(fromLanguageCode)
        .setTargetLanguage(toLanguageCode)
        .build();

    FirebaseTranslator translator = FirebaseNaturalLanguage.getInstance().getTranslator(translatorOptions);
    FirebaseModelDownloadConditions conditions = new FirebaseModelDownloadConditions.Builder().build();
    translator.downloadModelIfNeeded(conditions).addOnSuccessListener(new OnSuccessListener<Void>() {
        @Override
        public void onSuccess(Void unused) {
            tvTransRes.setText("Translating...");
            translator.translate(sourceText).addOnSuccessListener(new OnSuccessListener<String>() {
                @Override
                public void onSuccess(String s) { tvTransRes.setText(s); }
            }).addOnFailureListener(new OnFailureListener() {
                @Override
                public void onFailure(@NonNull Exception e) {
                    Toast.makeText(context, TranslatorActivity.this, text: "Translation failed "+e.getMessage(), Toast.LENGTH_SHORT).show();
                }
            });
        }
    }).addOnFailureListener(new OnFailureListener() {
        @Override
        public void onFailure(@NonNull Exception e) {
            Toast.makeText(context, TranslatorActivity.this, text: "Download failed "+e.getMessage(), Toast.LENGTH_SHORT).show();
        }
    });
}
```

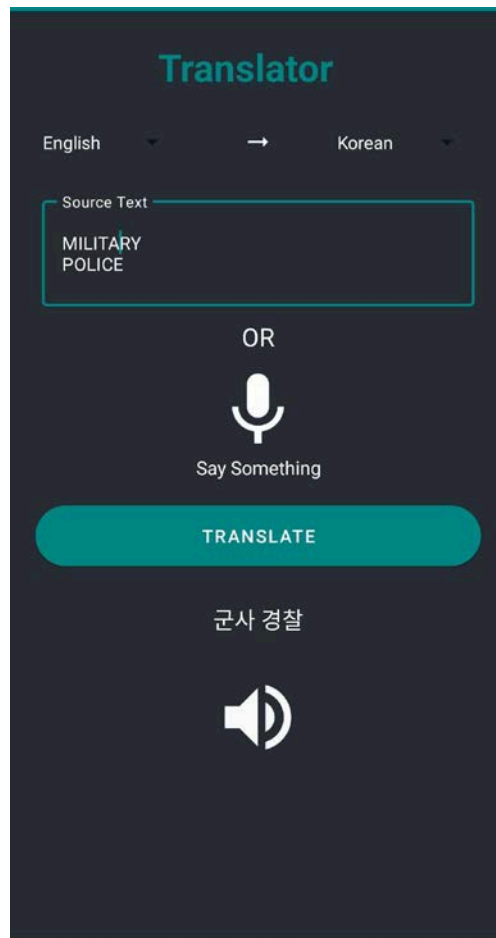
Εικόνα 7: Translate code

Ο κώδικας ξεκινά ορίζοντας το κείμενο ενός TextView που ονομάζεται "tvTransRes" σε "Downloading Model". Αυτό είναι πιθανό να ενημερώσει τον χρήστη σχετικά με τη διαδικασία λήψης μοντέλου που βρίσκεται σε εξέλιξη. Δημιουργεί ένα αντικείμενο FirebaseTranslatorOptions με το όνομα "translatorOptions" χρησιμοποιώντας το FirebaseTranslatorOptions.Builder(). Αυτό το αντικείμενο καθορίζει τη γλώσσα πηγής (fromLanguageCode) και τη γλώσσα προορισμού (toLanguageCode) για μετάφραση. Ο κώδικας αρχικοποιεί ένα αντικείμενο FirebaseTranslator που ονομάζεται "translator" καλώντας το