

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Πτυχιακή Εργασία

Θέμα: «Βαθιά Μάθηση σε ανθρωπιστικά Δεδομένα»

Ονοματεπώνυμο: Μαρία Δήμα

ΑΜ: Π2013029

Επιβλέπων: Βλάμος Π.

Κέρκυρα, Σεπτέμβριος 2021

Περιεχόμενα

Κεφάλαιο 1: Εισαγωγή	7
1.1 Αντικείμενο της εργασίας	7
1.2 Στόχος Εργασίας	10
1.3 Σημαντικότητα Εργασίας	11
1.4 Μεθοδολογία της εργασίας	16
Κεφάλαιο 2: Έννοιες Εξόρυξης γνώσης	17
2.1 Εξόρυξη Δεδομένων	17
2.1.1 Ορισμός εξόρυξης Δεδομένων	17
2.1.2 Βάσεις Δεδομένων	18
2.1.3 Φάσεις Εξόρυξης Γνώσης	20
2.1.3 Τεχνικές Εξόρυξης δεδομένων	24
2.2 Εξόρυξη Μουσικής Πληροφορίας	25
2.2.1 Πηγές Μουσικής Πληροφορίας και ορισμός Εξόρυξης Μουσικής Πληροφορίας	25
2.2.2 Επισκόπηση Εξόρυξης Μουσικής πληροφορίας	27
2.2.3 Διαχείριση Μουσικών Δεδομένων	27
2.2.4 Οπτικοποίηση Μουσικής	28
2.2.5 Εξόρυξη Συσχετίσεων	28
2.2.6 Εξόρυξη ακολουθίας	29
2.2.7 Ταξινόμηση	29
2.2.8 Ομαδοποίηση	30
2.2.9 Σύνοψη Μουσικής	30
2.3 Ανάκτηση Μουσικής Πληροφορίας	30
2.3.1 Γενικά	30
2.3.2 Αναζήτηση Ομοιότητας Μουσικής	31

2.4 Μηχανική Μάθηση	32
2.4.1 Ορισμός	32
2.4.2 Διαφορά Μηχανικής Μάθησης με τον Παραδοσιακό Προγραμματισμό	33
2.4.3 Λειτουργία μάθησης για έναν υπολογιστή	33
2.4.4 Κατηγοριοποίηση της Μηχανικής Μάθησης	34
2.4.5 Τα δεδομένα στην μηχανική μάθηση	36
Κεφάλαιο 3: Αλγόριθμοι.....	39
3.1 Κατηγοριοποίηση.....	39
3.1.2 Ορισμός	39
3.1.2 Βήματα Κατηγοριοποίησης.....	39
3.1.3 Εκπαίδευση και δοκιμή	39
3.1.4 Συντακτικό.....	41
3.1.5 Ταξινομητές Μηχανικής Μάθησης και εργαλεία.....	42
3.1.6 Εφαρμογές	43
3.1.7 Πλεονεκτήματα και Μειονεκτήματα.....	43
3.2 Δένδρα Αποφάσεων	44
3.2.1 Βασικά Στοιχεία	44
3.2.2 Αλγόριθμος ID3.....	48
3.3 Ταξινομητές Bayes.....	54
3.1.1 Βασικά Στοιχεία	54
3.3.2 Θεώρημα Bayes.....	56
3.3.3 Αφελής Υπόθεση.....	56
3.3.3 Πλεονεκτήματα και Μειονεκτήματα.....	58
3.4 Δίκτυα Bayesian.....	58
4.4.1 Πιθανότητες.....	59

4.4.2 Περιγραφή του δικτύου	59
4.4.3 Συμπερασμοί.....	60
4.4.4 Πλεονεκτήματα και Μειονεκτήματα	61
3.5 Νευρωνικά Δίκτυα	62
3.5.1 Βασικά	62
3.5.2 Βασικά συστατικά ενός Νευρωνικού Δικτύου	63
3.5.3 Από την είσοδο στην έξοδο	65
3.5.4 Εκπαίδευση Νευρωνικού Δικτύου	67
3.5.5 Πλεονεκτήματα και Μειονεκτήματα Νευρωνικών Δικτύων.....	72
3.6 Αλγόριθμος KNN (K-Nearest Neighbors)	73
3.6.1 Ο Αλγόριθμος.....	73
3.6.2 Επιλογή του κατάλληλου K.....	74
3.6.3 Πλεονεκτήματα και Μειονεκτήματα	75
Κεφάλαιο 4: Εξόρυξη γνώσης από Μουσικά Δεδομένα	76
4.1 Αναπαράσταση Μουσικής Πληροφορίας	76
4.1.1 Εισαγωγικά Στοιχεία	76
4.1.2 Αναπαράσταση Τραγουδιών	78
4.2 Εξόρυξη Μουσικής Πληροφορίας	79
4.2.1 Βαθιά Μάθηση για ανάλυση μουσικού περιεχομένου	79
4.2.2 Deep Belief Networks	80
4.2.3 Convolutional Neural Networks – CNN	84
4.2.4 Βαθιά Νευρωνικά Δίκτυα (DNN – Deep Neural Networks).....	87
Βιβλιογραφία	90

