

Κεφάλαιο 1: Εισαγωγή στη Θεωρία Αποφάσεων

1.1 Εισαγωγή

Η ευκολία με την οποία αναγνωρίζουμε ένα πρόσωπο, αντιλαμβανόμαστε τον προφορικό λόγο, διαβάζουμε χειρόγραφα κείμενα, αναγνωρίζουμε με την αφή τα κλειδιά του αυτοκινήτου στην τσέπη μας και αποφασίζουμε αν ένα μήλο έχει σαπίσει από τη μυρωδιά του, περιγράφεται από πολύπλοκες διαδικασίες που ανήκουν και περιγράφονται από αυτό που ονομάζεται Αναγνώριση Προτύπων. Η Αναγνώριση Προτύπων - η διαδικασία κατά την οποία παρατηρούμε απλά δεδομένα και λαμβάνουμε αποφάσεις βασισμένες στην «κατηγορία» του προτύπου- έχει ζωτική σημασία ακόμη και για την επιβίωση μας. Άλλωστε ως ανθρώπινα όντα έχουμε διανύσει πάνω από δέκα εκατομμύρια χρόνια εξελίσσοντας υψηλού επιπέδου νευρικά και νοητικά συστήματα για την επιτυχή ολοκλήρωση τέτοιου είδους διαδικασιών.

1.2 Αντίληψη Μηχανής

Είναι φυσικό να θέλουμε να σχεδιάζουμε και να «χτίζουμε» μηχανές που αναγνωρίζουν πρότυπα. Η αυτόματη αναγνώριση φωνής, η αναγνώριση μέσω δαχτυλικών αποτυπωμάτων, η οπτική αναγνώριση χαρακτήρων κειμένου, η αναγνώριση ακολουθιών DNA και τόσα άλλα καθιστούν σαφές το γεγονός ότι η αξιόπιστη και ακριβής αναγνώριση προτύπων από μια μηχανή είναι πάρα πολύ χρήσιμη. Επιπλέον, από την επίλυση των διαφόρων προβλημάτων που προκύπτουν κατά την υλοποίηση τέτοιων συστημάτων, αποκτούμε βαθύτερη κατανόηση και εκτίμηση για τα συστήματα αναγνώρισης προτύπων στο φυσικό κόσμο- και πολύ περισσότερο στον άνθρωπο. Για κάποια προβλήματα όπως είναι η ομιλία και η οπτική αναγνώριση, οι προσπάθειες σχεδιασμού μπορεί στην πραγματικότητα να εμπνευστούν από τη γνώση του πώς αυτά επιλύονται-λειτουργούν στη φύση, τόσο ως προς τους αλγόριθμους σχεδιασμού (π.χ. οι Γενετικοί Αλγόριθμοι μιμούνται τη διαδικασία της φυσικής εξέλιξης), όσο και ως προς το σχεδιασμό υλικού (hardware) ειδικού σκοπού.

1.3 Ένα Παράδειγμα

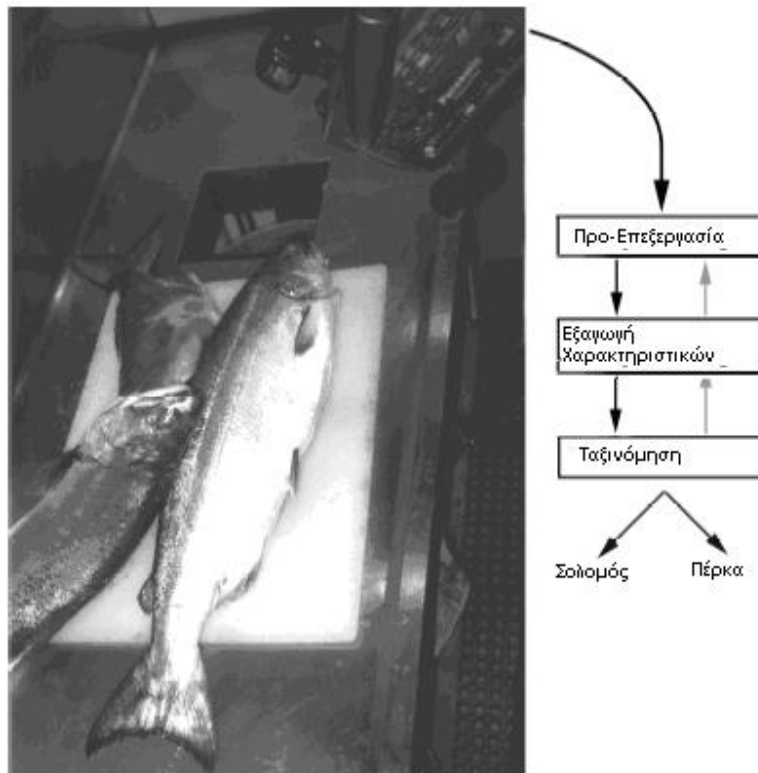
Για να γίνει κατανοητή η πολυπλοκότητα ορισμένων τύπων προβλημάτων με τα οποία θα ασχοληθούμε σε αυτό το βιβλίο, ας θεωρήσουμε το ακόλουθο φανταστικό παράδειγμα. Υποθέστε ότι ένα εργοστάσιο πακεταρίσματος επιθυμεί να αυτοματοποιήσει τη διαδικασία ταξινόμησης των εισερχομένων ψαριών επάνω σε έναν μάντα μεταφοράς, ανάλογα με το είδος του ψαριού.

Σαν πιλοτική εφαρμογή θα θεωρήσουμε ότι θέλουμε να διαχωρίζονται μέσω οπτικών αισθητήρων τα ψάρια «σολομός» και «πέρκα». Στήνουμε μια κάμερα, παίρνουμε κάποιες εικόνες δειγματοληπτικά και αρχίζουμε να παρατηρούμε κάποιες φυσικές διαφορές μεταξύ των δύο αυτών ειδών ψαριού. Αυτές μεταξύ άλλων μπορεί να είναι το μήκος, ο χρωματισμός, το πλάτος, ο αριθμός και το σχήμα των λεπιών τους, η θέση του στόματος κ.ά., οι οποίες συνιστούν τα χαρακτηριστικά (features) που θα χρησιμοποιηθούν στον ταξινομητή μας. Επίσης, παρατηρείται η ύπαρξη ανεπιθύμητου θορύβου και κάποιων διαταραχών στις εικόνες – π.χ. διαφορές στο χρωματισμό, στη θέση των ψαριών στον

ιμάντα μεταφοράς- που μπορεί να οφείλονται ακόμη και στο θόρυβο που προκαλεί η λειτουργία της ίδιας της ηλεκτρονικής κάμερας.

Δοθέντος του γεγονότος ότι υπάρχουν στην πραγματικότητα διαφορές μεταξύ των πληθυσμών της πέρκας και του σολομού μπορούμε να πούμε για αυτούς τους πληθυσμούς ότι έχουν διαφορετικά μοντέλα (models) – διαφορετική περιγραφή με βάση κάποιο μαθηματικό τύπο. Ο βασικός σκοπός και ο τρόπος προσέγγισης στα προβλήματα αναγνώρισης προτύπων είναι πρώτα να υποθέσει κανείς τις διάφορες κλάσεις αυτών των μοντέλων, στη συνέχεια να επεξεργαστεί τα δεδομένα που λαμβάνει από τους αισθητήρες, ώστε αυτά να μην περιέχουν θόρυβο, και τέλος για κάθε πρότυπο που αντιλαμβάνεται να το αντιστοιχεί στο μοντέλο που το περιγράφει καλύτερα. Κάθε τεχνική που βοηθάει και συντελεί στην ακόμα καλύτερη ολοκλήρωση του τελευταίου μέρους θα πρέπει να ληφθεί σοβαρά υπόψη από το σχεδιαστή συστημάτων αναγνώρισης προτύπων.

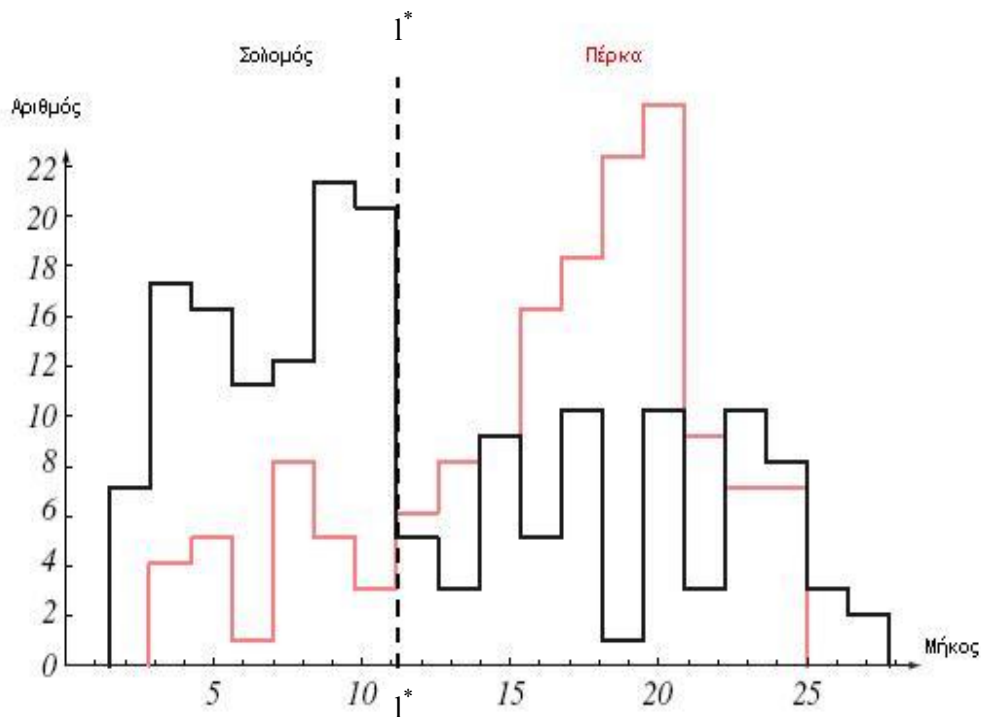
Το πρωτότυπο σύστημα, που επιτελεί τα παραπάνω, μπορεί να είναι όπως αυτό που φαίνεται στην εικόνα 1.1. Αρχικά η κάμερα λαμβάνει μια εικόνα από το ψάρι. Έπειτα, τα σήματα της κάμερας προεπεξεργάζονται (preprocessed) για να απλοποιηθούν οι επόμενες πράξεις χωρίς να χαθούν όμως ζωτικές πληροφορίες. Συγκεκριμένα μπορούμε να χρησιμοποιήσουμε την τμηματοποίηση (segmentation) κατά την οποία εικόνες από διαφορετικά ψάρια κατά κάποιο τρόπο απομονώνονται η μια από την άλλη και από το φόντο (background). Η πληροφορία από ένα μόνο ψάρι στέλνεται τότε σε ένα εξαγωγέα χαρακτηριστικών (feature extractor), του οποίου η λειτουργία επικεντρώνεται στο να μειώσει τον όγκο των δεδομένων, λαμβάνοντας υπόψη του μόνο κάποια σημαντικά «χαρακτηριστικά» ή «ιδιότητες».



