A logo of a university of patras

Description automatically generated

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ

 ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Τεχνητά Νευρωνικά Δίκτυα για την ανάλυση συναισθημάτων σε έντυπες κριτικές ή και σε κοινωνικά δίκτυα**

**ΑΡΙΣΤΕΙΔΗΣ ΜΠΟΛΙΑΣ**

**ΑΜ : 1069910**

**Επιβλέπων Καθηγητής : Οικονόμου Γεώργιος**

**Πάτρα 2025**

Copyright ©–All rights reserved Αριστείδης Μπολιάς, 2025.

Με την επιφύλαξη παντός δικαιώματος.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Το περιεχόμενο αυτής της εργασίας δεν απηχεί απαραίτητα τις απόψεις του Τμήματος, του Επιβλέποντα, ή της επιτροπής που την ενέκρινε.

Ευχαριστίες

Περίληψη

Abstract

Περιεχόμενα

[1 Μηχανική Μάθηση 7](#_Toc211288793)

[1.1 Τι είναι Μηχανική Μάθηση 7](#_Toc211288794)

# 1 Μηχανική Μάθηση

## Τι είναι Μηχανική Μάθηση

Η Μηχανική Μάθηση αποτελεί έναν από τους βασικούς κλάδους της Τεχνητής Νοημοσύνης και επικεντρώνεται στην δημιουργία συστημάτων που έχουν τη δυνατότητα να μαθαίνουν και να βελτιώνονται αυτόματα μέσα από την εμπειρία χωρίς να απαιτείται ρητός προγραμματισμός όλων των κανόνων. Στόχος της είναι η ανάπτυξη αλγορίθμων που μπορούν να αξιοποιούν δεδομένα, να αναγνωρίζουν πρότυπα, να προβλέπουν αποτελέσματα και να λαμβάνουν αποφάσεις.

Η έννοια της Μηχανικής Μάθηση δεν είναι νέα . Ο όρος εισήχθη πρώτη φορά από τον Arthur Samuel το 1959 , ο οποίος όρισε την μηχανική μάθηση ως :

Η Μηχανική Μάθηση είναι ο τομέας της μελέτης που δίνει στους υπολογιστές τη δυνατότητα να μαθαίνουν χωρίς να έχουν προγραμματιστεί ρητά.

O Arthur Samuel ανέπτυξε ένα πρόγραμμα για το παιχνίδι της ντάμας το οποίο βελτίωνε σταδιακά τις επιδόσεις του μέσα από την εμπειρία των παιχνιδιών. Το έργο του αποτέλεσε θεμέλιο για την κατανόηση της μάθησης μέσω υπολογιστών και δείχνει ξεκάθαρα την διαφορά ανάμεσα στον παραδοσιακό προγραμματισμό και την διαδικασία μάθησης από δεδομένα.

Σε αντίθεση με τον παραδοσιακό προγραμματισμό , όπου ο προγραμματιστής καθορίζει ρητά τους κανόνες λειτουργίας ενός συστήματος, η Μηχανική Μάθηση βασίζεται στην ιδέα ότι οι κανόνες μπορούν να εξαχθούν από δεδομένα. Ένα σύστημα μηχανικής μάθησης εκπαιδεύεται σε ένα σύνολο παραδειγμάτων (training data) και στη συνέχεια εφαρμόζει τη γνώση που απέκτησε για να κάνει προβλέψεις σε νέα άγνωστα δεδομένα (test data).

Με την πάροδο του χρόνου και την εξέλιξη της τεχνολογίας η Μηχανική Μάθηση εξελίχθηκε σε από τους πιο δυνατούς τομείς της Πληροφορικής, με εφαρμογές που ποικίλουν από την αναγνώριση φωνής και εικόνας μέχρι τις συστάσεις προϊόντων και τα αυτόνομα οχήματα. Η Μηχανική Μάθηση σήμερα αποτελεί το θεμέλιο πάνω στο οποίο χτίζεται η Τεχνητή Νοημοσύνη.

## **Είδη Μηχανικής Μάθησης**

Η Μηχανική Μάθηση διακρίνεται σε τρεις βασικές κατηγορίες ανάλογα με το αν τα δεδομένα εκπαίδευσης έχουν μια “βοηθητική” ετικέτα, με τον τρόπο με τον οποίο αλληλοεπιδρά με αυτά και με τον στόχο που επιδιώκεται. Οι κατηγορίες αυτές είναι η Επιβλεπόμενη Μάθηση (Supervised Learning), η Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning) και η Ενισχυτική Μάθηση (Reinforcement Learning).

### **1.2.1 Επιβλεπόμενη Μάθηση (Supervised Learning)**

H Επιβλεπόμενη Μάθηση είναι η πιο διαδεδομένη μορφή Μηχανικής Μάθησης. Κατά την διάρκεια της εκπαίδευσης ο αλγόριθμος του μοντέλου επεξεργάζεται μεγάλο όγκο δεδομένων με ετικέτες (labels) που δίνουν την απάντηση. Έπειτα η απόδοση του μοντέλου ελέγχεται με νέα δεδομένα (test data) για να κατανοήσουμε άμα έγινε σωστά η εκπαίδευση. Με λίγα λόγια η επιβλεπόμενη μάθηση είναι όταν υπάρχουν κάποιες μεταβλητές εισόδου (Χ) , μια μεταβλητή εξόδου (Υ) και ένας αλγόριθμος που εκπαιδεύεται ώστε να προβλέπει την έξοδο με βάση τις εισόδους:

f : X -> Y

Στόχος της Επιβλεπόμενης Μάθησης είναι εκπαιδεύσει ένα μοντέλο τόσο καλά ώστε όταν υπάρχουν νέα δεδομένα εισόδου (Χ) να μπορούν να προβλεφθούν οι μεταβλητές εξόδου (Y) για αυτά τα δεδομένα.