Πρόλογος

Κεφάλαιο 1: Εισαγωγή

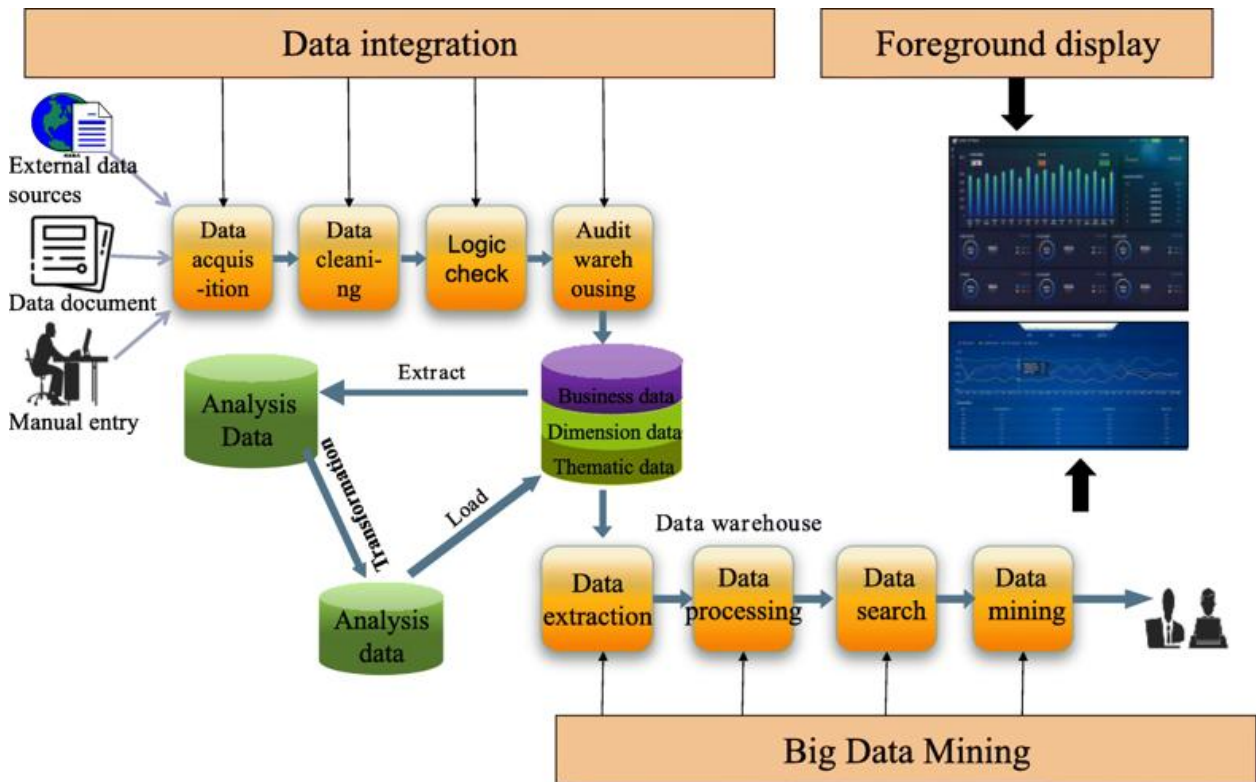
1.1 Αντικείμενο της εργασίας

Αντικείμενο της παρούσας εργασίας είναι να γίνει περιγραφή της χρήσης της βαθιάς μάθησης για την εξαγωγή ανθρωπιστικών δεδομένων. Μερικές από τις εφαρμογές της βαθιάς μάθησης σε ανθρωπιστικά δεδομένα είναι:

1. Εξαγωγή συναισθήματος από τον κοινωνικό ιστό:

Οι άνθρωποι κρίνουν τον κόσμο γύρω τους όταν ζουν σε μια κοινωνία. Έχουν θετικές και αρνητικές απόψεις για ανθρώπους, προϊόντα, μέρη και γεγονότα. Αυτές οι απόψεις μπορούν να χαρακτηριστούν ως συναισθήματα. Η ανάλυση συναισθήματος είναι η μελέτη αυτοματοποιημένων τεχνικών για εξαγωγή συναισθήματος από γραπτό κείμενο.

2. Εξόρυξη μουσικής πληροφορίας



Εικόνα 1: Εξόρυξη μουσικής Πληροφορίας

Στην εποχή του διαδικτύου ένας τεράστιος όγκος πληροφοριών είναι εύκολα προσβάσιμος. Παραδείγματα μουσικής πληροφορίας που μπορούν να βρεθούν στο διαδίκτυο είναι τα εξής:

- 1) Πληροφορίες για μουσική και καλλιτέχνες σε ιστοσελίδες δισκογραφικών εταιρειών
- 2) Στίχοι τραγουδιών μπορούν να κατέβουν από ιστοσελίδες βάσεων δεδομένων στίχων
- 3) Μουσικές κριτικές είναι διαθέσιμες σε διάφορες ιστοσελίδες στον παγκόσμιο ιστό (TaoLi, 2010)