Εικόνα 1.1

Αυτά τα χαρακτηριστικά (ή πιο συγκεκριμένα, οι τιμές αυτών των χαρακτηριστικών) περνάνε στη συνέχεια σε ένα ταξινομητή που αποτιμά τις ενδείξεις που του παρουσιάζουμε και παίρνει την τελική απόφαση όσον αφορά το είδος του ψαριού.

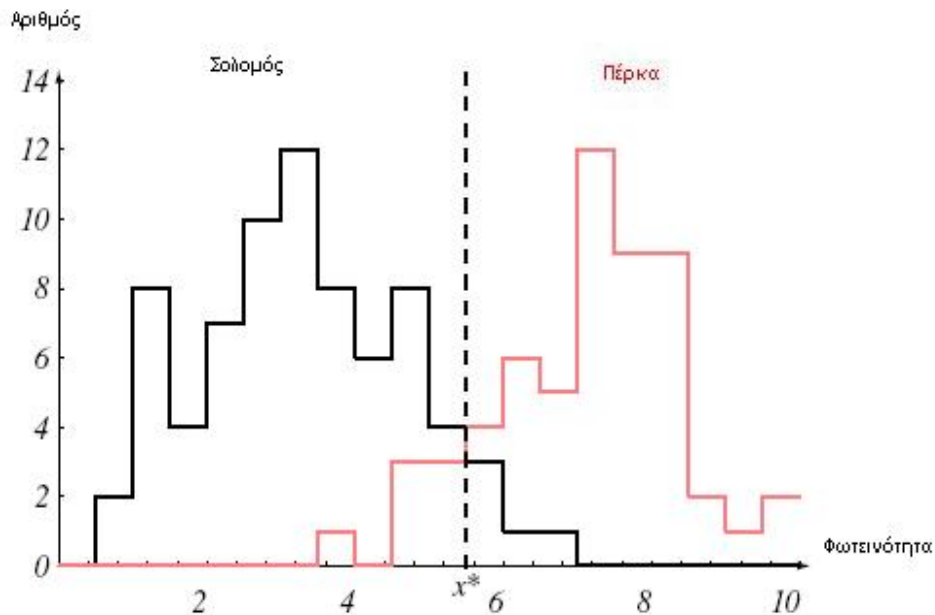
Ο προεπεξεργαστής μπορεί αυτομάτως να προσαρμόσει το μέσο όρο του επιπέδου της φωτεινότητας ή το κατώφλι της εικόνας, ώστε να απομακρύνει το φόντο του μάντα κ.ο.κ. Προς στιγμήν, ας μας επιτραπεί να αφήσουμε για λίγο στην άκρη το θέμα του πώς μπορεί μία εικόνα ψαριού να τμηματοποιηθεί και ας ασχοληθούμε με το σχεδιασμό του ταξινομητή και του εξαγωγέα χαρακτηριστικών. Ας υποθέσουμε ότι ένας εργάτης που δουλεύει στο τμήμα διαχωρισμού των ψαριών μας αποκαλύψει ότι γενικά η πέρκα είναι μεγαλύτερη σε μήκος από το σολομό. Αυτό μας δίνει κάποια δοκιμαστικά μοντέλα για τα ψάρια: Οι πέγκες έχουν κάποιο τυπικό μήκος και αυτό είναι μεγαλύτερο από το μήκος των σολομών. Το μήκος λοιπόν είναι ένα προφανές χαρακτηριστικό και μπορούμε να προσπαθήσουμε να ταξινομήσουμε ένα ψάρι ανάλογα με το αν το μήκος του I ξεπερνάει κάποιο κρίσιμο μήκος I^* ή όχι. Για να επιλέξουμε το κατώφλι I^* θα μπορούσαμε να πάρουμε κάποια δείγματα εκπαίδευσης ή σχεδίασης από ψάρια των δύο διαφορετικών ειδών, να κάνουμε μετρήσεις ως προς το μήκος και να επιθεωρήσουμε τα αποτελέσματα.



Εικόνα 1.2

Ας υποθέσουμε ότι κάνουμε τα παραπάνω και παίρνουμε τα ιστογράμματα της εικόνας 1.2 Αυτά τα «απογοητευτικά» ιστογράμματα δείχνουν όντως ότι η πέρκα είναι κάπως πιο μεγάλη σε μήκος από το σολομό, κατά μέσο όρο, αλλά είναι ξεκάθαρο ότι μόνο αυτό το κριτήριο δεν είναι επαρκές. Το συμπέρασμα είναι ότι: «Ανεξάρτητα από το πως (πόσο) επιλέγουμε το I^* , δεν μπορούμε αξιόπιστα να ξεχωρίσουμε την πέρκα από το σολομό μόνο βάση του μήκους».

Προχωράμε έτσι στην επιλογή άλλου χαρακτηριστικού. Αυτό είναι ο μέσος όρος φωτεινότητας του χρώματος του ψαριού. Τώρα πρέπει να είμαστε πολύ προσεκτικοί ώστε να εξαλείψουμε τις αποκλίσεις της φωτεινότητας, επειδή μπορούν να μπερδέψουν το μοντέλο μας αλλά και να δυσκολέψουν το έργο του ταξινομητή. Τα αντίστοιχα ιστογράμματα και η κρίσιμη τιμή x^* , φαίνονται στην εικόνα 1.3 όπου και είναι προφανές ότι είναι πιο ικανοποιητικά από την περίπτωση της προηγούμενης παραγράφου. Οι κλάσεις διαχωρίζονται καλύτερα.



Εικόνα 1.3

Μέχρις εδώ έχουμε υποθέσει κάπως ακίνδυνα ότι οι συνέπειες των πράξεων μας (εύρεση μοντέλου-ταξινόμηση) έχουν το ίδιο κόστος: Αποφασίζοντας πέρκα αντί για το σωστό (σολομός) έχει το ίδιο κόστος με το αντίστροφο (αποφασίζω σολομό αντί για το σωστό (πέρκα)). Αυτή η συμμετρία στο κόστος λανθασμένης απόφασης είναι συχνά αληθής, αλλά υπάρχουν φυσικά και πολλές εξαιρέσεις. (Φανταστείτε το παράδειγμα ενός ταξινομητή για νάρκες ενεργές ή μη...). Στο παράδειγμά μας, μια εταιρεία κονσερβοποίησης ψαριών μπορεί να γνωρίζει ότι οι πελάτες της δέχονται περιστασιακά κομμάτια σολομού στις κονσέρβες τους με τίτλο «πέρκα», όμως δυσανασχετούν όταν κομμάτια πέρκας βρεθούν στην κονσέρβα τους με όνομα «σολομός». Αν λοιπόν θέλει η παραπάνω εταιρία να παραμείνει ανταγωνιστική, θα πρέπει να προσαρμόσει τις αποφάσεις της, ώστε να ευχαριστεί τους πελάτες της, ακόμα και αυτό σημαίνει ότι θα αυξηθούν τα κομμάτια σολομού που θα μπαίνουν ενίοτε σε κονσέρβες με τίτλο «πέρκα». Σε αυτήν την περίπτωση λοιπόν θα πρέπει να μεταφερθεί το όριο απόφασης σε χαμηλότερες τιμές φωτεινότητας με αποτέλεσμα να μειωθεί ο αριθμός των ψαριών πέρκας που ταξινομούνται ως σολομοί. (εικόνα 1.3). Όσο περισσότερο διαμαρτύρονται οι πελάτες ότι παίρνουν πέρκα σε κονσέρβες που έχουν τον τίτλο «σολομός» (δηλαδή όσο μεγαλύτερο γίνεται το κόστος του λάθους ταξινόμησης) τόσο χαμηλότερα θα πρέπει να τίθεται το κατώφλι απόφασης x^* στην εικόνα 1.3.

Τέτοιες θεωρήσεις προτείνουν ότι υπάρχει ένα συνολικά μοναδικό κόστος που σχετίζεται με την απόφασή μας και η δουλειά μας στην πραγματικότητα είναι να δημιουργήσουμε ένα κανόνα απόφασης (για παράδειγμα να θέσουμε ένα όριο απόφασης) τέτοιο ώστε να ελαχιστοποιήσουμε το κόστος αυτό. Αυτή είναι η πρόκληση με την οποία ασχολείται η Θεωρία Αποφάσεων της οποίας η ταξινόμηση προτύπων είναι ίσως η σημαντικότερη υπο-περιοχή.

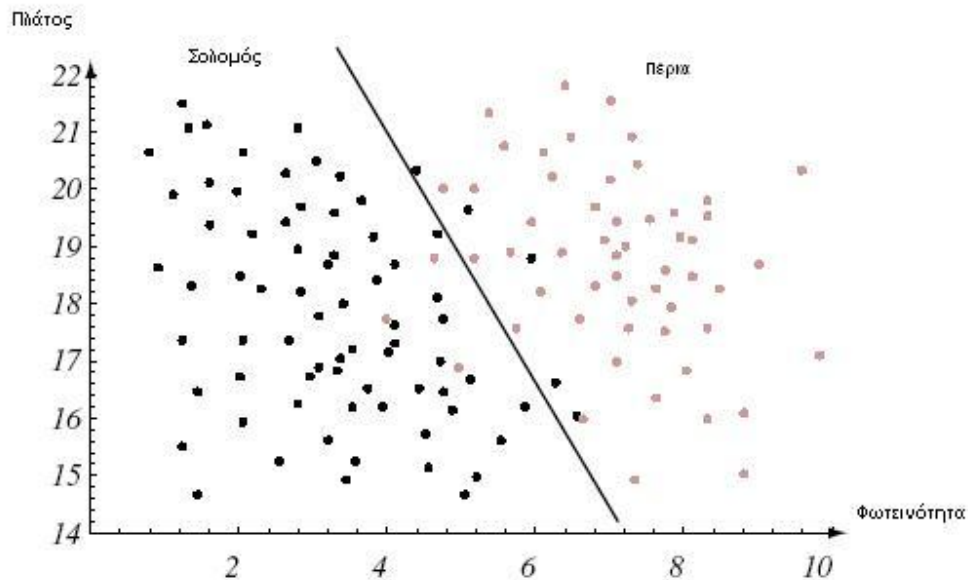
Ακόμη και αν γνωρίζουμε το κόστος το οποίο συνδέεται με τις αποφάσεις μας και διαλέξουμε το κρίσιμο σημείο x^* , μπορεί να απογοητευτούμε με το ανακριβές αποτέλεσμα που θα προκύψει. Η πρώτη μας προσπάθεια τότε είναι να ψάξουμε για κάποιο άλλο διαφορετικό χαρακτηριστικό πάνω στο οποίο θα στηριχτούμε για να διαχωρίσουμε τα ψάρια. Ας υποθέσουμε όμως ότι κανένα άλλο οπτικό χαρακτηριστικό από μόνο του δεν λειτουργεί καλύτερα από τη φωτεινότητα. Για να βελτιώσουμε την αναγνώριση, τότε πρέπει να καταφύγουμε στη χρήση περισσότερων του ενός χαρακτηριστικών.