FirebaseNaturalLanguage.getInstance().getTranslator(translatorOptions). Αυτό ανακτά μια παρουσία του μεταφραστή με τις καθορισμένες επιλογές μετάφρασης. Δημιουργεί ένα αντικείμενο FirebaseModelDownloadConditions με το όνομα "conditions" χρησιμοποιώντας το νέο FirebaseModelDownloadConditions.Builder().build(). Αυτό το αντικείμενο ορίζει τις συνθήκες για τη λήψη του μοντέλου μετάφρασης, όπως τη διαθεσιμότητα δικτύου. Ο κωδικός καλεί το translator.downloadModelIfNeeded(conditions) για λήψη του μοντέλου μετάφρασης εάν δεν έχει ήδη ληφθεί. Προσθέτει ένα OnSuccessListener στην εργασία, το οποίο ενεργοποιείται όταν η λήψη του μοντέλου είναι επιτυχής. Μέσα στο OnSuccessListener, ο κώδικας ενημερώνει το κείμενο του "tvTransRes" σε "Translating...". Αυτό πιθανώς υποδηλώνει ότι η διαδικασία μετάφρασης βρίσκεται σε εξέλιξη. Ο κώδικας καλεί το translator.translate(sourceText) για να μεταφράσει το "sourceText" στη γλώσσα-στόχο. Προσθέτει ένα OnSuccessListener στην εργασία μετάφρασης, η οποία ενεργοποιείται όταν η μετάφραση είναι επιτυχής. Μέσα στο OnSuccessListener, το μεταφρασμένο κείμενο (αποθηκευμένο στη μεταβλητή "s") ορίζεται σε "tvTransRes", αντικαθιστώντας το προηγούμενο κείμενο. Αυτό εμφανίζει το μεταφρασμένο αποτέλεσμα στον χρήστη. Εάν η εργασία μετάφρασης αποτύχει, προστίθεται ένα OnFailureListener για να χειριστεί την εξαίρεση. Εμφανίζει ένα μήνυμα Toast που υποδεικνύει την αποτυχία και το σχετικό μήνυμα σφάλματος. Ομοίως, εάν η εργασία λήψης μοντέλου αποτύχει, προστίθεται ένα άλλο OnFailureListener για να χειριστεί την εξαίρεση. Εμφανίζει ένα μήνυμα Toast που υποδεικνύει την αποτυχία και το σχετικό μήνυμα σφάλματος.





Εικόνα 8: Αποτέλεσμα Μετάφρασης στα Ελληνικά



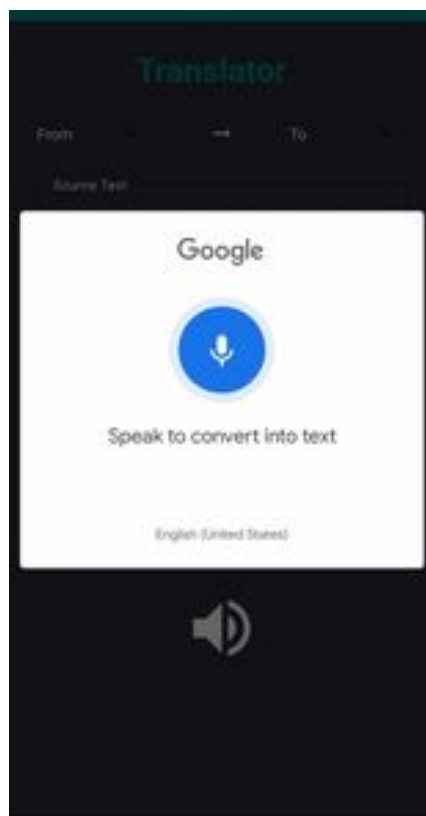
Εικόνα 9:Αποτέλεσμα Μετάφρασης στα Κορεάτικα

Το πάτημα της εικόνας του μικροφώνου ξεκινά τη δραστηριότητα αναγνώρισης ομιλίας και χειρίζεται τυχόν εξαιρέσεις που ενδέχεται να προκύψουν κατά τη διάρκεια της διαδικασίας. Σε κώδικα μεταφράζεται ως:

```
ivMic.setOnClickListener(new View.OnClickListener() {  
    @Override  
    public void onClick(View view) {  
        Intent intent = new Intent(RecognizerIntent.ACTION_RECOGNIZE_SPEECH);  
        intent.putExtra(RecognizerIntent.EXTRA_LANGUAGE_MODEL, RecognizerIntent.LANGUAGE_MODEL_FREE_FORM);  
        intent.putExtra(RecognizerIntent.EXTRA_LANGUAGE, Locale.getDefault());  
        intent.putExtra(RecognizerIntent.EXTRA_PROMPT, value: "Speak to convert into text");  
        try {  
            startActivityForResult(intent, REQUEST_PERMISSION_CODE);  
        } catch (Exception e) {  
            e.printStackTrace();  
            Toast.makeText(context: TranslatorActivity.this, e.getMessage(), Toast.LENGTH_SHORT).show();  
        }  
    }  
});
```

Εικόνα 10: Mic code

Όταν γίνει κλικ στη συσχετισμένη Προβολή, ενεργοποιείται η μέθοδος `onClick()`. Μέσα στη μέθοδο `onClick()`, δημιουργείται ένα αντικείμενο `Intent` με το όνομα "intent" με την ενέργεια `RecognizerIntent.ACTION_RECOGNIZE_SPEECH`. Αυτή η ενέργεια υποδεικνύει ότι η πρόθεση προορίζεται για την αναγνώριση ομιλίας. Ο κώδικας ορίζει πρόσθετα για την πρόθεση (intent) διαμόρφωσης της διαδικασίας αναγνώρισης ομιλίας: `intent.putExtra(RecognizerIntent.EXTRA_LANGUAGE_MODEL, RecognizerIntent.LANGUAGE_MODEL_FREE_FORM)`: Αυτή η γραμμή ορίζει το γλωσσικό μοντέλο σε `LANGUAGE_MODEL_FREE_FORM`, υποδεικνύοντας ότι η αναγνώριση ομιλίας θα πρέπει να προσαρμόζει την ομιλία χωρίς καμία ελεύθερη ερμηνεία της ομιλίας. `intent.putExtra(RecognizerIntent.EXTRA_LANGUAGE, Locale.getDefault())`: Αυτή η γραμμή ορίζει τη γλώσσα για την αναγνώριση ομιλίας στην προεπιλεγμένη τοπική ρύθμιση της συσκευής, διασφαλίζοντας ότι η αναγνώριση εκτελείται στην προτιμώμενη γλώσσα του χρήστη. `intent.putExtra(RecognizerIntent.EXTRA_PROMPT, "Speak to convert into text")`: Αυτή η γραμμή ορίζει ένα μήνυμα προτροπής που θα εμφανίζεται στον χρήστη για να τον ενθαρρύνει να μιλήσει για αναγνώριση ομιλίας. Στη συνέχεια, ο κώδικας επιχειρεί να ξεκινήσει τη δραστηριότητα για την αναγνώριση ομιλίας καλώντας το `startActivityForResult` (πρόθεση, `REQUEST_PERMISSION_CODE`). Αυτή η μέθοδος ξεκινά τη δραστηριότητα και αναμένει να επιστραφεί ένα αποτέλεσμα όταν ολοκληρωθεί η διαδικασία αναγνώρισης ομιλίας. Εάν παρουσιαστεί οποιαδήποτε εξαίρεση κατά την εκτέλεση του μπλοκ κώδικα (π.χ., εάν η συσκευή δεν υποστηρίζει αναγνώριση ομιλίας), η εξαίρεση καταγράφεται στο μπλοκ σύλληψης (catch block). Το μήνυμα της εξαίρεσης εκτυπώνεται στο ίχνος στοίβας και εμφανίζεται ένα μήνυμα Toast στον χρήστη που υποδεικνύει το μήνυμα εξαίρεσης



Εικόνα 11: Microphone View

Τέλος το πάτημα του μεγάφωνου εκτελεί μετατροπή κειμένου σε ομιλία με βάση το περιεχόμενο ενός TextView. Καθορίζει την κατάλληλη γλώσσα με βάση την τιμή "toLanguageCode" και χρησιμοποιεί το αντικείμενο TextToSpeech για να εκφωνήσει το κείμενο. Σε κώδικα μεταφράζεται ως:

```
ivSpeakerT.setOnClickListener(new View.OnClickListener() {
    @Override
    public void onClick(View view) {
        String text = tvTransRes.getText().toString();
        switch(toLanguageCode) {
            case 9:
                t2s.setLanguage(Locale.GERMAN);
                break;
            case 17:
                t2s.setLanguage(Locale.FRENCH);
                break;
            case 28:
                t2s.setLanguage(Locale.ITALIAN);
                break;
            case 32:
                t2s.setLanguage(Locale.KOREAN);
                break;
            case 29:
                t2s.setLanguage(Locale.JAPANESE);
                break;
            case 10:
                t2s.setLanguage(Locale.forLanguageTag("el-GR"));
                break;
            case 44:
                t2s.setLanguage(Locale.forLanguageTag("ru-RU"));
                break;
            case 13:
                t2s.setLanguage(Locale.forLanguageTag("es-ES"));
                break;
            default:
                t2s.setLanguage(Locale.US);
                Timber.d("message: "Language Code -> %d", languageCode);
                break;
        }
        t2s.speak(text, TextToSpeech.QUEUE_FLUSH, params: null);
    }
});
```

Εικόνα 12: Speaker Code

Όταν γίνει κλικ στη συσχετισμένη Προβολή, ενεργοποιείται η μέθοδος onClick(). Μέσα στη μέθοδο onClick(), δημιουργείται μια μεταβλητή String με το όνομα "text" και της εκχωρείται η τιμή του κειμένου που ανακτήθηκε από ένα TextView με το όνομα "tvTransRes". Μετατρέπει το κείμενο σε συμβολοσειρά χρησιμοποιώντας τη μέθοδο toString(). Στη συνέχεια, ο κώδικας χρησιμοποιεί μια δήλωση διακόπτη (switch statement) για να χειριστεί διαφορετικές περιπτώσεις με βάση την τιμή της μεταβλητής "toLanguageCode". Κάθε μπλοκ περίπτωσης

(case) αντιστοιχεί σε έναν συγκεκριμένο κώδικα γλώσσας και με βάση τον κώδικα γλώσσας, η μέθοδος `setLanguage()` ενός αντικειμένου `TextToSpeech` καλείται για να ορίσει την κατάλληλη γλώσσα για τη μετατροπή κειμένου σε ομιλία. Μετά τη ρύθμιση της γλώσσας, ο κώδικας καλεί το `tts.speak(text, TextToSpeech.QUEUE_FLUSH, null)` για να μετατρέψει τη συμβολοσειρά "text" σε ομιλία χρησιμοποιώντας τη μετατροπή κειμένου σε ομιλία. Κάνει ουρά στην ομιλία για άμεση αναπαραγωγή, αντικαθιστώντας οποιαδήποτε ομιλία που εκφωνείται αυτήν τη στιγμή.

#### 4.2.5 Object Detector

Με το πάτημα του κουμπιού ανοίγει το ομώνυμο Activity. Για την επίτευξη της αναγνώρισης αντικειμένων χρησιμοποιήθηκε το Open CV για την όραση της συσκευής και το Tensorflow Lite για την εγκατάσταση του μοντέλου.



Εικόνα 13: Object Detector

Η παραπάνω λειτουργία χωρίζεται σε δύο μέρη. Στη Ανάλυση Εικόνας και το Object Detection. Το Object Detection επεξεργάζεται μια εικόνα αρχικοποιώντας έναν ανιχνευτή αντικειμένων, αντιγράφοντας δεδομένα εικόνας, πραγματοποιώντας ανίχνευση αντικειμένων, σχεδιάζοντας τα αντικείμενα που εντοπίστηκαν σε καμβά και ενημερώνοντας τη διεπαφή χρήστη με το χρόνο επεξεργασίας. Είναι μια κοινή δομή για την υλοποίηση εργασιών όρασης

υπολογιστή σε πραγματικό χρόνο σε εφαρμογές Android. Σε κώδικα μεταφράζεται ως:

```

override fun analyze(image: ImageProxy) {
    if (image.planes.isEmpty()) {return}
    if (detectorAddr == 0L) {
        detectorAddr = initDetector(this.assets)
    }

    val buffer = image.planes[0].buffer
    val size = buffer.capacity()
    if (!::rgbaFrame.isInitialized) {
        rgbaFrame = ByteArray(size)
    }

    buffer.position(newPosition: 0)
    buffer.get(rgbaFrame, offset: 0, size)

    val start = System.currentTimeMillis()
    val res = detect(detectorAddr, rgbaFrame, image.width, image.height)
    val span = System.currentTimeMillis() - start

    val canvas = binding.surfaceView.holder.lockCanvas()
    if (canvas != null) {
        canvas.drawColor(Color.TRANSPARENT, PorterDuff.Mode.MULTIPLY)

        // Draw the detections, in our case there are only 3
        for (i in 0..until < res[0]) {
            this.drawDetection(canvas, image.width, image.height, res, i)
        }

        binding.surfaceView.holder.unlockCanvasAndPost(canvas)
    }

    runOnUiThread {
        binding.lblStatus.text = "$span ms"
    }

    image.close()
}

```

Εικόνα 14- Object Detection code

- Χειρισμός σφαλμάτων: `if (image.planes.isEmpty()) {return}`: Αυτή η συνθήκη ελέγχει εάν το αντικείμενο εικόνας δεν έχει επίπεδα εικόνας. Εάν δεν υπάρχουν επίπεδα, η συνάρτηση επιστρέφει νωρίς, υποδεικνύοντας ότι δεν υπάρχουν δεδομένα εικόνας για επεξεργασία.
- Αρχικοποίηση ανιχνευτή αντικειμένων: `if (detectorAddr == 0L) {...}`: Αυτή η συνθήκη ελέγχει εάν μια μεταβλητή `detectorAddr` είναι μηδέν. Εάν είναι, υποδηλώνει ότι ένας ανιχνευτής αντικειμένων δεν έχει αρχικοποιηθεί ακόμα. Εάν δεν είναι αρχικοποιημένο, ο κώδικας καλεί το `initDetector(this.assets)` για να το προετοιμάσει.
- Χειρισμός buffer εικόνας: `val buffer = image.planes[0].buffer`: Αυτή η γραμμή ανακτά το buffer δεδομένων εικόνας από το πρώτο επίπεδο εικόνας. `val size = buffer.capacity ()`: Υπολογίζει το μέγεθος του buffer δεδομένων εικόνας.
- Εκκίνηση του rgbaFrame: `if (!::rgbaFrame.isInitialized) {...}`: Αυτό ελέγχει εάν η μεταβλητή `rgbaFrame` δεν έχει αρχικοποιηθεί. Εάν δεν έχει αρχικοποιηθεί, το αρχικοποιεί ως πίνακα byte με μέγεθος ίσο με το μέγεθος του buffer δεδομένων εικόνας.

- Αντιγραφή δεδομένων εικόνας: `buffer.position(0)`: Ορίζει τη θέση του `buffer` στην αρχή (θέση 0). `buffer.get (rgbaFrame, 0, μέγεθος)`: Αντιγράφει τα δεδομένα εικόνας από το `buffer` στον πίνακα `byte rgbaFrame`.
- Ανίχνευση αντικειμένου χρονισμού: `val start = System.currentTimeMillis()`: Καταγράφει την τρέχουσα ώρα πριν από την έναρξη της ανίχνευσης αντικειμένου. `Val res = detect (detectorAddr, rgbaFrame, image.width, image.height)`: Καλεί μια συνάρτηση με το όνομα `detect` με παραμέτρους που περιλαμβάνουν τη διεύθυνση του ανιχνευτή, τα δεδομένα εικόνας, το πλάτος και το ύψος. Αυτό πιθανώς εκτελεί τη διαδικασία ανίχνευσης αντικειμένων και επιστρέφει ένα αποτέλεσμα στη μεταβλητή `res`. Υπολογίζει το χρονικό διάστημα της διαδικασίας ανίχνευσης: `val span = System.currentTimeMillis() - start`.
- Σχέδιο σε καμβά: Προετοιμάζει έναν καμβά για σχέδιο στο UI: `val canvas = binding.surfaceView.holder.lockCanvas()`. Εάν ληφθεί έγκυρος καμβάς, τότε: Καθαρίζει τον καμβά με ένα διαφανές χρώμα: `canvas.drawColor(Color.TRANSPARENT, PorterDuff.Mode.MULTIPLY)`. Επαναλαμβάνει τα αντικείμενα που έχουν εντοπιστεί και καλεί `this.drawDetection` για να τα σχεδιάσει στον καμβά.
- Ενημέρωση διεπαφής χρήστη: `binding.lblStatus.text = "$span ms"`: Ενημερώνει μια ετικέτα διεπαφής χρήστη με το χρόνο που απαιτείται για τη διαδικασία ανίχνευσης αντικειμένων.
- Εικόνα κλεισίματος: `image.close()`: Κλείνει την εικόνα, απελευθερώνοντας σχετικούς πόρους.

Τέλος ο κώδικας στην ανάλυση εικόνας ρυθμίζει και διαμορφώνει την κάμερα της συσκευής Android για προεπισκόπηση και ανάλυση εικόνας χρησιμοποιώντας τη βιβλιοθήκη `CameraX`. Χειρίζεται πτυχές όπως η περιστροφή της οθόνης, η αναλογία διαστάσεων, η στρατηγική αντίθλιψης και το `resolution`. Σε κώδικα μεταφράζεται ως:

```

val cameraProviderFuture = ProcessCameraProvider.getInstance(context, this)

cameraProviderFuture.addListener(Runnable {
    // Used to bind the lifecycle of cameras to the lifecycle owner
    val cameraProvider: ProcessCameraProvider = cameraProviderFuture.get()

    val rotation = binding.viewFinder.display.rotation

    // Preview
    val preview = Preview.Builder() Preview.Builder
        .setTargetAspectRatio(AspectRatio.RATIO_4_3)
        .setTargetRotation(rotation)
        .build() Preview
        .also { it: Preview
            it.setSurfaceProvider(binding.viewFinder.surfaceProvider)
        }

    // ImageAnalysis
    val imageAnalyzer = ImageAnalysis.Builder() ImageAnalysis.Builder
        .setBackpressureStrategy(STRATEGY_KEEP_ONLY_LATEST)
        .setTargetResolution(Size( width: 768, height: 1024))
        .setTargetRotation(rotation)
        .setOutputImageRotationEnabled(true)
        .setOutputImageFormat(ImageAnalysis.OUTPUT_IMAGE_FORMAT_RGBA_8888)
        .build() ImageAnalysis
        // The analyzer can then be assigned to the instance
        .also { it: ImageAnalysis
            it.setAnalyzer(cameraExecutor, analyzer, this)
        }

    try {
        // Unbind use cases before rebinding
        cameraProvider.unbindAll()