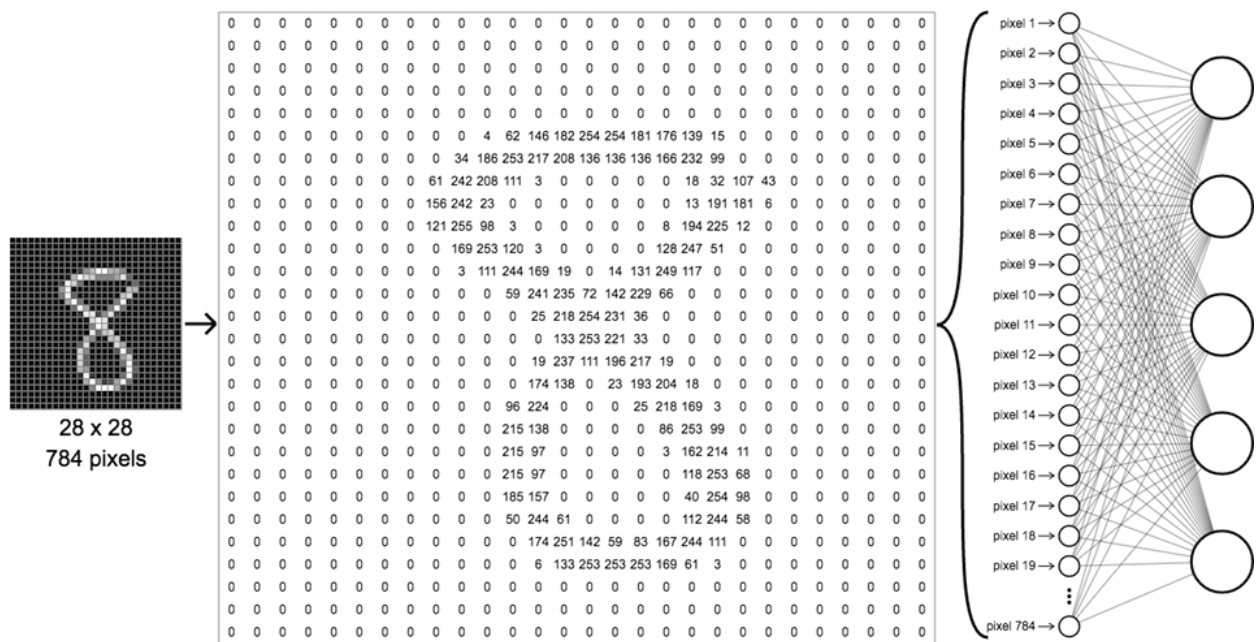
3. Αναγνώριση χειρόγραφων χαρακτήρων

Όπως το λέει και ο τίτλος με την αναγνώριση χειρόγραφων χαρακτήρων μπορούμε να αναπτύξουμε ένα σύστημα, με την χρήση νευρωνικών δικτύων, για την αναγνώριση χειρόγραφων χαρακτήρων. Για παράδειγμα δοθέντος ενός χειρόγραφου αριθμού π.χ. του 1 το σύστημα πρέπει να αναγνωρίσει ότι ο αριθμός που δόθηκε είναι το 1 και όχι κάποιος άλλος.

Για να γίνει καλύτερη κατανόηση των παραπάνω όρων πρέπει να γίνει αναφορά στις έννοιες της μηχανικής μάθησης, της βαθιάς μάθησης, των νευρωνικών δικτύων αλλά και της εξόρυξης δεδομένων.

- **Μηχανική Μάθηση** είναι ο τομέας, ο οποίος δίνει την δυνατότητα στους ηλεκτρονικούς υπολογιστές να μάθουν χωρίς να έχουν προγραμματιστεί ρητώς. Με βάση τα παραπάνω γίνεται κατανοητό ότι με χρήση της μηχανικής μάθησης οι ηλεκτρονικοί υπολογιστές αποκτούν ένα χαρακτηριστικό που μέχρι προσφάτως ήταν προτέρημα των ανθρώπων: την ικανότητα να μάθουν.(geeksforgeeks, n.d.)
- Η **Βαθιά Μάθηση** είναι μια τεχνική μηχανικής μάθησης η οποία μαθαίνει στους υπολογιστές να κάνουν αυτό που στους ανθρώπους συμβαίνει φυσικά: να μάθουν με παραδείγματα. Στην βαθιά μάθηση ένας υπολογιστής μαθαίνει να εκτελεί εργασίες ταξινόμησης από εικόνες, κείμενο και ήχο. Τα μοντέλα της βαθιάς μάθησης έχουν εκπληκτική ακρίβεια, η οποία μερικές φορές ξεπερνάει και τον άνθρωπο. Τα μοντέλα εκπαιδεύονται χρησιμοποιώντας ένα μεγάλο σύνολο από δεδομένα με ετικέτες χρησιμοποιώντας νευρωνικά δίκτυα.(What Is Deep Learning?, n.d.)
- **Νευρωνικά Δίκτυα** είναι ένα σύνολο από αλγόριθμους, οι οποίοι βασίζονται, ως ένα σημείο, στον ανθρώπινο εγκέφαλο. Οι αλγόριθμοι αυτοί είναι σχεδιασμένοι για να αναγνωρίσουν πρότυπα. Ερμηνεύουν τα αισθητήρια δεδομένα μέσω ενός είδους αντίληψης μηχανής, σήμανσης ή ομαδοποίησης της εισόδου. Τα πρότυπα που αναγνωρίζουν οι αλγόριθμοι είναι αριθμητικά, τα οποία περιέχονται σε πίνακες στους οποίους όλα τα δεδομένα (εικόνα, ήχος, κείμενο), πρέπει να μετατρέπονται. Τα νευρωνικά δίκτυα μας βοηθάνε να ομαδοποιούμε και να ταξινομούμε δεδομένα. Στην ουσία μας βοηθούν να ομαδοποιήσουμε δεδομένα χωρίς ετικέτα (δηλαδή δεδομένα τα οποία αρχικά δεν ξέρουμε τι αναπαριστούν) με βάση ομοιότητες που παρατηρούνται με τα δεδομένα που τροφοδοτούνται στο νευρωνικό δίκτυο. Επιπλέον, τα νευρωνικά δίκτυα μας βοηθάνε στην ταξινόμηση των δεδομένων όταν βέβαια υπάρχει ένα σύνολο δεδομένων με ετικέτα για να εκπαιδευτούν. Τέλος τα νευρωνικά δίκτυα μπορούν να εξάγουν χαρακτηριστικά τα οποία τροφοδοτούνται σε άλλους αλγορίθμους για ομαδοποίηση και ταξινόμηση. Με βάση τα παραπάνω, τα νευρωνικά δίκτυα μπορούν να θεωρηθούν σαν συνιστώσες σε μεγαλύτερες εφαρμογές μηχανικής μάθησης, οι οποίες

συμπεριλαμβάνουν αλγορίθμους για ενίσχυση της μάθησης και ομαδοποίησης.(Nicholson, n.d.)