Κατά την αναζήτηση αυτών των άλλων χαρακτηριστικών μπορεί να δοκιμάσουμε να στηριχθούμε στο γεγονός ότι η πέρκα είναι γενικά πιο μεγάλη σε πλάτος από το σολομό. Τώρα έχουμε δύο χαρακτηριστικά για να ταξινομήσουμε το κάθε ψάρι – την φωτεινότητα x_1 και το πλάτος x_2 . Αν αγνοήσουμε τον τρόπο με τον οποίο αυτά τα χαρακτηριστικά μετρώνται πρακτικά, διαπιστώνουμε ότι ο εξαγωγέας χαρακτηριστικών έχει «μειώσει» την εικόνα για κάθε ψάρι σε ένα σημείο ή αλλιώς σε ένα διάνυσμα χαρακτηριστικών x , σε ένα δύο διαστάσεων χώρο χαρακτηριστικών, όπου:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Το πρόβλημα μας στη συνέχεια, έγκειται στο να διαχωρίσουμε αυτό το χώρο χαρακτηριστικών σε δύο περιοχές. Αυτές οι περιοχές θα είναι διαχωρισμένες έτσι ώστε για όλα τα σημεία εντός της μίας περιοχής να ονομάζουμε το ψάρι «πέρκα» και εντός της άλλης περιοχής να το ονομάζουμε «σολομό». Ας υποθέσουμε ότι μετράμε τα διανύσματα των χαρακτηριστικών για τα δείγματα μας και ότι παίρνουμε τα σημεία που φαίνονται στην εικόνα 1.4. Αυτό το σχέδιο προτείνει τον ακόλουθο κανόνα για να διαχωρίσουμε τα ψάρια: Ταξινομούμε τα ψάρια ως «πέρκα» αν το διάνυσμα των χαρακτηριστικών του είναι πάνω από το όριο απόφασης που φαίνεται στην εικόνα 1.4 και ως «σολομό» διαφορετικά.

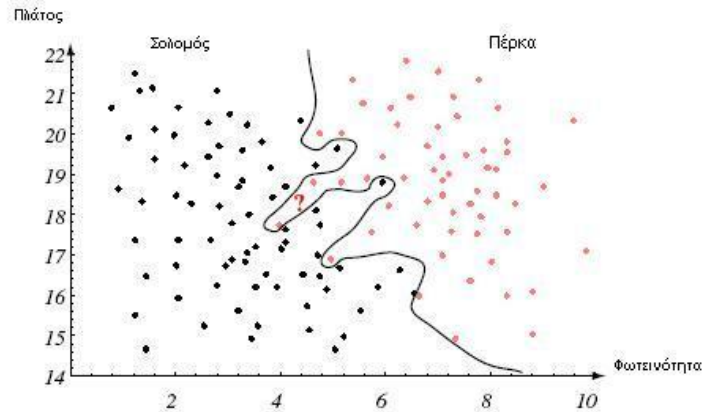
Αυτός ο κανόνας δείχνει να είναι καλός στο διαχωρισμό των δειγμάτων και προτρέπει στο να χρησιμοποιήσουμε ακόμη περισσότερα χαρακτηριστικά. Εκτός από τη φωτεινότητα και το πλάτος των ψαριών μπορούμε ακόμη να συμπεριλάβουμε κάποια σχηματική παράμετρο, όπως είναι η κυρτή γωνία του ραχιαίου πτερυγίου ή ακόμη την τοποθεσία των οφθαλμών (όπως αυτή εκφράζεται από την αναλογία της απόστασης στόματος-ουράς) κ.ο.κ. Πως όμως γνωρίζουμε εκ των προτέρων ποιο από τα χαρακτηριστικά αυτά θα δουλέψει καλύτερα; Κάποια χαρακτηριστικά μπορεί να είναι πλεονάζοντα. Για παράδειγμα, αν το χρώμα των ματιών των ψαριών σχετίζεται άφογα και με το πλάτος τους, τότε η επιλογή του χρώματος του ματιού ως χαρακτηριστικό θα είναι πλεονάζουσα. Ακόμη και αν η χρήση πολλών χαρακτηριστικών υποθέσουμε ότι δεν επιβαρύνει το υπολογιστικό κόστος μήπως υπάρχει κάποια «κατάρρα» η οποία επιβαρύνει το να δουλεύουμε σε πολλές διαστάσεις;



Εικόνα 1.4

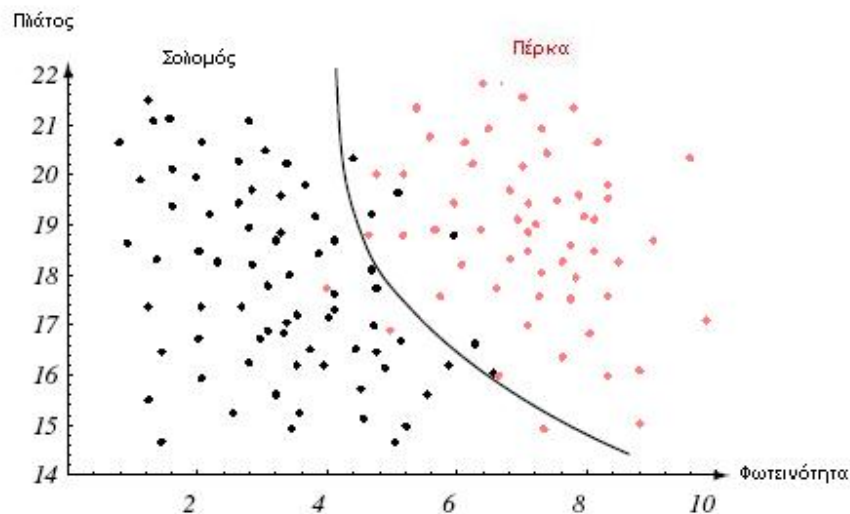
Έστω ότι τα άλλα χαρακτηριστικά είναι δύσκολο (ή ακριβό) να μετρηθούν ή απλώς προσφέρουν μικρή βελτίωση στον ταξινομητή ή ακόμη χειροτερεύουν την ταξινόμηση. Ας υποθέσουμε επίσης ότι είμαστε αναγκασμένοι να χρησιμοποιήσουμε τα δύο χαρακτηριστικά της εικόνας 1.4. Αν τα μοντέλα μας ήταν πολύ περίπλοκα, ο ταξινομητής μας θα είχε όριο απόφασης πιο πολύπλοκο από αυτήν την απλή ευθεία γραμμή. Στην περίπτωση αυτή όλα τα δείγματα εκπαίδευσης θα χωρίζονταν τέλεια όπως φαίνεται στο σχήμα 1.5. Με μια τέτοια λύση όμως, η ικανοποίησή μας θα ήταν πρώιμη, επειδή ο κεντρικός σκοπός μας, όταν κατασκευάζουμε ένα ταξινομητή, είναι να μας προτείνει δράσεις όταν παρουσιάζονται σε αυτόν νέα πρότυπα όπως π.χ. κάποιο ψάρι που δεν έχει ξαναδεί. Αυτό είναι γνωστό με τον όρο γενίκευση. Είναι απίθανο το περίπλοκο όριο απόφασης του σχήματος 1.5 να παρέχει καλή γενίκευση - μοιάζει να είναι ρυθμισμένο να δουλεύει καλά μόνο για τα συγκεκριμένα δείγματα εκπαίδευσης παρά να βασίζεται στα γενικά χαρακτηριστικά ή στο πραγματικό μοντέλο που χαρακτηρίζει όλες τις πέρκες και τους σολομούς που πρέπει να διαχωρίσει.

Φυσικά, μια άλλη προσέγγιση θα ήταν να πάρουμε όσο το δυνατόν περισσότερα δείγματα εκπαίδευσης για να έχουμε την όσο το δυνατόν καλύτερη εκτίμηση των σχετικών χαρακτηριστικών, όπως για παράδειγμα την κατανομή της πιθανότητας κάθε κατηγορίας. Σε κάποια προβλήματα αναγνώρισης προτύπων, πάντως, το πλήθος των δεδομένων που μπορούμε να έχουμε είναι περιορισμένο. Ακόμη και αν είχαμε ένα μεγάλο πλήθος δεδομένων εκπαίδευσης όμως, σε έναν συνεχές χώρο χαρακτηριστικών, αν ακολουθούσαμε την προσέγγιση της εικόνας 1.5 ο ταξινομητής μας θα παρουσίαζε ένα πολύ περίπλοκο όριο απόφασης που επίσης δε θα δούλευε καλά για νέα πρότυπα.



Εικόνα 1.5

Αντί για αυτό, ψάχνουμε να απλοποιήσουμε τον ταξινομητή μας και το όριο απόφασης ώστε να μην είναι τόσο πολύπλοκο όσο αυτό του σχήματος 1.5. Όντως μπορούμε να είμαστε ικανοποιημένοι με μια κάπως πιο χαμηλή απόδοση του ταξινομητή μας στα δεδομένα εκπαίδευσης, αν αυτό οδηγήσει σε έναν ταξινομητή με καλύτερη απόδοση σε καινούρια πρότυπα. Όπως έλεγε και ο William of Occam (1284-1347) "Entities are not to be multiplied without necessity". Δηλαδή, οι αποφάσεις που βασίζονται σε συνολικά πολύπλοκα μοντέλα, συχνά οδηγούν σε χαμηλότερη ακρίβεια όσον αφορά τον ταξινομητή. Αλλά αν με το να σχεδιάσουμε ένα περίπλοκο ταξινομητή είναι απίθανο τελικά να έχουμε μια καλή γενίκευση, πώς μπορούμε τελικά να αποφασίζουμε για έναν απλό ταξινομητή; Πως το σύστημά μας αυτόματα θα καθορίζει ότι η απλή καμπύλη του σχήματος 1.6 είναι καταλληλότερη από την ευθεία του σχήματος 1.4 ή το πολύπλοκο όριο του σχήματος 1.5; Ας υποθέσουμε ότι κατά κάποιο τρόπο επιτυγχάνουμε να καθορίσουμε την βελτίωση του ταξινομητή μας λαμβάνοντας υπόψη μας τα παραπάνω. Μπορούμε να προβλέψουμε πόσο καλά το σύστημά μας θα αναγνωρίσει νέα πρότυπα; Αυτά είναι κάποια από τα κυρίαρχα προβλήματα στη στατιστική αναγνώριση προτύπων.



