

Thesis Title

Institution Name



Νησιωτη Ελένη

Day Month Year

## 0.1 Μηχανική Μάθηση

### 0.1.1 Η έννοια της μηχανικής μάθησης

Μια έννοια που χαρακτηρίζεται από πολύ υψηλό λόγο χρησιμότητας προς κατανοησιμότητας είναι αυτή της μηχανικής μάθησης. Χάρης στην εμπορική της αξία και την ευρεία εφαρμοστικότητα, η μηχανική μάθηση υπάρχει παντού: από αβέβαιες οικονομικές αγορές και βιολογικά εργαστήρια με τεράστιους όγκους δεδομένων μέχρι ιστοσελίδες όπως το facebook και το netflix. Ας προσπαθήσουμε να κατανοήσουμε όμως αυτήν την επιστήμη πέρα από το κλισέ της περιγραφής "μηχανές που σκέφτονται όπως οι άνθρωποι".

Η μηχανική μάθηση προέκυψε από τον επιστημονικό τομέα της Τεχνητής Νοημοσύνης, που όπως δηλώνει και το όνομά της, μελετά τη νοημοσύνη που έχει τη δυνατότητα να επιδειξεί μία μηχανή, είτε ως hardware είτε ως software. Στο paper Υπολογιστική Μηχανική και Νοημοσύνη (Computing Machinery and Intelligence) ο Allan Turing αναρωτιέται πόσο ασαφείς είναι οι λέξεις "σκέφτομαι" και "μηχανή" και μέσω ενός συλλογισμού, που αποκάλεσε "Το παιχνίδι της μίμησης", κατέληξε στη διατύπωση του ερωτήματος ως εξής: "Μπορούν οι μηχανές να κάνουν αυτό που εμείς, ως σκεπτόμενα όντα, κάνουμε;". Αυτή τη συλλογιστική ακολούθησε ο Tom M. Mitchel στον φορμαλιστικό και όχι τόσο διδακτικό ορισμό του για τη μηχανική μάθηση:

*"Λέμε πως ένα πρόγραμμα υπολογιστή μαθαίνει από μια εμπειρία  $E$ , αναφερόμενοι σε ένα σύνολο καθηκόντων  $T$  και ένα μέτρο απόδοσης  $P$ , αν η απόδοση του στα καθήκοντα  $T$ , όπως μετράται από το  $P$ , βελτιώνεται καθώς αποκτά εμπειρία  $E$ ."*

Βασικό αντικείμενο της μηχανικής μάθησης είναι η μελέτη και ο σχεδιασμός αλγορίθμων που μπορούν να "μάθουν" από τα δεδομένα και να κάνουν προβλέψεις για μελλοντικές καταστάσεις. Είναι λογικό λοιπόν να κληρονομεί στοιχεία από την επιστήμη υπολογιστών, για το σχεδιασμό των αλγορίθμων, καθώς και τη στατιστική, για την ανάλυση των δεδομένων. Αποτελεί εξέλιξη της Αναγνώρισης Προτύπων, που εξερευνά μεγάλους όγκους δεδομένων προς εύρεση μοτίβων, καθώς και της Θεωρίας Υπολογιστικής Μάθησης, η οποία μελετά αλγορίθμους. Και οι δυο αυτές επιστήμες έχουν πλέον ενσωματωθεί στην έννοια της μηχανικής μάθησης.

Τα προβλήματα με τα οποία ασχολείται η μηχανική μάθηση παίρνουν διάφορες μορφές, ανάλογα με το είδος της πρόβλεψης που επιτελούν:

- *Προβλήματα Παλινδρόμησης.* Πρόκειται για προβλήματα πρόβλεψης μιας συνεχούς τιμής, για παράδειγμα της τιμής πώλησης οικοπέδων.
- *Προβλήματα Ταξινόμησης.* Εδώ ενδιαφερόμαστε να αναγνωρίσουμε την κατηγορία στην οποία ανήκει ένα δεδομένο. Για παράδειγμα ένα εργοστάσιο ενδιαφέρεται για την πρόβλεψη ελαττωματικών εξαρτημάτων σύμφωνα με τα χαρακτηριστικά τους.
- *Προβλήματα Ομαδοποίησης.* Κι εδώ προβλέπουμε την κατηγορία στην οποία ανήκει μία είσοδος, χωρίς όμως εκ των προτέρων γνώση για τις κατηγορίες. Μία τράπεζα για παράδειγμα που θέλει να αναγνωρίσει τα είδη των πελατών της, θα αναγνωρίσει ομάδες με βάση την κοινή τους συμπεριφορά και στη συνέχεια θα αναλύσει τη φύση των ομάδων που ανέκυψαν.