Εικόνα 2: Νευρωνικό Δίκτυο για αναγνώριση χειρόγραφου χαρακτήρα

- **Εξόρυξη δεδομένων** καλούμε την διαδικασία της ανάλυσης κρυμμένων προτύπων δεδομένων σύμφωνα με διαφορετικές προοπτικές για κατηγοριοποίηση σε χρήσιμα δεδομένα τα οποία συλλέγονται σε διάφορες περιοχές όπως π.χ. «αποθήκες δεδομένων» με σκοπό την αποτελεσματική ανάλυση και την διευκόλυνση της λήψης επιχειρηματικών αποφάσεων αλλά και άλλες απαιτήσεις πληροφοριών, με στόχο την μείωση του κόστους και την αύξηση των εσόδων. Η εξόρυξη γνώσης είναι γνωστή και ως ανακάλυψη πληροφορίας και ανακάλυψη γνώσης.(Twin, 2019)

Τις παραπάνω έννοιες θα τις αναλύσουμε σε επόμενα κεφάλαια.

1.2 Στόχος Εργασίας

Στόχος της συγκεκριμένης εργασίας είναι η ενημέρωση του αναγνώστη για την χρήση της εξόρυξης γνώσης και της βαθιάς μάθησης σε ανθρωπιστικά δεδομένα. Θα γίνει προσπάθεια η εργασία να γραφτεί με όσο το δυνατόν πιο απλά λόγια ούτως ώστε να μπορεί να την

παρακολουθήσει ο οποιοσδήποτε, ακόμα και κάποιος που δεν έχει κάποια σχέση με την πληροφορική και την επιστήμη των υπολογιστών.

1.3 Σημαντικότητα Εργασίας

Τα τελευταία χρόνια έχουμε παρατηρήσει μια αλματώδη ανάπτυξη στην τεχνητή νοημοσύνη και στην μηχανική μάθηση. Με τον τεράστιο όγκο των δεδομένων που έχουμε διαθέσιμα την σημερινή εποχή, η ανάπτυξη συστημάτων που θα κάνει εξόρυξη μουσικής πληροφορίας αλλά και θα εξάγει συναίσθημα από τα κοινωνικά δίκτυα είναι πολύ σημαντικό. Έτσι η εργασία αυτή θα ενημερώσει τους χρήστες για 3 σημαντικά θέματα της τεχνητής νοημοσύνης.

Συγκεκριμένα για την εξαγωγή συναισθήματος:

Blogs, forums στο διαδίκτυο αλλά και ιστοσελίδες κοινωνικής δικτύωσης όπως είναι το Facebook και το twitter μπορούν να θεωρηθούν μέσα κοινωνικής δικτύωσης. Αυτά τα μέσα κοινωνικής δικτύωσης μπορούν να χρησιμοποιηθούν για την διάδοση των απόψεων χιλιάδων αλλά και εκατομμυρίων ανθρώπων. Τα μέσα κοινωνικής δικτύωσης χρησιμοποιούνται όλο και περισσότερο για πηγές πληροφοριών για τις επιχειρήσεις. Από την άλλη οι άνθρωποι είναι διατεθειμένοι να μοιραστούν γεγονότα για την ζωή τους, γνώση, εμπειρίες και σκέψεις με ολόκληρο τον κόσμο χρησιμοποιώντας τα μέσα κοινωνικής δικτύωσης. Συμμετέχουν ενεργά σε εκδηλώσεις με το να εκφράζουν την άποψή τους. Αυτός ο τρόπος μοιράσματος των γνώσεων και των συναισθημάτων με την κοινωνία έχει ως αποτέλεσμα οι επιχειρήσεις να συλλέγουν όλο και περισσότερες πληροφορίες για τα προϊόντα τους. Επίσης τις βοηθάει στο να μάθουν πόσο καλή φήμη έχουν στην κοινωνία και να παίρνουν αποφάσεις για την εύρυθμη λειτουργία τους .

Με βάση τα παραπάνω είναι πλέον φανερό ότι η ανάλυση συναισθήματος είναι ένας βασικός παράγοντας για την καινοτόμα διαχείριση πελατειακής εμπειρίας. Επιπλέον η ανάλυση συναισθήματος στα κοινωνικά μέσα βοηθάει τις επιχειρήσεις που θέλουν να προωθήσουν τα προϊόντα τους, να βρουν νέες ευκαιρίες εξέλιξης αλλά και να διαχειριστούν την φήμη τους. Στην σημερινή εποχή με την εξέλιξη του διαδικτύου το να έχει κάποιος πρόσβαση σε έναν μεγάλο αριθμό πληροφοριών δεν είναι πια ζήτημα. Είναι αδιαμφισβήτητο ότι η πληροφορία έχει γίνει το κύριο μέσο ανταλλαγής για πολλές επιχειρήσεις. Αυτό έχει σαν αποτέλεσμα, η ανάλυση συναισθήματος να είναι εξέχουσας