Εικόνα 1.6

Για τα ίδια εισαγόμενα στον ταξινομητή πρότυπα μπορεί να θέλουμε να χρησιμοποιήσουμε άλλη συνάρτηση κόστους ή να θέλουμε να χρησιμοποιήσουμε τον ταξινομητή για να κάνουμε διαφορετικό διαχωρισμό. Μπορεί για παράδειγμα, να θέλουμε αντί να ξεχωρίζει σολομούς από πέγκες, να κατατάσσει τα ψάρια ανάλογα με το φύλο τους – όλα τα θηλυκά (ό,τι είδος και να είναι) από όλα τα αρσενικά ψάρια - αν θέλουμε να πουλήσουμε αυγοτάραχο. Εναλλακτικά, μπορεί να θέλουμε να επιλέξουμε τα αλλοιωμένης εμφάνισης ψάρια για τροφή για γάτες ή και άλλα πολλά σενάρια. Διαφορετικές αποφάσεις μπορεί να απαιτούν χαρακτηριστικά και όρια διαφορετικά από αυτά του αρχικού μας προβλήματος ταξινόμησης.

Αυτό κάνει ξεκάθαρο το ότι οι αποφάσεις μας βασίζονται στο κόστος και την εργασία που επιθυμούμε και ότι το να δημιουργήσουμε ένα μοντέλο αναγνώρισης προτύπων για γενική χρήση -δηλαδή ένα που να ενεργεί ακριβέστατα, βασισμένο σε μια μεγάλη ποικιλία από εργασίες- είναι όντως δύσκολο εγχείρημα. Αυτό επίσης μας κάνει να εκδηλώσουμε τον θαυμασμό μας στην ικανότητα του ανθρώπου να «δουλεύει» καλά σε διαφορετικά προβλήματα αναγνώρισης προτύπων και να εναλλάσσεται στιγμιαία μεταξύ τους.

Επειδή η ταξινόμηση είναι βασικά μια διαδικασία εύρεσης του μοντέλου που γεννάει τα πρότυπα, διάφορες τεχνικές αναγνώρισης είναι χρήσιμες ανάλογα με το είδος των υποψηφίων μοντέλων. Στην στατιστική αναγνώριση προτύπων επικεντρωνόμαστε στις στατιστικές ιδιότητες των προτύπων (όπως γενικά εκφράζονται από την πυκνότητα πιθανότητας). Αυτό θα είναι το κύριο μέλημα μας στο βιβλίο που έχετε στα χέρια σας. Εδώ το μοντέλο για ένα πρότυπο μπορεί να είναι ένα μοναδικό σύνολο χαρακτηριστικών, αν και το πραγματικό πρότυπο, όταν μετρήθηκε, μπορεί να έχει επηρεαστεί από κάποιο τυχαίο θόρυβο. Σε μερικές περιπτώσεις θεωρείται ότι η αναγνώριση προτύπων μέσω Νευρωνικών Δικτύων είναι μια προσέγγιση που ευσταθεί από μόνη της αλλά εκτός από μια κάπως διαφορετική, πιο διανοητική προέλευση δεν έχει κάτι διαφορετικό που να μας οδηγήσει στο να την ξεχωρίσουμε από την Στατιστική Αναγνώριση Προτύπων. Οι λόγοι θα γίνουν προφανείς στη συνέχεια. Σε περίπτωση που το μοντέλο μας αποτελείται από ένα σύνολο από λογικούς κανόνες, τότε εφαρμόζουμε τις μεθόδους της Συντακτικής Αναγνώρισης Προτύπων, όπου γραμματικοί κανόνες περιγράφουν την απόφαση μας. Για παράδειγμα μπορεί να θέλουμε να ταξινομήσουμε μια πρόταση στα Αγγλικά αν είναι γραμματικώς σωστή ή όχι. Εδώ οι λογικοί κανόνες, αντί για τους στατιστικούς (συχνότητες λέξεων κ.λ.π.) είναι χρησιμότεροι.

Ήταν χρήσιμο στο παράδειγμα μας με τα ψάρια να επιλέξουμε τα χαρακτηριστικά μας προσεχτικά και να καταλήξουμε έτσι σε μια γραφική αναπαράσταση (όπως αυτή της Εικόνας 1.6), η οποία μας επέτρεψε να έχουμε μια επιτυχή ταξινόμηση. Ένα σημαντικό θέμα σε κάθε πρόβλημα αναγνώρισης προτύπων είναι να κατασκευάσουμε μια τέτοια «επιτυχή» αναπαράσταση, στην οποία οι δομικές σχέσεις ανάμεσα στα χαρακτηριστικά να είναι απλή και να διαφαίνεται απλά. Επίσης, θα πρέπει μέσω αυτής της αναπαράστασης να εκφράζεται και το πραγματικό (αλλά άγνωστο) μοντέλο των προτύπων. Σε μερικές περιπτώσεις τα πρότυπα θα πρέπει να αναπαρίστανται ως διανύσματα πραγματικών αριθμών, σε άλλες ως λίστες με ιδιότητες, σε άλλες ως περιγραφές των μερών και των σχέσεων τους κ.ά. Εμείς αναζητούμε να βρούμε μια αναπαράσταση, στην οποία τα πρότυπα που οδηγούν στην ίδια απόφαση να είναι κατά κάποιο τρόπο κοντά το ένα με το άλλο, αλλά επίσης και όσο γίνεται πιο μακριά από τα πρότυπα που οδηγούν σε διαφορετική απόφαση. Το πώς δημιουργούμε ή «μαθαίνουμε» μια σωστή αναπαράσταση και το πόσο καλά ορίζουμε το «μακριά» και το «κοντά» είναι οι βασικοί παράγοντες που

καθορίζουν την επιτυχία ενός ταξινομητή προτύπων. Κάποια επιπρόσθετα χαρακτηριστικά είναι σχεδόν πάντα χρήσιμα για την αναπαράσταση. Όμως συνήθως επιλέγουμε να χρησιμοποιήσουμε έναν μικρό αριθμό χαρακτηριστικών, τα οποία μπορεί α) να οδηγήσουν σε απλούστερες περιοχές απόφασης και β) να απλοποιήσουν το χρόνο εκπαίδευσης του ταξινομητή. Μπορεί επίσης να θέλουμε να χρησιμοποιήσουμε χαρακτηριστικά που είναι εύρωστα, ανθεκτικά δηλαδή σε θόρυβο ή άλλα λάθη. Σε πρακτικές εφαρμογές εντούτοις μπορεί να θέλουμε να έχουμε ένα γρήγορο ταξινομητή ή έναν που να χρησιμοποιεί λίγα ηλεκτρονικά μέσα (αισθητήρες, πολλούς H/Y κ.ά.), λίγη μνήμη ή ακόμα και λίγα υπολογιστικά βήματα.

Μια κεντρική τεχνική, όταν έχουμε στη διάθεσή μας ανεπαρκή δεδομένα εκπαίδευσης, είναι να ενσωματώσουμε γνώση για την περιοχή του προβλήματος. Όντως, όσο λιγότερα είναι τα δεδομένα εκπαίδευσης, τόσο επιτακτικότερη είναι και η ανάγκη για την ενσωμάτωση αυτής της γνώσης –όπως για παράδειγμα, το πως τα ίδια τα πρότυπα έχουν παραχθεί. Μία μέθοδος που λαμβάνει τα παραπάνω υπόψην της είναι η ανάλυση από τη σύνθεση, όπου στην ιδανική περίπτωση έχουμε ένα μοντέλο για το πως το κάθε πρότυπο παράγεται. Θεωρήστε το παράδειγμα της αναγνώριση φωνής. Ανάμεσα στη διαφορετικότητα της εκφοράς της συλλαβής "dee" από διάφορους ανθρώπους, το κοινό στοιχείο είναι ότι για να την παράγουν όλοι κάνουν τα εξής: χαμηλώνουν λίγο το σαγόκι, ανοίγουν το στόμα, τοποθετούν τη γλώσσα τους στην «οροφή» του στόματός τους μετά από ένα καθορισμένο χρονικό διάστημα κτλ. Μπορούμε να υποθέσουμε ότι όλες οι ακουστικές παραλλαγές της συλλαβής οφείλονται στο αν ο ομιλητής είναι άνδρας ή γυναίκα, νέος ή γέρος, με διαφορετικό τόνο στο λόγο τους κ.ά. Σε κάποιο επίπεδο το μοντέλο «φυσιολογίας» (ή αλλιώς «μηχανικό» μοντέλο) για την παραγωγή του "dee" είναι κατάλληλο και διαφορετικό από αυτό για την παραγωγή της συλλαβής "doo" και από όλες τις άλλες συλλαβές. Αν το μοντέλο αυτό της παραγωγής του "dee" εξαρτάται από τον ήχο (και αυτό είναι όντως ένα μεγάλο αν), τότε μπορούμε να ταξινομήσουμε τις συλλαβές ανάλογα με το πως παράγονται. Αυτό σημαίνει ότι η αναπαράσταση της παραγωγής μπορεί να είναι η καλύτερη αναπαράσταση για να γίνει ταξινόμηση. Το σύστημα αναγνώρισης προτύπων θα πρέπει να αναλύσει (και έτσι να ταξινομήσει) τα πρότυπα εισόδου βασισμένο στο πώς συντίθεται κάθε πρότυπο. Το κόλπο είναι φυσικά, το σύστημα να μπορεί να ανακτήσει τις απαραίτητες παραμέτρους από τους αισθητήρες που εφαρμόζονται στα πρότυπα.

Αναλογιστείτε τη δυσκολία της δημιουργίας μιας συσκευής αναγνώρισης όλων των ειδών καρεκλών - τυπικές καρέκλες γραφείου, σύγχρονες καρέκλες σαλονιού, καρέκλες με μορφή σάκου κ.ά. - η οποία να βασίζεται σε φωτογραφίες. Δοθείσης της ποικιλίας στον αριθμό των ποδιών, το υλικό κατασκευής, το σχήμα κ.ο.κ. απελιζόμαστε στην ιδέα του να βρούμε μια αναπαράσταση που να βοηθάει στην κατάταξη των καρεκλών. Ίσως το πιο «ενωτικό» χαρακτηριστικό των καρεκλών είναι λειτουργικό: η καρέκλα είναι ένα σταθερό αντικείμενο το οποίο υποστηρίζει έναν καθιστό άνθρωπο και έχει στήριγμα για την πλάτη του. Έτσι, θα μπορούσαμε να προσπαθήσουμε να εξάγουμε τέτοια λειτουργικά χαρακτηριστικά από τις φωτογραφίες που έχουν τραβηχτεί. Και η έκφραση «ένα σταθερό αντικείμενο το οποίο υποστηρίζει έναν καθιστό άνθρωπο» είναι έμμεσα συνδεδεμένη με τον προσανατολισμό της μεγάλης επιφάνειας της καρέκλας, και θα πρέπει να ισχύει ακόμα και αν πρόκειται για καρέκλα σε μορφή σάκου. Αυτό το τελευταίο απαιτεί παραπέρα εξήγηση, αλλά η συζήτηση αυτή είναι περισσότερο θέμα της επιστήμης της

μηχανικής όρασης (computer vision) και δεν είναι τόσο σχετικό με την αναγνώριση προτύπων.