Η Επιβλεπόμενη Μάθηση έχει δυο βασικές κατηγορίες προβλημάτων, την ταξινόμηση(Classification) και την Παλινδρόμηση (Regression)

**Ταξινόμηση(Classification) :** Ένα πρόβλημα ταξινόμησης είναι όταν η μεταβλητή εξόδου είναι μια κατηγορία . Το ζητούμενο είναι το μοντέλο να την είσοδο σε μια από τις προκαθορισμένες κλάσεις. Για παράδειγμα ταξινόμηση email σε “spam” ή “μη spam”.

**Παλινδρόμηση(Regression) :** Ένα πρόβλημα παλινδρόμησης είναι όταν η μεταβλητή εξόδου είναι συνεχής ποσοτική τιμή, όπως “ευρώ” ή “βάρος”. Στόχος είναι το μοντέλο να προβλέψει μια αριθμητική τιμή όπως η τιμή ενός σπιτιού ή το βάρος ενός αντικειμένου.

### **1.2.2 Μη Επιβλεπόμενη Μάθηση**

Η Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning) είναι ένας από τους σημαντικότερους τομείς της Μηχανικής Μάθησης , όπου σε αντίθεση με την επιβλεπόμενη μάθηση δεν υπάρχουν διαθέσιμες ετικέτες ( labels) που να υποδεικνύουν τη σωστή απάντηση. Υπάρχουν μόνο δεδομένα εισόδου (X) και καμία αντίστοιχη μεταβλητή εξόδου. Ο στόχος του αλγορίθμου είναι να εντοπίσει κρυμμένα μοτίβα, ομάδες ή και συσχετίσεις μέσα στα δεδομένα χωρίς καμία καθοδήγηση.

Υπάρχουν τρία βασικά είδη μη επιβλεπόμενης μάθησης :η ομαδοποίηση(clustering), οι κανόνες συσχέτισης (Association Rules) και η ανίχνευση ανωμαλιών (Anomaly Detection).

**Ομαδοποίηση (Clustering) :** Η πιο γνωστή κατηγορία μην επιβλεπόμενης μάθησης , στόχος της είναι η δημιουργία ομάδων (clusters) δεδομένων με βάσει τις ομοιότητες τους. Τα δεδομένα που ανήκουν στο ίδιο cluster μοιράζονται κοινά χαρακτηριστικά , ενώ διαφέρουν από τα δεδομένα άλλων clusters.Υπάρχουν διάφοροι τύποι αλγορίθμων ομαδοποίησης :

* **K-means Clustering :** Ένας από τους πιο γνωστούς αλγόριθμους που βασίζεται στην εύρεση Κ κέντρων και στην ελαχιστοποίηση της απόστασης κάθε σημείου από το πλησιέστερο κέντρο.
* **Ιεραρχική Ομαδοποίηση (Hierarchical Clustering) :** Τα δεδομένα οργανώνονται σε μια ιεραρχία από clusters, δηλαδή αντί να καθορίζουμε εκ των προτέρων πόσες ομάδες θέλουμε , η μέθοδος δημιουργεί ένα δέντρο που δείχνει πως ενώνονται ή χωρίζονται τα δεδομένα σε διαφορετικά επίπεδα ομοιότητας.

**Κανόνες συσχέτισης (Association Rules) :** Οι κανόνες συσχέτισης αποσκοπούν στην ανίχνευση σχέσεων μεταξύ διαφορετικών μεταβλητών μέσα σε μεγάλα σύνολα δεδομένων. Χαρακτηριστικοί αλγόριθμοι είναι οι εξής :

* **Apriori :** Βασίζεται στην ιδέα ότι αν ένα σύνολο στοιχείων είναι συχνό, τότε και όλα τα υποσύνολα του θα είναι επίσης συχνά. Λειτουργεί επαναληπτικά βρίσκει μοτίβα σε δεδομένα και στην συνέχεια παράγει κανόνες της μορφής A -> B.
* **FP- Growth :** Είναι μια πιο αποδοτική εκδοχή του Apriori . Χρησιμοποιεί μια δομή δέντρου (FP-tree) για να αποθηκεύσει τα δεδομένα και να βρει πιο γρήγορα συχνά μοτίβα χωρίς να παράγει όλους τους πιθανούς συνδυασμούς όπως κάνει το Apriori.

**Ανίχνευση Ανωμαλιών (Anomaly Detection) :** Η ανίχνευση ανωμαλιών στοχεύει στον εντοπισμό δεδομένων που αποκλίνουν πολύ από το κανονικό μοτίβο. Αυτές οι ανωμαλίες μπορεί να υποδηλώνουν σφάλματα ή ξαφνικές αλλαγές. Υπάρχουν τρεις βασικές κατηγορίες ανίχνευσης ανωμαλιών :

* **Μέθοδοι βασισμένες σε στατικά μοντέλα :** Ελέγχεται αν μια παρατήρηση αποκλίνει από την κανονική συμπεριφορά.
* **Μέθοδοι βασισμένες σε αποστάσεις :** Παρατηρήσεις που απέχουν πολύ από τα υπόλοιπα δεδομένα σε ένα πολυδιάστατο χώρο χαρακτηρίζονται ως ανώμαλες.
* **Μέθοδοι Μηχανικής Μάθησης :** Χρησιμοποιούνται μοντέλα που μαθαίνουν την κανονική συμπεριφορά και εντοπίζουν ότι αποκλείνει.