σημασίας διότι ο τρόπος αυτός μας δίνει μηχανισμούς για την αναζήτηση και την ανάκτηση σχετικών δεδομένων και πληροφοριών για τους λόγους που αναφέραμε παραπάνω.(R. A. S. C. Jayasanka, 2013)

Όσον αφορά την εξόρυξη μουσικής πληροφορίας:

Η μουσική παίζει σημαντικό ρόλο στην καθημερινή ζωή για πολλούς ανθρώπους. Με την ψηφιοποίηση, μεγάλες συλλογές από μουσικά δεδομένα έχουν δημιουργηθεί και συλλέγονται από μουσικόφιλους. Αυτό έχει σαν αποτέλεσμα την δημιουργία συλλογών, όχι μόνο στην μορφή μουσικών CD, αλλά και αποθηκευμένες σε σκληρούς δίσκους και στο διαδίκτυο. Έτσι οι μουσικές συλλογές είναι τόσο μεγάλες που είναι φυσικά αδύνατον για τους ανθρώπους να παρακολουθούν την μουσική που υπάρχει αλλά και τις πιθανές σχέσεις μεταξύ των διαφορετικών τραγουδιών. Αυτό το πρόβλημα λύνεται με την χρήση μηχανικής μάθησης και εξόρυξης μουσικών δεδομένων.(Hall, 2012)

Τέλος όσον αφορά την σημαντικότητα της συγκεκριμένης εργασίας εκτός από αυτά που αναφέρθηκαν παραπάνω, γίνεται αναφορά και στην μηχανική μάθηση την οποία την συναντάμε συνέχεια στην επιστήμη των υπολογιστών. Συγκεκριμένα η μηχανική μάθηση έχει εφαρμογές οι οποίες έχουν φέρει την επανάσταση στην τεχνολογία. Μερικές από αυτές είναι:

1) Αναγνώριση Εικόνας

Η αναγνώριση εικόνας είναι το πιο σημαντικό παράδειγμα μηχανικής μάθησης. Βασικά, είναι μια μέθοδος για την αναγνώριση ενός χαρακτηριστικού σε μια ψηφιακή εικόνα. Επιπροσθέτως, η τεχνική αυτή μπορεί να χρησιμοποιηθεί για περαιτέρω ανάλυση όπως είναι η αναγνώριση προτύπου, αναγνώριση προσώπου, οπτική αναγνώριση χαρακτήρων και πολλά άλλα. Παρόλο που πολλές τεχνικές για την αναγνώριση εικόνας είναι διαθέσιμες, προτιμάται η μηχανική μάθηση.

2) Ταξινόμηση Ειδήσεων

Η ταξινόμηση ειδήσεων είναι άλλη μια περίπτωση χρήσης μηχανικής μάθησης. Όπως έχουμε αναφέρει αρκετές φορές στην παρούσα εργασία, ο όγκος των δεδομένων που είναι διαθέσιμος στο διαδίκτυο έχει αυξηθεί κατακόρυφα. Όμως, κάθε άνθρωπος έχει τα δικά του ενδιαφέροντα. Αυτό έχει σαν αποτέλεσμα, η επιλογή κατάλληλης πληροφορίας από αυτό τον όγκο των δεδομένων να έχει μετατραπεί σε μια επίπονη διαδικασία για τους χρήστες. Με την παροχή

αυτόματων κατηγοριών νέων στους αναγνώστες θα αυξηθεί η επισκεψιμότητα των ειδησεογραφικών ιστοσελίδων. Επιπλέον, οι αναγνώστες έχουν την δυνατότητα να ψάξουν για συγκεκριμένες ειδήσεις με αποτελεσματικό τρόπο.



Εικόνα 3: Ταξινόμηση Ειδήσεων

3) Βίντεο Παρακολούθησης

Ένα μικρό βίντεο περιέχει περισσότερη πληροφορία σε σχέση με έγγραφα κειμένου και άλλα αρχεία πολυμέσων όπως είναι ο ήχος και η εικόνα. Για τον λόγο αυτό, η εξαγωγή χρήσιμων πληροφοριών από ένα βίντεο, όπως για παράδειγμα ένα αυτοματοποιημένο σύστημα παρακολούθησης, είναι ένα σημαντικό ζήτημα. Η παρουσία ενός ανθρώπου σε ένα διαφορετικό πλαίσιο σε ένα βίντεο είναι ένα κοινό σενάριο. Σε μια εφαρμογή βασισμένη στην ασφάλεια, η ταυτοποίηση ενός ανθρώπου από τα βίντεο είναι ένα σημαντικό θέμα. Το μοτίβο του προσώπου είναι η πιο ευρέως χρησιμοποιούμενη παράμετρος για την αναγνώριση του ατόμου.

4) Ταξινόμηση email και φιλτράρισμα ανεπιθύμητων

Υπάρχουν αλγόριθμοι μηχανικής μάθησης οι οποίοι ταξινομούν ένα email και φιλτράρουν ανεπιθύμητα μηνύματα. Υπάρχουν πολλές τεχνικές που χρησιμοποιούνται για το φιλτράρισμα ανεπιθύμητων email.

5) Αναγνώριση Ομιλίας

Η αναγνώριση ομιλίας είναι η διαδικασία της μετατροπής ομιλούμενων λέξεων σε κείμενο. Επιπλέον καλείται και αυτόματη αναγνώριση ομιλίας, αναγνώριση ομιλίας υπολογιστή ή ομιλία σε κείμενο. Αυτός ο τομέας έχει ευεργετηθεί από την εξέλιξη της μηχανικής μάθησης και των μεγάλων δεδομένων. Προς το παρόν, όλα τα συστήματα αναγνώρισης ομιλίας που χρησιμοποιούνται για διαφημιστικούς σκοπούς κάνουν χρήση μηχανική μάθησης επειδή τα συστήματα αυτά έχουν καλύτερη επίδοση από τα κλασσικά συστήματα αναγνώρισης ομιλίας.