Η ουσία της μηχανικής μάθησης βρίσκεται στο εξής ερώτημα: "Πότε μπορώ να την εφαρμόσω, δεδομένου ενός προβλήματος" Οι προϋποθέσεις που οφείλω να ελέγξω ότι πληρούνται είναι οι εξής:

- *Υπάρχει κάποιο μοτίβο στα δεδομένα.* Σε περιπτώσεις που το πρόβλημα ζει σε κάποιο άναρχο, τυχαίο χώρο, η μηχανική μάθηση είναι καταδικασμένη να αποτύχει, όπως και κάθε νοημοσύνη.
- *Το πρόβλημα δεν χαρακτηρίζεται από κάποια μαθηματική εξίσωση.* Σε αυτή την περίπτωση η μηχανική μάθηση θα σου δώσει τη σωστή λύση, την οποία όμως θα μπορούσες να είχες βρει πολύ πιο εύκολα. Αν για παράδειγμα συλλέξεις δεδομένα και ζητήσεις από

κάποιο αλγόριθμο μηχανικής μάθησης να αποφασίσει αν σχηματίζουν κύκλο, τότε, με όρους μηχανικού, έχεις μια μη αποδοτική λύση.

- *Υπάρχουν δεδομένα σχετικά με το πρόβλημα.* Ειδικά και με βάση τον ορισμό που δώσαμε, η μηχανική μάθηση δεν μπορεί να εφαρμοστεί.

Μπορούμε να αναλύσουμε τα είδη των προβλημάτων που απασχολούν αυτήν την επιστήμη με βάση την πληροφορία που προσφέρεται στον αλγόριθμο:

- *Επιβλεπόμενη μάθηση.* Στον αλγόριθμο δίνονται δεδομένα, που περιέχουν τόσο τα χαρακτηριστικά που έχουμε διαθέσιμα, όσο και την πρόβλεψη για αυτά. Για παράδειγμα, αν θέλουμε να προβλέψουμε την τιμή κατοικιών σε μια περιοχή, θα προσφέραμε στον αλγόριθμο χαρακτηριστικά όπως η τοποθεσία, τα τετραγωνικά μέτρα και η τιμή κάποιων κατοικιών, θα παίρναμε το μοντέλο και θα το χρησιμοποιούσαμε για να προβλέψουμε την τιμή άλλων κατοικιών με βάση την τοποθεσία τους και το μέγεθός τους.
- *Μη Επιβλεπόμενη μάθηση.* Εδώ, δε δίνεται πληροφορία στον αλγόριθμο σχετικά με το χαρακτηριστικό που θέλουμε να προβλέψουμε. Αντιθέτως, το ζητούμενο είναι η μελέτη των δεδομένων ώστε να ανακαλυφτούν δομικά πρότυπα σε αυτά. Αν λοιπόν η Amazon θέλει να αναγνωρίσει τους τύπους των πελατών της, ώστε να μεγιστοποιήσει το κέρδος της προβάλλοντας σε κάθε τύπο διαφορετικές, προσαρμοσμένες διαφημίσεις, θα χρειαστεί ένα τέτοιο αλγόριθμο, που με βάση τις αγορές, τις συνήθειες, την καταγωγή και άλλα χαρακτηριστικά των πελατών θα τους ομαδοποιήσει σε ομοιογενείς ομάδες.
- *Ενισχυτική μάθηση.* Ούτε εδώ έχουμε πληροφορία για το χαρακτηριστικό που προβλέπουμε, ωστόσο υπάρχει η έννοια της ανταμοιβής: ο αλγόριθμος δρα σε κάποιο δυναμικό περιβάλλον, όπου κάνει κάποιες αβέβαιες επιλογές-προβλέψεις, για τις οποίες θα ανταμειφθεί ή τιμωρηθεί εκ των υστέρων. Όπως ακριβώς ένα νεογέννητο μωρό χρειάζεται να ακουμπήσει μερικές φορές κάτι καυτό για να μάθει ότι δεν πρέπει να το ξανακάνει, έτσι και ένας πράκτορας λογισμικού χρειάζεται να παίξει διάφορες κινήσεις στο σκάκι για να μπορεί να κερδίζει.

Στην παρούσα εργασία θα εστιάσουμε στην επιβλεπόμενη μάθηση σε προβλήματα ταξινόμησης, καθώς είναι η πιο ευρέως χρησιμοποιούμενη και μελετημένη κατηγορία μηχανικής μάθησης.