        // Bind use cases to camera
        cameraProvider.bindToLifecycle(
            lifecycleOwner, this,
            CameraSelector.DEFAULT_BACK_CAMERA,
            preview,
            imageAnalyzer
        )
    } catch (exc: Exception) {
        Log.e(TAG, msg: "Use case binding failed", exc)
    }
}, ContextCompat.getMainExecutor(context, this))

```

Εικόνα 15- Ανάλυση Εικόνας

- Val cameraProviderFuture = ProcessCameraProvider.getInstance (this): Αυτή η γραμμή προετοιμάζει ένα ProcessCameraProviderFuture για να αποκτήσει ένα στιγμιότυπο του ProcessCameraProvider, το οποίο χρησιμοποιείται για τη διαχείριση της λειτουργικότητας της κάμερας.
- cameraProviderFuture.addListener(Runnable{...},ContextCompat.getMainExecutor(this)): Αυτός ο κώδικας προσθέτει ένα πρόγραμμα ακρόασης στο cameraProviderFuture που θα εκτελεστεί όταν ο πάροχος της κάμερας είναι έτοιμος. Χρησιμοποιεί το ContextCompat.getMainExecutor(this) για να διασφαλίσει ότι ο ακροατής εκτελείται στο κύριο νήμα.
- Μέσα στον ακροατή: val cameraProvider: ProcessCameraProvider = cameraProviderFuture.get(): Λαμβάνει τον πάροχο κάμερας όταν είναι έτοιμος.val rotation = binding.viewFinder.display.rotation: Αυτή η γραμμή υπολογίζει την



τρέχουσα περιστροφή της οθόνης της συσκευής. Η περιστροφή της οθόνης είναι ζωτικής σημασίας για τη διαμόρφωση της προεπισκόπησης της κάμερας.

- Προεπισκόπηση διαμόρφωσης: `val preview = Preview.Builder() ...`: Διαμορφώνει μια περίπτωση χρήσης προεπισκόπησης με συγκεκριμένες ρυθμίσεις: Ο λόγος διαστάσεων ορίστηκε σε 4:3. Η περιστροφή στόχου ορίστηκε στην υπολογισμένη περιστροφή της οθόνης. Δημιουργεί την περίπτωση χρήσης προεπισκόπησης και ορίζει τον πάροχο επιφάνειας στο `binding.viewFinder.surfaceProvider`. Αυτό υποδεικνύει πού πρέπει να εμφανίζεται η προεπισκόπηση της κάμερας.
- Διαμόρφωση ImageAnalysis: `imageAnalyzer = ImageAnalysis.Builder() ...`: Αυτός ο κώδικας διαμορφώνει μια περίπτωση χρήσης ImageAnalysis με διάφορες ρυθμίσεις: Η στρατηγική αντίθλιψης ορίστηκε σε STRATEGY\_KEEP\_ONLY\_LATEST. Η ανάλυση στόχος ορίστηκε στα 768x1024 pixel. Η περιστροφή στόχου ορίστηκε στην υπολογισμένη περιστροφή της οθόνης. Ενεργοποιήθηκε η περιστροφή εικόνας εξόδου. Μορφή εικόνας εξόδου ορίστηκε σε RGBA 8888. Δημιουργεί την περίπτωση χρήσης ανάλυσης εικόνας. Ο αναλυτής εκχωρείται στο στιγμιότυπο, πράγμα που σημαίνει ότι η λογική ανάλυσης που ορίζεται αλλού στην κλάση θα εκτελεστεί όταν τα πλαίσια είναι διαθέσιμα για ανάλυση.
- `try { ... } catch (exc: Exception) { ... };` `try { ... } catch (εκτός: Εξαίρεση) { ... };`: Αυτή η ενότητα είναι υπεύθυνη για τη δέσμευση περιπτώσεων χρήσης στον πάροχο της κάμερας. Κάνει τα εξής: Αποσυνδέει τυχόν θήκες χρήσης κάμερας που είχαν δεσμευτεί προηγουμένως χρησιμοποιώντας την `cameraProvider.unbindAll()`. Αυτό το βήμα είναι απαραίτητο πριν επανασυνδεθούν οι θήκες χρήσης στην κάμερα. Συνδέει τις διαμορφωμένες περιπτώσεις χρήσης (προεπισκόπηση και ανάλυση εικόνας) στον πάροχο της κάμερας χρησιμοποιώντας το `cameraProvider.bindToLifecycle(...)`. `CameraSelector.DEFAULT_BACK_CAMERA` επιλέγει την προεπιλεγμένη κάμερα που βλέπει προς τα πίσω. Εάν υπάρχει εξαίρεση κατά τη διαδικασία δέσμευσης περιπτώσεων χρήσης, καταγράφει ένα μήνυμα σφάλματος χρησιμοποιώντας το `Log.e (TAG, "Η δέσμευση περίπτωσης χρήσης απέτυχε", exc)`.

## Συμπεράσματα

---

Συμπερασματικά, η χρήση μιας εφαρμογής Android με δυνατότητες μηχανικής μάθησης προσφέρει πολλά οφέλη και ευκαιρίες. Συνολικά, η ενσωμάτωση μηχανικής μάθησης σε εφαρμογές Android ανοίγει έναν κόσμο δυνατοτήτων, επιτρέποντας έξυπνες αλληλεπιδράσεις, αυτοματισμό, εξατομίκευση και λήψη αποφάσεων βάσει δεδομένων. Αξιοποιώντας τη δύναμη της μηχανικής μάθησης, οι εφαρμογές Android μπορούν να παρέχουν βελτιωμένες εμπειρίες χρήστη, να βελτιώσουν την αποτελεσματικότητα και να ξεκλειδώσουν νέα επίπεδα λειτουργικότητας και καινοτομίας.

## Βιβλιογραφία

---

- [1] Tan, P.N., Steinbach, M. and Kumar, V. (2006) *Introduction to Data Mining*. Pearson: Addison Wesley, Boston.