Εικόνα 4: Αναγνώριση Ομιλίας

6) Διαδικτυακή Ανίχνευση Απάτης

Η διαδικτυακή ανίχνευση απάτης είναι μια προχωρημένη εφαρμογή των αλγορίθμων μηχανικής μάθησης. Αυτή η προσέγγιση παρέχει κυβερνοασφάλεια στους χρήστες με αποτελεσματικό τρόπο. Πρόσφατα, η PayPal χρησιμοποίησε έναν αλγόριθμο μηχανικής μάθησης και τεχνητής νοημοσύνης για να ανιχνεύσει ξέπλυμα βρώμικου χρήματος. Αυτό βοηθά την μείωση του κόστους και την αύξηση του κέρδους.



7) Αναγνώριση Συγγραφέα

Με την ανάπτυξη του διαδικτύου, η παράνομη χρήση online μηνυμάτων για ακατάλληλους ή παράνομους σκοπούς έχει γίνει μια μεγάλη ανησυχία για την κοινωνία. Για αυτό τον λόγο είναι απαραίτητη η πιστοποίηση του συγγραφέα. Ένα σύστημα αναγνώρισης συγγραφέα μπορεί να χρησιμοποιήσει πολλά πεδία όπως το ποινικό μητρώο, την ακαδημαϊκή κοινότητα και την ανθρωπολογία.

8) Πρόβλεψη

Η πρόβλεψη είναι η διαδικασία του να ειπωθεί κάτι το οποίο να βασίζεται σε προηγούμενο γεγονός. Μπορεί να είναι πρόβλεψη του καιρού, πρόβλεψη της κίνησης στους δρόμους και πολλά άλλα. Όλες οι προβλέψεις μπορούν να γίνουν χρησιμοποιώντας μηχανική μάθηση.

9) Υπηρεσίες στα Μέσα Κοινωνικής Δικτύωσης

Τα μέσα κοινωνικής δικτύωσης χρησιμοποιούν μηχανική μάθηση για να δημιουργήσουν ελκυστικά χαρακτηριστικά. Μερικά από αυτά είναι: άτομα που μπορεί να ξέρετε, προτάσεις, επιλογές αντίδρασης για τους χρήστες. Για παράδειγμα το Facebook συνεχώς παρακολουθεί την δραστηριότητα ενός ατόμου, όπως είναι με ποιον συνομιλεί, τα Like του, τον τόπο εργασίας και τον τόπο σπουδών. Με βάση αυτά τα δεδομένα και χρησιμοποιώντας αλγορίθμους μηχανικής μάθησης το Facebook δίνει προτάσεις βασισμένες σε αυτή την δραστηριότητα.

10) Ιατρικές Υπηρεσίες

Τα εργαλεία και οι αλγόριθμοι της μηχανικής μάθησης χρησιμοποιούνται εκτεταμένα στην ιατρική. Για παράδειγμα χρησιμοποιείται για την αναγνώριση μιας ασθένειας, για τον σχεδιασμό μιας θεραπείας, πρόβλεψη της εξέλιξης μιας ασθένεια κ.α.

Φυσικά η μηχανική μάθηση δεν χρησιμοποιείται μόνο στους παραπάνω τομείς αλλά και σε πολλούς άλλους. Μπορεί το θέμα της εργασίας να μην είναι η ανάλυση όλων των παραπάνω μεθόδων αλλά το σημαντικό είναι ότι με την μελέτη της εργασίας να παρακινήσει τον αναγνώστη να εντρυφήσει περεταίρω σε αυτό τον άκρως ενδιαφέροντα κλάδο της τεχνητής νοημοσύνης. (Hasan, n.d.)

1.4 Μεθοδολογία της εργασίας

Για την ανάπτυξη της εργασίας θα χρησιμοποιηθούν πηγές από το διαδίκτυο. Θα γίνει αναζήτηση πηγών και βιβλίων σχετικών με το θέμα στο διαδίκτυο έτσι ώστε να γίνει μια σωστή και επιστημονικά τεκμηριωμένη εργασία.

Κεφάλαιο 2: Έννοιές Εξόρυξης γνώσης

Στο κεφάλαιο αυτό θα αναπτύξουμε την έννοια της εξόρυξης γνώσης, και τα βήματα που περιλαμβάνει η τεχνική της εξόρυξης γνώσης. Στην συνέχεια θα αναφερθούμε στις τεχνικές ανάπτυξης μουσικής πληροφορίας και τέλος θα μιλήσουμε για έναν σημαντικό κλάδο της τεχνητής νοημοσύνης: την μηχανική μάθηση.