Χωρίς να χρειαστεί να εισχωρήσουμε σε τόσο ακραία παραδείγματα, βλέπουμε ότι πολλά συστήματα αναγνώρισης προτύπων του πραγματικού κόσμου αναζητούν και ενσωματώνουν τουλάχιστον κάποια γνώση, όσον αφορά τη μέθοδο παραγωγής των προτύπων, ή την πρακτική τους χρήση, ώστε να εξασφαλίσουν μια καλή αναπαράσταση, παρότι ο στόχος της αναπαράστασης είναι η ταξινόμηση και όχι η αναπαραγωγή. Για παράδειγμα, στην οπτική αναγνώριση χαρακτήρων κειμένου (OCR), κάποιος μπορεί να υποθέσει ότι οι χαρακτήρες κειμένου –γραμμένοι με το χέρι από άνθρωπο- γράφονται ως μια ακολουθία από απλές γραμμές και μπορεί να προσπαθήσει να ανακτήσει αναπαραστάσεις γραμμών από την εικόνα που έχει ληφθεί και να συμπεράνει τον χαρακτήρα με βάση αυτές τις αναγνωρισμένες γραμμές.

1.3.1 Σχετικά πεδία

Η ταξινόμηση προτύπων διαφέρει από την κλασική στατιστική δοκιμή της υπόθεσης, όπου τα δεδομένα που δίνουν οι αισθητήρες χρησιμοποιούνται για να δεχτούμε ή να απορρίψουμε μια αρχική υπόθεση για χάρη μίας άλλης εναλλακτικής υπόθεσης. Μιλώντας γενικά, αν η πιθανότητα του να ισχύει μια υπόθεση πέφτει κάτω από ένα κατώφλι τότε απορρίπτουμε την υπόθεση αυτή προς όφελος της εναλλακτικής. Αυτή η μέθοδος του ελέγχου μίας υπόθεσης συναντάται συχνά για να προσδιορίσουμε αν ένα φάρμακο είναι αποτελεσματικό, όπου η αρχική υπόθεση είναι ότι το φάρμακο δεν έχει καμία επίδραση. Ο έλεγχος της υπόθεσης μπορεί να χρησιμοποιηθεί για να καθορίσει αν ένα ψάρι επάνω στον ιμάντα μεταφοράς ανήκει σε μια μόνο τάξη (για παράδειγμα, όλα τα ψάρια να είναι σολομοί) -η αρχική υπόθεση- ή αν ανήκει σε μια από τις δύο τάξεις (η εναλλακτική υπόθεση).

Η αναγνώριση προτύπων διαφέρει και από την επεξεργασία εικόνας. Στην επεξεργασία εικόνας η είσοδος είναι μια εικόνα και η έξοδος μια εικόνα. Τα βήματα στην επεξεργασία εικόνας περιλαμβάνουν συχνά την περιστροφή, την αύξηση του contrast (αντίθεσης) και άλλους μετασχηματισμούς που διατηρούν όμως όλη την αρχική πληροφορία της εικόνας.

Η εξαγωγή χαρακτηριστικών, όπως είναι η εύρεση κορυφών ή των περιοχών της έντασης, οδηγεί σε απώλεια πληροφοριών (αλλά ελπίζουμε να διατηρεί όλα τα σχετικά στοιχεία για την εργασία μας).

Όπως έχει ήδη αναφερθεί η εξαγωγή χαρακτηριστικών παίρνει ως είσοδο ένα πρότυπο και εξάγει τιμές χαρακτηριστικών. Ο αριθμός των χαρακτηριστικών επιλέγεται έτσι ώστε να είναι μικρότερος από τον συνολικό αριθμό που απαιτείται για να περιγράψει απόλυτα ο στόχος του ενδιαφέροντος μας. Αυτό όμως μας οδηγεί σε απώλεια πληροφορίας. Στα πλαίσια της βοηθητικής μνήμης, το σύστημα παίρνει ως είσοδο ένα πρότυπο και εξάγει ένα άλλο το οποίο είναι αντιπροσωπευτικό μιας γενικής ομάδας προτύπων. Έτσι, μειώνει την πληροφορία κατά ένα ποσοστό αλλά πολύ λιγότερο από ότι η διαδικασία αναγνώρισης προτύπων. Λόγω του κρίσιμου ρόλου της απόφασης στην αναγνώριση προτύπων, είναι βασική μια διαδικασία μείωσης της πληροφορίας. Δεν μπορείς να ανακατασκευάσεις ένα πρότυπο γνωρίζοντας μόνο την κατηγορία στην οποία ανήκει. Το βήμα της ταξινόμησης αναπαριστά ένα ακόμη πιο σημαντικό χάσιμο πληροφορίας μειώνοντας τα αρχικά μερικά χιλιάδες bits που αναπαριστούν όλα τα χρώματα καθενός από τα χιλιάδες pixels της εικόνας ενός προτύπου, σε λίγα μόνο bits που αναπαριστούν την κατηγορία στην οποία ανήκει το πρότυπο (στο παράδειγμά μας με τα ψάρια, είναι ένα μόνο bit).

Τρεις στενά συνδεδεμένες τεχνικές οι οποίες λαμβάνουν μέρος στην αναγνώριση προτύπων είναι η ανάδραση-αναδρομή, η παρεμβολή και η εκτίμηση πυκνότητας πιθανότητας. Στην αναδρομή μπορεί να ψάχνουμε να βρούμε κάποια λειτουργική περιγραφή των δεδομένων, συχνά με σκοπό να προβλέψουμε τις μελλοντικές τιμές από νέες εισόδους. Η γραμμική αναδρομή (linear regression) -στην οποία η συνάρτηση είναι γραμμική στο πεδίο ορισμού των μεταβλητών της- είναι μακράν η πιο δημοφιλής και καλά μελετημένη μορφή της ανάδρασης. Μπορεί για παράδειγμα να «δούμε» ότι το μήκος του σολομού μεταβάλλεται ανάλογα με την ηλικία του και το βάρος του και στη συνέχεια να πάρουμε μετρήσεις της ηλικίας και του μήκους από πολλούς τυπικούς σολομούς και να χρησιμοποιήσουμε γραμμική αναδρομή για να βρούμε τους συντελεστές.

Στην παρεμβολή ξέρουμε ή μπορούμε εύκολα να βρούμε τη συνάρτηση για δεδομένες τιμές της εισόδου. Το πρόβλημα είναι να παρεμβάλλουμε τη συνάρτηση για τις ενδιάμεσες τιμές της εισόδου. Έτσι, μπορεί να γνωρίζουμε πως το μήκος του σολομού ποικίλει ανάλογα με την ηλικία του, τις πρώτες δύο εβδομάδες ζωής του, και μετά από τα δύο χρόνια ζωής του. Στη συνέχεια εφαρμόζοντας μεθόδους γραμμικής παρεμβολής βρίσκουμε με ποιο τρόπο το μήκος εξαρτάται από την ηλικία του, στο διάστημα από δύο εβδομάδες έως και δύο χρόνια ζωής του.

Η εκτίμηση πυκνότητας (πιθανότητας) είναι το πρόβλημα υπολογισμού της πυκνότητας (ή της πιθανότητας) του ότι ένα μέλος μίας συγκεκριμένης κατηγορίας θα βρεθεί να έχει συγκεκριμένα χαρακτηριστικά.

Αυτές οι τρεις προαναφερθέντες τεχνικές συχνά χρησιμοποιούνται -αποκλειστικά ή σε συνδυασμό- σαν πρώτα βήματα στην αναγνώριση προτύπων. Για παράδειγμα, θα συναντήσουμε αρκετές μεθόδους για την εκτίμηση των πυκνοτήτων διαφορετικών κατηγοριών. Τότε, ένα άγνωστο πρότυπο ταξινομείται ανάλογα με το σε πια κατηγορία είναι πιο πιθανό να ανήκει. Παρόλο που αυτά τα πεδία έχουν αναπτυχθεί σε μεγάλο βαθμό και έχουν αποδείξει ότι είναι πολύ χρήσιμα, εμείς θα τα χρησιμοποιήσουμε μόνο έμμεσα, στο βαθμό που σχετίζονται με την αναγνώριση προτύπων.

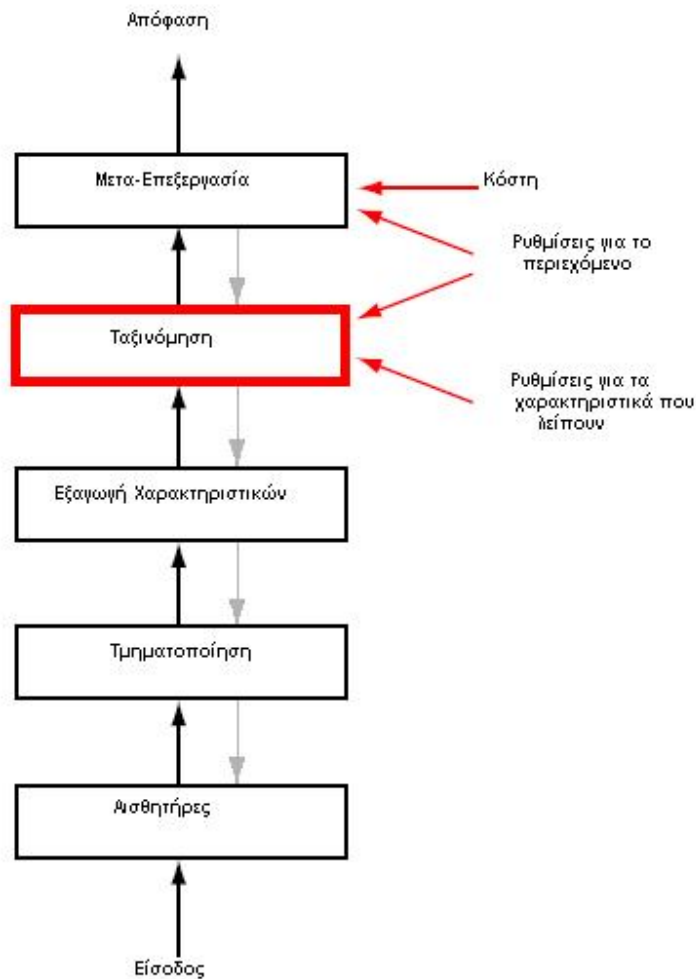
1.4 Συστήματα Αναγνώρισης Προτύπων

Στην περιγραφή του υποθετικού μας συστήματος αυτόματης ταξινόμησης ψαριών διαχωρίσαμε τις τρεις διαφορετικές λειτουργίες: την προεπεξεργασία, την εξαγωγή χαρακτηριστικών και την ταξινόμηση (εικόνα 1.1). Η εικόνα 1.7 δείχνει ένα πιο εκλεπτυσμένο διάγραμμα των συνιστωσών ενός συστήματος ταξινόμησης προτύπων. Για να κατανοήσουμε το πρόβλημα του σχεδιασμού ενός τέτοιου συστήματος πρέπει να κατανοήσουμε πρώτα τα προβλήματα που κάθε μια από αυτές τις συνιστώσες πρέπει να επιλύσει. Στη συνέχεια θα δούμε τις λειτουργίες της κάθε συνιστώσας με τη σειρά και θα περιγράψουμε επιμέρους προβλήματα που μπορεί να προκύψουν.