- [2] Bishop, Christopher M. *Pattern Recognition and Machine Learning*. Springer, 2016.
- [3] Oscar Deniz Suarez, et al. *OpenCV Essentials*. Packt Publishing Ltd, 25 Aug. 2014.
- [4] Singh, Anubhav, and Rimjhim Bhadani. *Mobile Deep Learning with TensorFlow Lite, ML Kit and Flutter*. Packt Publishing Ltd, 6 Apr. 2020.
- [5] “ML Kit.” *Google Developers*, developers.google.com/ml-kit.
- [6] “Recognize Text and Facial Features with ML Kit: Android.” *Google Codelabs*, codelabs.developers.google.com/codelabs/mlkit-android.
- [7] Marsicano, Kristin, et al. *Android Programming : The Big Nerd Ranch Guide*. Atlanta, Ga, Big Nerd Ranch, 2019.
- [8] Dougherty, James, et al. “Supervised and Unsupervised Discretization of Continuous Features.” *Machine Learning Proceedings 1995*, 1995, pp. 194–202, <https://doi.org/10.1016/b978-1-55860-377-6.50032-3>.
- [9] Koul, Anirudh, et al. *Practical Deep Learning for Cloud, Mobile, and Edge*. “O’Reilly Media, Inc.,” 14 Oct. 2019.
- [10] “11.2 - the Impurity Function | STAT 508.” *PennState: Statistics Online Courses*, online.stat.psu.edu/stat508/lesson/11/11.2.
- [11] Loaiza, Steven. “Gini Impurity Measure.” *Medium*, 23 Mar. 2020, [towardsdatascience.com/gini-impurity-measure-dbd3878ead33](https://towardsdatascience.com/gini-impurity-measure-dbd3878ead33).
- [12] Kumar, Narender. “Entropy in Machine Learning.” *Spark by {Examples}*, 6 Apr. 2023, [sparkbyexamples.com/machine-learning/entropy-in-machine-learning/](https://sparkbyexamples.com/machine-learning/entropy-in-machine-learning/).
- [13] Mehta, Sourabh. “What Is Pruning in Tree Based ML Models and Why Is It Done?” *Analytics India Magazine*, 6 July 2022, [analyticsindiamag.com/what-is-pruning-in-tree-based-ml-models-and-why-is-it-done/](https://analyticsindiamag.com/what-is-pruning-in-tree-based-ml-models-and-why-is-it-done/).
- [14] Shah, Tarang. “About Train, Validation and Test Sets in Machine Learning.”

*Towards Data Science*, Towards Data Science, 6 Dec. 2017,  
[towardsdatascience.com/train-validation-and-test-sets-72cb40cba9e7](https://towardsdatascience.com/train-validation-and-test-sets-72cb40cba9e7)

- [15] “Generalization | Machine Learning Crash Course.” *Google Developers*, [developers.google.com/machine-learning/crash-course/generalization/video-lecture](https://developers.google.com/machine-learning/crash-course/generalization/video-lecture).
- [16] Jayaswal, Vaibhav. “Laplace Smoothing in Naïve Bayes Algorithm.” *Medium*, 22 Nov. 2020, [towardsdatascience.com/laplace-smoothing-in-na%C3%AFve-bayes-algorithm-9c237a8bdece](https://towardsdatascience.com/laplace-smoothing-in-na%C3%AFve-bayes-algorithm-9c237a8bdece).
- [17] “What Is Gaussian Distribution.” *Deepchecks*, [deepchecks.com/glossary/gaussian-distribution/](https://deepchecks.com/glossary/gaussian-distribution/).
- [18] Gupta, Binay. “Binning in Machine Learning.” *Scaler Topics*, 13 Jan. 2023, [www.scaler.com/topics/machine-learning/binning-in-machine-learning/](https://www.scaler.com/topics/machine-learning/binning-in-machine-learning/).
- [19] bhadani, goutam. “Haar Cascade Classifier vs Histogram of Oriented Gradients(HOG).” *Medium*, 13 May 2020, [medium.com/@goutam0157/haar-cascade-classifier-vs-histogram-of-oriented-gradients-hog-6f4373ca239b](https://medium.com/@goutam0157/haar-cascade-classifier-vs-histogram-of-oriented-gradients-hog-6f4373ca239b).
- [20] “Predictive Maintenance Using Machine Learning | Implementations | AWS Solutions.” *Amazon Web Services, Inc.*, [aws.amazon.com/solutions/implementations/predictive-maintenance-using-machine-learning/](https://aws.amazon.com/solutions/implementations/predictive-maintenance-using-machine-learning/).
- [21] “TensorFlow 20 Frozen Graph Support | Saturn Cloud Blog.” *Saturncloud.io*, 13 June 2023, [saturncloud.io/blog/tensorflow-20-frozen-graph-support/](https://saturncloud.io/blog/tensorflow-20-frozen-graph-support/).