2.1 Εξόρυξη Δεδομένων

2.1.1 Ορισμός εξόρυξης Δεδομένων

Η **εξόρυξη Δεδομένων** ψάχνει για κρυφά, έγκυρα αλλά και πιθανώς χρήσιμα πρότυπα σε μεγάλα σύνολα δεδομένων. Στην ουσία ανακαλύπτει σχέσεις μεταξύ των δεδομένων οι οποίες ήταν άγνωστες μέχρι εκείνη την στιγμή. Είναι μια ικανότητα η οποία συνδυάζει πολλούς κλάδους όπως είναι μηχανική μάθηση, η στατιστική, και η τεχνολογία βάσεων δεδομένων. Οι γνώσεις που προκύπτουν από την εξόρυξη δεδομένων μπορούν να χρησιμοποιηθούν για:

- Μάρκετινγκ
- Ανίχνευση απάτης
- Ανακαλύψεις στην επιστημονική κοινότητα

Η εξόρυξη γνώσης είναι επίσης γνωστή ως: Ανακάλυψη γνώσης, εξόρυξη γνώσης, ανάλυση δεδομένων/προτύπων, συλλογή πληροφοριών κ.α.

Η εξόρυξη δεδομένων μπορεί να εφαρμοστεί στα παρακάτω κατηγορίες δεδομένων:

- Σχεσιακές Βάσεις Δεδομένων
- Αποθήκες Δεδομένων
- Αντικειμενοστραφείς βάσεις δεδομένων
- Χωρικές Βάσεις δεδομένων
- Ετερογενείς και παλαιές βάσεις δεδομένων
- Πολυμεσικές Βάσεις δεδομένων
- Βάσεις Δεδομένων κειμένου

- Εξόρυξη κειμένου και εξόρυξη Ιστού

Στις περισσότερες κατηγορίες δεδομένων βλέπουμε πολύ συχνά την έννοια της βάσης δεδομένων. Πριν προχωρήσουμε στην ανάλυση της εξόρυξης δεδομένων είναι χρήσιμο να αναφερθούμε εν συντομία στην έννοια της Βάσης δεδομένων.(Data Mining Tutorial: Process, Techniques, Tools, EXAMPLES, n.d.)

2.1.2 Βάσεις Δεδομένων

Μια βάση δεδομένων είναι μια συστηματική συλλογή από δεδομένα. Οι βάσεις δεδομένων υποστηρίζουν αποθήκευση και χειρισμό δεδομένων. Οι βάσεις δεδομένων κάνουν πολύ εύκολη την διαδικασία της διαχείρισης δεδομένων. Παρακάτω αναφέρονται μερικά παραδείγματα:

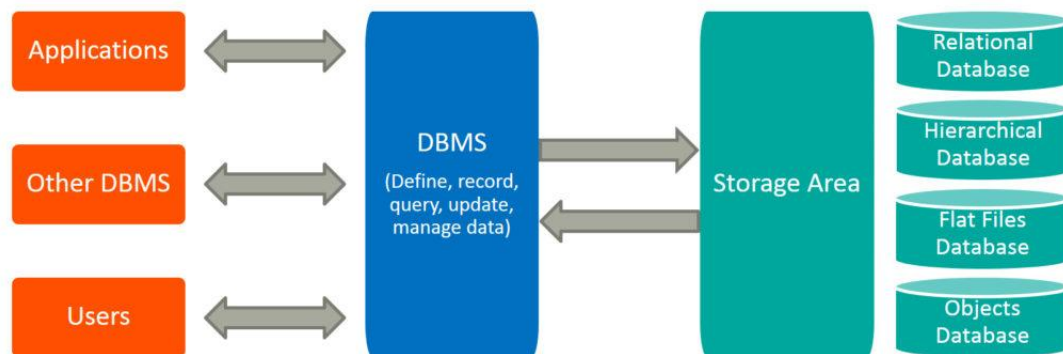
- Ένας online τηλεφωνικός κατάλογος θα χρησιμοποιούσε μια βάση δεδομένων για την αποθήκευση δεδομένων σχετικά με ανθρώπους, τηλεφωνικούς αριθμούς και άλλα στοιχεία επικοινωνίας.
- Ένας πάροχος ηλεκτρικής ενέργειας χρησιμοποιεί βάση δεδομένων για να διαχειρίζεται τους λογαριασμούς και άλλα στοιχεία σχετικά με τους πελάτες του.
- Ένα άλλο παράδειγμα όπου χρησιμοποιούνται οι βάσεις δεδομένων είναι το Facebook. Αυτό το μέσο κοινωνικής δικτύωσης χρειάζεται να αποθηκεύει, να επεξεργάζεται και να παρουσιάζει δεδομένα σχετικά με τα μέλη του, τους φίλους τους καθώς και λεπτομέρειες για την δραστηριότητα των μελών, μηνύματα και πολλά άλλα.

Αυτά είναι λίγα παραδείγματα όπου χρησιμοποιούνται οι βάσεις δεδομένων. Στην σημερινή κοινωνία υπάρχουν αμέτρητα τέτοια παραδείγματα καθώς οι περισσότεροι οργανισμοί, επιχειρήσεις ακόμα και κυβερνήσεις χρησιμοποιούν βάσεις δεδομένων.

Συστήματα Διαχείρισης Βάσεων Δεδομένων

Ένα **Σύστημα Διαχείρισης Βάσεων Δεδομένων** (ΣΔΒΔ – Database Management System) είναι μια συλλογή από προγράμματα τα οποία δίνουν την δυνατότητα στους χρήστες τους να έχουν πρόσβαση σε βάσεις δεδομένων, να επεξεργάζονται τα δεδομένα, και να αναπαριστούν τα δεδομένα στον χρήστη. Επιπλέον βοηθούν τον έλεγχο πρόσβασης στην βάση. Το πρώτο σύστημα βάσεων διαχείρισης βάσεων δεδομένων υλοποιήθηκε την δεκαετία του 1960.