Ας ορίσουμε κάποιες έννοιες που χρησιμοποιούνται συχνά στη βιβλιογραφία. Έστω το πρόβλημα πρόβλεψης της κακοήθειας ενός όγκου, με βάση την ηλικία του και το μέγεθός του. Έχουμε συλλέξει δεδομένα και έχουμε απεικονίσει με κόκκινο αυτά που αντιστοιχούν σε κακοήγη και με μπλε σε καλοήγη όγκο. Η γραμμή που βλέπουμε αποτελεί την έξοδο ενός μοντέλου ταξινόμησης, μπορούμε δηλαδή με βάση αυτήν να προβλέψουμε το είδος ενός νέου όγκου.

- *Χαρακτηριστικά  $x_n$ .* Τα στοιχεία που χαρακτηρίζουν τα δεδομένα ενός προβλήματος. Στο πρόβλημά μας κάποια χαρακτηριστικά μπορεί να είναι το μέγεθός του, το χρώμα του και η δομή του.
- *Κλάση  $y_n$ .* Πρόκειται για το στοιχείο που θέλουμε να προβλέψουμε, δηλαδή στην παραπάνω περίπτωση την κακοήθεια του όγκου.
- *Παραδείγματα  $(x_n, y_n)$ .* Τα δεδομένα του προβλήματος δίνονται συνήθως σε μορφή πίνακα: κάθε γραμμή είναι ένα παράδειγμα και οι στήλες περιέχουν τα χαρακτηριστικά και την κλάση.
- *Συνάρτηση-στόχος  $f$ .* Είναι η άγνωστη συνάρτηση που προσπαθούμε μέσω εκπαίδευσης να βρούμε. Ορίζεται ως  $f : X \rightarrow Y$ , στο συγκεκριμένο πρόβλημα δηλαδή καθορίζει πως προκύπτει η κακοήθεια ενός όγκου από τα χαρακτηριστικά του. Η συνάρτηση αυτή ζει στον πραγματικό κόσμο, επομένως αναμένουμε να είναι περίπλοκη, γεγονός που σε συνδυασμό με το πεπερασμένο πλήθος δεδομένων από το συγκεκριμένο χώρο περιορίζει τις προσδοκίες μας: δεν προσπαθούμε να την ορίσουμε ακριβώς, αλλά να την προσεγγίσουμε.

- *Υπόθεση h*. Το αποτέλεσμα της εκπαίδευσης, συνήθως μια ευθεία στο χώρο των χαρακτηριστικών των δεδομένων που τον χωρίζει σε 2 υποχώρους: νέα χαρακτηριστικά που φθάνουν κατατάσσονται ως κακοήγη ή καλοήγη ανάλογα με την υπόθεση. Η υπόθεση μπορεί να έχει οποιαδήποτε μορφή, ώστε να διαχωρίσει τα δεδομένα, όπως προσεγγιστικά θα έκανε και η άγνωστη συνάρτηση-στόχος.
- *Μοντέλο Εκπαίδευσης*. Προκειμένου να κάνουμε προβλέψεις σε άγνωστα δεδομένα, χρειαζόμαστε ένα μοντέλο, μία μαθηματική διαδικασία, η οποία έχει εκπαιδευθεί και παραμετροποιηθεί πάνω στο συγκεκριμένο πρόβλημα και λαμβάνοντας τα χαρακτηριστικά ενός νέου δεδομένου μπορεί να δώσει την κλάση του. Το μοντέλο αποτελείται από δύο συστατικά:
  - *Σετ υπόθεσης*  $H = \{h\}$ . Κάθε μοντέλο προσπαθεί να προσεγγίσει τη συνάρτηση-στόχο με διαφορετικό τρόπο. Το σετ υπόθεσης περιέχει όλες τις πιθανές υποθέσεις που μπορούν να γίνουν. Κάθε διαφορετική υπόθεση αντιστοιχεί σε διαφορετική ρύθμιση κάποιας παραμέτρου του μοντέλου και επιτελεί διαφορετική πρόβλεψη για τα δεδομένα. Διάφορα είδη μοντέλων πρόβλεψης είναι οι perceptrons, τα νευρωνικά δίκτυα, οι μηχανές διανυσματικής στήριξης, ο ταξινομητής bayes...
  - *Αλγόριθμος μάθησης*. Ανάλογα με το μοντέλο που έχουμε επιλέξει, υπάρχει μια γκάμα αλγορίθμων, οι οποίοι επιτελούν τη διαδικασία της μάθησης, προσπαθώντας να βελτιστοποιήσουν τις παραμέτρους της υπόθεσης. Για παράδειγμα οι perceptrons χρησιμοποιούν τον αλγόριθμο PLA, τα νευρωνικά τον αλγόριθμο backpropagation...