1.4.1 Αισθητήρες

Η είσοδος σε ένα ΣΑΠ (Σύστημα Αναγνώρισης Προτύπων) είναι αρκετά συχνά κάποιος αναμεταδότης, όπως μια κάμερα ή μια συστοιχία μικροφώνων. Η δυσκολία του προβλήματος εδώ μπορεί να εξαρτάται από τα χαρακτηριστικά αυτών των αισθητήρων και ειδικά από κάποιους περιορισμούς όπως το εύρος συχνοτήτων, η ανάλυση της εικόνας, η ευαισθησία, η παραμόρφωση, η αναλογία σήματος προς θόρυβο, η καθυστέρηση του σήματος κ.ά. Όσο σημαντικό και αν είναι αυτό το πρόβλημα στην

πράξη, ο σχεδιασμός αισθητήρων για ΣΑΠ είναι πέρα από τους σκοπούς αυτού του βιβλίου.



Εικόνα 1.7

1.4.2 Τμηματοποίηση και Ομαδοποίηση.

Στο παράδειγμά μας με τα ψάρια, υπονοήσαμε ότι κάθε ψάρι πάνω στον ιμάντα μεταφοράς είναι απομονωμένο, χωριστά από τα υπόλοιπα και έτσι μπορούσε εύκολα να διαχωριστεί και από τον ίδιο τον ιμάντα.

Στην πράξη όμως τα ψάρια μπορεί κάτω από τον αισθητήρα να είναι επικαλυπτόμενα οπότε το σύστημά μας θα έπρεπε με κάποιο τρόπο να μπορεί να καταλάβει που ένα ψάρι τελειώνει και που αρχίζει το επόμενο. Το κάθε πρότυπο πρέπει να τμηματοποιηθεί. Αν είχαμε αναγνωρίσει το κάθε ψάρι θα ήταν πιο εύκολο να τμηματοποιήσουμε την εικόνα του. Αλλά πώς γίνεται να τμηματοποιήσουμε τις εικόνες πριν αυτές να έχουν κατηγοριοποιηθεί ή πώς να τις κατηγοριοποιήσουμε αν αυτές δεν έχουν προηγουμένως τμηματοποιηθεί; Φαίνεται να χρειαζόμαστε ένα τρόπο να ξέρουμε πότε έχουμε αλλάξει από το ένα μοντέλο στο άλλο ή να ξέρουμε πότε έχουμε απλώς φόντο ή «καμία κατηγορία». Πώς όμως μπορεί να γίνει κάτι τέτοιο;

Η τμηματοποίηση είναι ένα από τα μεγαλύτερα προβλήματα στην αναγνώριση προτύπων.

1.4.3 Εξαγωγή Χαρακτηριστικών

Το θεωρητικό όριο ανάμεσα στην εξαγωγή χαρακτηριστικών και την ταξινόμηση είναι κάπως αυθαίρετο. Ένας ιδανικός εξαγωγέας χαρακτηριστικών θα παράγαγε μια αναπαράσταση, η οποία θα έκανε αρκετά εύκολο το έργο του ταξινομητή. Αντίστροφα ένας καλός ταξινομητής δε θα χρειαζόταν βοήθεια από έναν πολύπλοκο εξαγωγέα χαρακτηριστικών. Η διάκριση αυτών των διαδικασιών γίνεται κυρίως για πρακτικούς λόγους παρά για θεωρητικούς.

Ο παραδοσιακός στόχος του εξαγωγέα χαρακτηριστικών είναι να χαρακτηρίσει ένα αντικείμενο που πρόκειται να αναγνωριστεί από μετρήσεις των οποίων οι τιμές είναι αρκετά παρόμοιες, για αντικείμενα που ανήκουν στην ίδια κατηγορία, και πολύ διαφορετικές, για αντικείμενα που ανήκουν σε διαφορετικές κατηγορίες. Αυτό οδηγεί στην ιδέα του να ψάχνεις για διαφορετικά χαρακτηριστικά τα οποία να είναι αμετάβλητα όταν υπόκεινται σε «άσχετους» μετασχηματισμούς της εισόδου. Στο παράδειγμά μας η ακριβής θέση του ψαριού στον μάντα μεταφοράς είναι άσχετη με την κατηγορία στην οποία αυτό ανήκει και έτσι πρέπει και η αναπαράσταση που θα επιλέξουμε να είναι ανεξάρτητη από τη θέση του ψαριού πάνω στον μάντα. Ιδανικά, σε αυτή την περίπτωση θέλουμε τα χαρακτηριστικά να είναι αμετάβλητα από τη μετάφραση της αναπαράστασης είτε δηλαδή βρίσκονται οριζόντια είτε κάθετα. Επειδή η περιστροφή είναι και αυτή άσχετη της ταξινόμησης, θέλουμε επίσης να είναι και τα χαρακτηριστικά αμετάβλητα από την περιστροφή. Τελικά, ακόμη και το μέγεθος του ψαριού μπορεί να μην είναι σημαντικό -για παράδειγμα ένας μικρός σολομός δεν παύει να είναι σολομός. Έτσι, θέλουμε τα χαρακτηριστικά να παραμένουν αμετάβλητα από την βάρθρωση. Γενικά χαρακτηριστικά που περιγράφουν ιδιότητες, όπως το σχήμα, το χρώμα και πολλά άλλα είδη υφής παραμένουν αμετάβλητα από την μετάφραση, την περιστροφή και τη βάρθρωση.

1.4.4 Ταξινόμηση

Ο βαθμός δυσκολίας του προβλήματος ταξινόμησης εξαρτάται από την μεταβλητότητα των τιμών των χαρακτηριστικών για αντικείμενα της ίδιας κατηγορίας σε σχέση με τη διαφορά μεταξύ των τιμών των χαρακτηριστικών για αντικείμενα που ανήκουν σε διαφορετικές κατηγορίες. Η μεταβλητότητα των τιμών των χαρακτηριστικών των αντικείμενων της ίδιας κατηγορίας μπορεί να οφείλεται στην πολυπλοκότητα αλλά και στο θόρυβο. Ορίζουμε το θόρυβο σύμφωνα με τους ακόλουθους γενικούς όρους: κάθε ιδιότητα του προτύπου που δεν οφείλεται στο πραγματικό υποκείμενο μοντέλο αλλά απεναντίας στην τυχαιότητα του πραγματικού κόσμου ή των αισθητήρων. Όλα τα δύσκολα προβλήματα της αναγνώρισης προτύπων αφορούν το θόρυβο. Όλες οι δύσκολες αποφάσεις και τα προβλήματα αναγνώρισης προτύπων εμπεριέχουν τον θόρυβο σε κάποια μορφή του. Ποιος όμως είναι ο καλύτερος τρόπος να σχεδιάσουμε ένα ταξινομητή ο οποίος θα συνεργάζεται με αυτήν την μεταβλητότητα; Ποια είναι η καλύτερη δυνατή απόδοση του ταξινομητή;

Ένα πρόβλημα που εμφανίζεται στην πράξη είναι ότι δεν μπορούμε πάντα να καθορίσουμε τις τιμές όλων των χαρακτηριστικών για μια δεδομένη είσοδο. Στο υποθετικό μας σύστημα για ταξινόμηση ψαριών, για παράδειγμα, μπορεί να μην είναι δυνατό να καθορίσουμε το πλάτος ενός ψαριού επειδή ίσως μέρος του κρύβεται από άλλο ψάρι. Πώς θα μπορούσε ο ταξινομητής να ισοσκελίσει αυτήν την απώλεια; Η αφελής μέθοδος της υπόθεσης ότι η τιμή του χαρακτηριστικού που λείπει είναι μηδέν ή ίση με τη

μέση τιμή των άλλων τιμών για τα πρότυπα που έχουν ήδη αναγνωριστεί, είναι μάλλον μη βέλτιστη. Παρομοίως, πώς μπορούμε να εκπαιδύσουμε ένα ταξινομητή ή να τον χρησιμοποιήσουμε όταν κάποια χαρακτηριστικά λείπουν;

1.4.5 Μετα-Επεξεργασία

Ένας ταξινομητής σπανίως υπάρχει αφ' εαυτού. Απεναντίας, γενικά χρησιμοποιείται για να προτείνει περαιτέρω δράσεις (όπως τοποθέτησε αυτό το ψάρι σε εκείνο το καλάθι, το άλλο ψάρι στο άλλο καλάθι) κάθε μια από τις οποίες έχει ένα σχετικό κόστος. Ο μετα-επεξεργαστής χρησιμοποιεί την έξοδο του ταξινομητή για να αποφασίσει για την προτεινόμενη δράση.

Εννοιολογικά, το απλούστερο μέτρο της απόδοσης ενός ταξινομητή είναι ο ρυθμός εμφάνισης λαθών, δηλαδή το ποσοστό των νέων προτύπων που ταξινομούνται λανθασμένα. Έτσι, είναι σύνηθες να ψάχνουμε για την ταξινόμηση με τον ελάχιστο ρυθμό λαθών. Πάντως, μπορεί να είναι πολύ καλύτερο να προτείνουμε δράσεις που θα ελαχιστοποιούν το συνολικό αναμενόμενο κόστος, το οποίο καλείται ρίσκο. Πώς μπορούμε να ενσωματώσουμε τη γνώση μας σχετικά με τα κόστη και πώς αυτή η γνώση θα επηρεάσει την τελική απόφαση του ταξινομητή; Μπορούμε να εκτιμήσουμε το συνολικό κόστος και έτσι να πούμε πότε ο ταξινομητής μας είναι σε επιτρεπτά όρια αποτυχίας; Μπορούμε να εκτιμήσουμε το ελάχιστο πιθανό ρίσκο κάθε ταξινομητή, για να δούμε πόσο κοντά στον ιδανικό είναι ο δικός μας ή αν τελικά το πρόβλημα που καλείται να λύσει είναι απλώς πολύ δύσκολο συνολικά;

Στο παράδειγμά μας με τα ψάρια είδαμε το ότι η χρήση περισσότερων του ενός χαρακτηριστικών μπορεί να οδηγήσει σε βελτιωμένη αναγνώριση. Μπορεί κάποιος να φανταστεί ότι θα μπορούσαμε να τα πάμε καλύτερα αν χρησιμοποιούμε περισσότερους του ενός ταξινομητές, με κάθε ταξινομητή να επενεργεί σε διαφορετικές απόψεις κάθε εισόδου. Για παράδειγμα, θα μπορούσαμε να συνδυάσουμε τα αποτελέσματα της ακουστικής αναγνώρισης προτύπων και του διαβάσματος των χειλιών για να βελτιώσουμε την απόδοση της αναγνώρισης προφορικού λόγου.

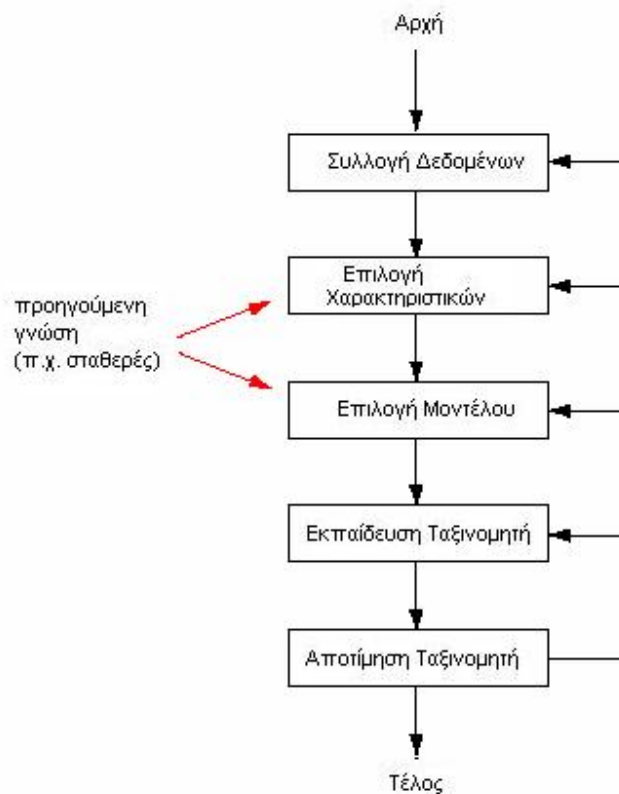
Αν όλοι οι ταξινομητές συμφωνούν σε ένα συγκεκριμένο πρότυπο τότε δεν υπάρχει καμία δυσκολία. Ας υποθέσουμε όμως ότι διαφωνούν. Πώς θα μπορούσε ένας υπερ-ταξινομητής να εξάγει τις απαραίτητες αποδείξεις από τις επιμέρους αναγνωρίσεις, ώστε να επιτύχει βέλτιστη απόδοση; Φανταστείτε ότι καλείτε δέκα ειδικούς για να καθορίζουν εάν ένα ψάρι είναι άρρωστο. Ενώ οι εννιά συμφωνούν ότι το συγκεκριμένο ψάρι είναι υγιές ο ένας διαφωνεί. Ποιος έχει δίκιο; Μπορεί να τύχει ο ένας που διαφωνεί με τους υπόλοιπους να είναι ο μόνος γνώστης πάνω στα συγκεκριμένα πολύ σπάνια συμπτώματα ασθένειας ψαριών και να έχει δίκιο. Πώς όμως ένας υπέρ-ταξινομητής θα μπορούσε να γνωρίζει πότε να βασιστεί μια απόφαση στην άποψη της μειοψηφίας ακόμα και όταν αυτή εκφράζεται από ένα ειδικό σε μια μικρή περιοχή γνώσης, ο οποίος δεν έχει τα προσόντα να παίρνει αποφάσεις;

Όμως στο σημείο αυτό πρέπει να αναφέρουμε ότι έχουμε εγείρει πολύ περισσότερα ερωτήματα από όσα έχουμε απαντήσει. Στόχος μας ήταν να δώσουμε έμφαση στην πολυπλοκότητα που διέπει τα προβλήματα αναγνώρισης προτύπων και να διαλύσουμε τις «αγαθές» ελπίδες κάποιων, ότι κάθε απλή προσέγγιση έχει τη δύναμη να λύσει όλα τα προβλήματα αναγνώρισης προτύπων. Οι μέθοδοι που παρουσιάζονται στο βιβλίο αυτό είναι πρωταρχικώς χρήσιμες για το βήμα της ταξινόμησης. Θα δούμε ότι αυτές έχουν σχέση με τα προβλήματα διαμερισμού, εξαγωγής χαρακτηριστικών και μετά-επεξεργασίας, τα οποία εξαρτώνται πολύ από την περιοχή. Πάντως, η απόδοση των

δύσκολων προβλημάτων αναγνώρισης προτύπων γενικά απαιτεί τη βέλτιστη εκμετάλλευση της οποιασδήποτε εξειδικευμένης γνώσης που αφορά το πεδίο.

1.5 Ο Κύκλος Σχεδίασης

Ο σχεδιασμός ενός συστήματος αναγνώρισης προτύπων (ΣΑΠ) συνήθως συνεπάγεται την επανάληψη ενός συνόλου διαφορετικών λειτουργιών: τη συλλογή δεδομένων, την επιλογή χαρακτηριστικών, την επιλογή μοντέλου που τα περιγράφει, την εκπαίδευση και τέλος την αποτίμηση. Σε αυτήν την ενότητα παρουσιάζεται μια γενική επισκόπηση του κύκλου σχεδίασης (Εικόνα 1.8) και εξετάζονται κάποια από τα προβλήματα που προκύπτουν συχνά.



Εικόνα 1.8

1.5.1 Συλλογή Δεδομένων

Η συλλογή των δεδομένων μπορεί να μετρήσει ιδιαίτερα σημαντικά στο συνολικό κόστος σχεδιασμού ενός ΣΑΠ. Είναι πιθανό να μπορεί κανείς να ελέγξει τις δυνατότητες ενός ΣΑΠ με ένα μικρό σύνολο από «τυπικά» παραδείγματα. Συνήθως όμως χρειάζονται πολλά περισσότερα δεδομένα για να επιτύχουμε μια καλή απόδοση του συστήματος μας. Πώς όμως γνωρίζουμε πότε έχουμε συγκεντρώσει ένα επαρκώς μεγάλο και αντιπροσωπευτικό σύνολο παραδειγμάτων για την εκπαίδευση και τη δοκιμή του συστήματός μας:

1.5.2 Επιλογή Χαρακτηριστικών

Η επιλογή των διακριτών χαρακτηριστικών είναι ένα κρίσιμο βήμα στη σχεδίαση και εξαρτάται από τα γενικότερα χαρακτηριστικά της περιοχής του προβλήματος. Η πρόσβαση σε χρήσιμα δεδομένα –όπως για παράδειγμα, εικόνες ψαριών πάνω στον ιμάντα μεταφοράς- θα είναι σίγουρα πολύτιμη στην επιλογή του συνόλου των χαρακτηριστικών. Πάντως η εκ των προτέρων γνώση παίζει επίσης πολύ σημαντικό ρόλο. Στο υποθετικό μας παράδειγμα η εκ των προτέρων γνώση για τη φωτεινότητα των διαφορετικών ειδών ψαριών βοήθησε στον σχεδιασμό του ταξινομητή αφού αποτέλεσε ένα πολλά υποσχόμενο χαρακτηριστικό.

1.5.3 Επιλογή Μοντέλου

Σε περίπτωση που δεν είμαστε ικανοποιημένοι με την απόδοση του ταξινομητή ψαριών που κατασκευάσαμε, όπως φαίνεται στα σχήματα 1.4 και 1.5, μπορούμε να μετακινηθούμε σε μια εντελώς διαφορετική κλάση μοντέλων. Η κλάση αυτή μπορεί για παράδειγμα να βασίζεται σε μια συνάρτηση του αριθμού και της θέσης των λεπιών, του χρώματος των ματιών, του βάρους, του σχήματος του στόματος κ.ά. Πώς γνωρίζουμε όμως πότε το μοντέλο που έχουμε υποθέσει διαφέρει σημαντικά από το πραγματικό μοντέλο, αυτό που υπόκειται στα πρότυπα μας, με αποτέλεσμα να χρειαζόμαστε ένα νέο μοντέλο; Εν συντομία, πώς μπορούμε να γνωρίζουμε πότε να απορρίπτουμε μια κλάση μοντέλων και να δοκιμάζουμε μία άλλη; Ως σχεδιαστές, είμαστε καταδικασμένοι να αναλωνόμαστε σε τυχαίες και κουραστικές δοκιμές και λάθη στην επιλογή των μοντέλων, μη γνωρίζοντας στην πραγματικότητα πότε να περιμένουμε βελτιωμένη απόδοση; Ή υπάρχουν μέθοδοι και κανόνες για να γνωρίζουμε πότε να προωθούμε μια κλάση και να απορρίπτουμε μία άλλη;

1.5.4 Εκπαίδευση

Γενικά η διαδικασία της χρησιμοποίησης των δεδομένων για τον καθορισμό του ταξινομητή αναφέρεται ως εκπαίδευση του ταξινομητή. Μεγάλο μέρος του παρόντος βιβλίου ασχολείται με τις πολλές διαφορετικές μεθόδους εκπαίδευσης ταξινομητών και επιλογής μοντέλων.

Είδαμε ήδη πολλά προβλήματα που προκύπτουν στην σχεδίαση ενός ΣΑΠ. Δεν έχουν βρεθεί δυστυχώς καθολικές μέθοδοι για την επίλυση όλων αυτών των προβλημάτων. Πάντως η επαναλαμβανόμενη εμπειρία των τελευταίων είκοσι χρόνων έχει δείξει ότι οι πιο αποδοτικές μέθοδοι για το σχεδιασμό ταξινομητών εμπεριέχουν την εκμάθηση που βασίζεται σε δείγματα προτύπων.

1.5.5 Αποτίμηση

Όταν από τη χρήση ενός χαρακτηριστικού προτιμήσαμε τη χρήση δύο στο πρόβλημα της ταξινόμησης των ψαριών, ήταν απολύτως απαραίτητη η αποτίμηση της συχνότητας λάθους για να δούμε ότι η χρήση ενός χαρακτηριστικού ήταν ανεπαρκής και ότι ήταν δυνατό να τα πάμε καλύτερα. Όταν μεταφερθήκαμε από τον απλό ταξινομητή ευθείας γραμμής του σχήματος 1.4 στον πιο πολύπλοκο μοντέλο του σχήματος 1.5 ήταν πάλι το αποτέλεσμα της αποτίμησης του λάθους που μας οδήγησε στο ότι μπορούσαμε να επιτύχουμε ακόμη καλύτερα αποτελέσματα ταξινόμησης. Η αποτίμηση είναι λοιπόν, σημαντική όχι μόνο για να μετρήσουμε την απόδοση του ΣΑΠ αλλά και για να αναγνωρίσουμε την δυνατότητα βελτιώσεων των συστατικών του.

Ενώ ένα πολύπλοκο σύστημα μπορεί να επιτύχει τέλεια ταξινόμηση των δειγμάτων εκπαίδευσης, είναι μάλλον απίθανο να έχει την ίδια επιτυχία σε καινούρια πρότυπα (για τα οποία δεν έχει εκπαιδευτεί). Η κατάσταση αυτή είναι γνωστή ως υπερ-εκπαίδευση (overfitting). Ένα από τα σημαντικότερα πεδία ερευνών στην στατιστική αναγνώριση προτύπων είναι να καθοριστεί ο τρόπος με τον οποίο θα προσαρμόσουμε την πολυπλοκότητα του μοντέλου –δεν θέλουμε ένα πολύ απλό μοντέλο το οποίο δεν θα μπορεί να διακρίνει τις διαφορές μεταξύ των κατηγοριών, αλλά ούτε και ένα πολύ σύνθετο το οποίο δεν θα κάνει καλή ταξινόμηση σε άγνωστα, νέα για το σύστημα, πρότυπα. Υπάρχουν μέθοδοι για την εύρεση της καλύτερης (μέσης) πολυπλοκότητας για την ταξινόμηση.

1.5.6 Υπολογιστική Πολυπλοκότητα

Κάποια προβλήματα αναγνώρισης προτύπων μπορούν να «επιλυθούν» με τη χρήση αλγορίθμων που είναι όμως μη πρακτικοί. Για παράδειγμα, μπορεί να θέλουμε να θέσουμε ετικέτες σε όλες τις εικόνες των 20x20 binary pixels με μια ετικέτα κατηγορίας για γραφική οπτική αναγνώριση χαρακτήρων (OCR) και να χρησιμοποιήσουμε ένα πίνακα, έτσι ώστε να κοιτάμε τις ετικέτες και στη συνέχεια να ταξινομούμε τα εισερχόμενα πρότυπα. Ενώ θεωρητικά μπορούμε να επιτύχουμε μηδενικό λάθος στην αναγνώριση, ο χρόνος για την «ετικετοποίηση» και οι απαιτήσεις σε αποθηκευτικό χώρο γι' αυτές μπορεί να είναι απαγορευτικά στοιχεία, αφού θα απαιτούνταν «ετικετοποίηση» καθενός από τα $2^{20 \times 20} = 10^{120}$ πρότυπα. Έτσι, γίνεται σαφές ότι πρέπει να παίρνονται σοβαρά υπόψη οι υπολογιστικοί πόροι που είναι απαραίτητοι και η υπολογιστική πολυπλοκότητα των διαφόρων αλγορίθμων.

1.6 Μάθηση και Προσαρμοστικότητα

Με την ευρεία έννοια κάθε μέθοδος που ενσωματώνει πληροφορίες από δείγματα εκπαίδευσης κατά το σχεδιασμό του ταξινομητή έχει να κάνει με τη μάθηση. Επειδή σχεδόν όλα τα πρακτικά ή ενδιαφέροντα προβλήματα αναγνώρισης προτύπων είναι τόσο δύσκολα που δεν μπορούμε να μαντέψουμε την καλύτερη απόφαση ταξινόμησης εκ των προτέρων (a priori), θα πρέπει να ξοδέσουμε τον περισσότερο από το χρόνο μας λαμβάνοντας υπόψη τις διαδικασίες μάθησης. Η μάθηση αναφέρεται στην εύρεση και χρήση κάποιου αλγόριθμου για την ελαχιστοποίηση του λάθους σε ένα συγκεκριμένο σύνολο δεδομένων εκπαίδευσης. Μια μεγάλη κατηγορία αλγορίθμων ταχύτερης καθόδου (gradient descent) που αλλάζουν τις παραμέτρους του ταξινομητή, με σκοπό να μειώσουν το λάθος, εισέρχονται τελευταία στο χώρο της στατιστικής αναγνώρισης προτύπων. Σε αυτούς αξίζει να δώσουμε ένα σημαντικό μέρος της προσοχής μας. Η διαδικασία μάθησης «εμφανίζεται» σε αρκετές γενικές μορφές, όπως θα δούμε και στη συνέχεια.

1.6.1 Επιβλεπόμενη μάθηση

Στην επιβλεπόμενη μάθηση ένας επιβλέπων δίνει ένα κόστος ή μια ετικέτα κατηγορίας σε κάθε πρότυπο ενός συνόλου εκπαίδευσης και ψάχνει να μειώσει το συνολικό κόστος γι' αυτά τα πρότυπα.

1.6.2 Μη Επιβλεπόμενη μάθηση

Στην μη επιβλεπόμενη μάθηση ή αλλιώς ομαδοποίηση (clustering) δεν υπάρχει επιβλέπων και το σύστημα σχηματίζει συστοιχίες (clusters) ή «φυσικούς σχηματισμούς ομάδων»

πάνω στα πρότυπα εισόδου. Ο όρος «φυσικός» ορίζεται πάντα άμεσα ή έμμεσα από το ίδιο το σύστημα ομαδοποίησης και δοθέντος ενός συγκεκριμένου συνόλου προτύπων ή συναρτήσεων κόστους, διαφορετικοί αλγόριθμοι ομαδοποίησης οδηγούν σε διαφορετικές συστοιχίες.

1.6.3 Ενισχυτική Μάθηση

Ο πιο συνήθης τρόπος για την εκπαίδευση ενός ταξινομητή είναι να βάλουμε μια είσοδο, να υπολογίσουμε την δοκιμαστική ετικέτα κατηγορίας, στην οποία μπορεί να ανήκει, και να χρησιμοποιήσουμε την γνωστή ετικέτα κατηγορίας για να βελτιώσουμε τον ταξινομητή. Για παράδειγμα σε ένα πρόβλημα OCR, η είσοδος μπορεί να είναι μια εικόνα ενός χαρακτήρα (γράμματος ή συμβόλου), η έξοδος του ταξινομητή να είναι η κατηγορία με ετικέτα 'R', ενώ η επιθυμητή έξοδος να είναι το 'B'. Στην ενισχυτική μάθηση ή αλλιώς μάθηση με κριτική, κανένα επιθυμητό «σήμα» κατηγορίας δεν δίνεται. Απεναντίας η μόνη επιβλεπόμενη ανατροφοδότηση είναι το αν όντως η δοκιμαστική κατηγορία είναι η σωστή ή όχι. Αυτό είναι ανάλογο με έναν κριτή, ο οποίος απλώς κρίνει αν κάτι είναι σωστό ή όχι, αλλά δεν υποδεικνύει ακριβώς το είδος του λάθους. Στην ταξινόμηση προτύπων είναι σύνηθες η ενισχυτική αυτή μάθηση να έχει δυαδικό χαρακτήρα (αν είναι σωστό 1, αν είναι λάθος 0). Πώς μπορεί ένα σύστημα να μάθει από μια τέτοια -όχι συγκεκριμένη- ανατροφοδότηση;

1.7 Βιβλιογραφία

- [1] Aristotle, Robin Waterfield. and David Bostock. Physics. Oxford University Press. Oxford, UK, 1996.
- [2] Allan Bloom. The Republic of Plato. Basic Books, New York, second edition, 1991.
- [3] Bodhidharma. The Zen Teachings of Bodhidharma. North Point Press, San Francisco, CA. 1989
- [4] Mikhail M. Bongard. Pattern Recognition. Spartan Books, Washington, D.C., 1970.
- [5] Chi-hau Chen, Louis Francois Pan, and Patrick S. P. Wang, editors. Handbook of Pattern Recognition & Computer Vision. World Scientific, Singapore, second edition, 1993.
- [6] Luc Devroye. Laszlo Gyorfi. and Gabor Lugosi. A Probabilistic Theory of Pattern Recognition. Springer, New York, 1996.
- [7] Many Fischler and Oscar Firschein. Readings in Computer Vision; Issues, Problems. Principles and Paradigms. Morgan Kaufmann, San Mateo, CA, 1987.
- [8] Keinosuke Fukunaga. Introduction to Statistical Pattern Recognition. Academic Press, New York, second edition, 1990.
- [9] John Hertz, Anders Krogh, and Richard G. Palmer. Introduction to the Theory of Neural Computation. Addison-Wesley, Redwood City, CA, 1991.
- [10] Douglas Hofstadter. Godel. Escher, Bach: An EternalGolden Braid. Basic Books, New York, 1979.
- [11] Eric R. Kandel and James H. Schwartz. Principles of Neural Science. Elsevier, New York, second edition, 1985.
- [12] Immanuel Kant. Critique of Pure Reason. Prometheus Books, New York, 1990.
- [13] George F. Luger. Cognitive Science: The Science of Intelligent Systems. Academic Press, New York, 1994.
- [14] Howard Margolis. Patterns, Thinking, and Cognition: A Theory of Judgment. University of Chicago Press. Chicago, IL, 1987.
- [15] Karl Raimund Popper. Popper Selections. Princeton University Press, Princeton. NJ, 1985.

- [16] Lawrence Rabiner and Biing-Hwang Juang. Fundamentals of Speech Recognition. Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [17] Jude W. Shavlik and Thomas G. Dietterich, editors. Readings in Machine Learning. Morgan Kaufmann, San Maieo, CA, 1990.
- [18] Brian Cantwell Smith. On the Origin of Objects. MIT Press, Cambridge. MA, 1996.
- [19] Louise Stark and Kevin Bowyer. Generic Object Recognition Using Form & Function. World Scientific, River Edge, NJ, 1996.
- [20] Donald R. Tsvet. The Pattern Recognition Basis of Artificial Intelligence. IEEE Press, New York, 1998.
- [21] William R. Uttal. The Psychology of Sensory Coding. HarperCollins, New York, 1973.
- [22] Satoshi Watanabe. Knowing and Guessing: A Quantitative Study of Inference and Information. Wiley, New York, 1969.