Database Management System



Εικόνα 1: Σύστημα Διαχείρισης Βάσεων Δεδομένων

Είδη Συστημάτων Διαχείρισης Βάσεων Δεδομένων

Υπάρχουν 4 κύρια είδη συστημάτων διαχείρισης βάσεων δεδομένων:

1. **Ιεραρχικό:** Αυτό το είδος χρησιμοποιεί την σχέση «γονέα-παιδιού» για να αποθηκεύει δεδομένα. Αυτός ο τύπος ΣΔΒΔ πλέον χρησιμοποιείται σπάνια. Η δομή του είναι σαν δέντρο με τις κορυφές να αναπαριστούν εγγραφές και τις ακμές να αναπαριστούν πεδία. Το μητρώο των Windows που χρησιμοποιούσαν τα Windows XP είναι ένα παράδειγμα μιας ιεραρχικής βάσης δεδομένων. Οι ρυθμίσεις διαμόρφωσης αποθηκεύονταν σαν δομές δέντρων με κορυφές.
2. **Δικτυακό ΣΔΒΔ:** Αυτός ο τύπος ΣΔΒΔ υποστηρίζει σχέσεις πολλά προς πολλά. Αυτό συνήθως έχει ως αποτέλεσμα δομικά περίπλοκες βάσεις δεδομένων. Ο RDM Server είναι μια περίπτωση Συστήματος Διαχείρισης Βάσεων Δεδομένων το οποίο υλοποιεί το δικτυακό μοντέλο.
3. **Σχεσιακό ΣΔΒΔ:** Αυτός ο τύπος συστημάτων διαχείρισης βάσεων δεδομένων ορίζει σχέσεις βάσεων δεδομένων με την μορφή πινάκων, οι οποίες είναι και γνωστές ως σχέσεις. Σε αντίθεση με τα Δικτυακά Συστήματα Βάσεων Δεδομένων, τα σχεσιακά ΣΔΒΔ δεν υποστηρίζουν σχέσεις πολλά προς πολλά. Τα σχεσιακά ΣΔΒ, έχουν συνήθως προκαθορισμένους τύπους δεδομένων τους οποίους μπορούν να υποστηρίξουν. Αυτός ο τύπος Συστημάτων Διαχείρισης Βάσεων Δεδομένων είναι ο πιο δημοφιλής στην αγορά.

Παραδείγματα σχεσιακών ΣΔΒΔ είναι τα: MySQL, Oracle, Microsoft SQL Server Database κ.α.

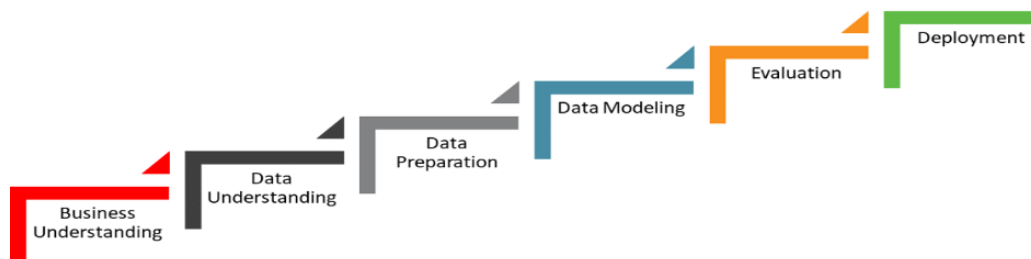
4. **ΣΔΒΔ Αντικειμενοστραφών Σχέσεων:** Αυτός ο τύπος ΣΔΒΔ υποστηρίζει την αποθήκευση νέων τύπων δεδομένων. Τα δεδομένα προς αποθήκευση έχουν την μορφή αντικειμένων. Τα αντικείμενα αυτά έχουν ιδιότητες και μεθόδους, οι οποίες ορίζουν τι πρέπει να κάνουν με τα δεδομένα. Ένα παράδειγμα Συστημάτων Διαχείρισης Βάσεων Δεδομένων αντικειμενοστραφών σχέσεων είναι το PostgreSQL.(What is Database? What is SQL?, n.d.)

2.1.3 Φάσεις Εξόρυξης Γνώσης

Αφού αναφερθήκαμε στις βάσεις δεδομένων, θα συνεχίσουμε την αναφορά στην εξόρυξη γνώσης και θα αναλύσουμε τις φάσεις που απαιτεί η διαδικασία Εξόρυξης γνώσεως:

Η διαδικασία εξόρυξης γνώσης χωρίζεται σε 6 φάσεις:

1. Κατανόηση Θέματος
2. Κατανόηση Δεδομένων
3. Προετοιμασία Δεδομένων
4. Μοντελοποίηση
5. Αξιολόγηση
6. Ανάπτυξη



Εικόνα 2: Φάσεις Εξόρυξης Δεδομένων

Ας δούμε αναλυτικά την κάθε φάση:

1. Κατανόηση Θέματος

Σε αυτή την φάση καθορίζονται οι στόχοι της διαδικασίας.

Αρχικά οι στόχοι του θέματος αλλά και των πελατών πρέπει να είναι κατανοητοί. Είναι αναγκαίο να έχει οριστεί επακριβώς το τι θέλει ο πελάτης (κάτι που τις περισσότερες φορές δεν ξέρει ούτε ο ίδιος). Στην συνέχεια πρέπει να γίνει ανασκόπηση του τρέχοντος σεναρίου εξόρυξης δεδομένων. Σε αυτή την φάση πρέπει να ληφθούν υπόψη στην εκτίμησή μας, οι πόροι, παραδοχές, περιορισμοί και άλλοι σημαντικοί παράγοντες. Η χρήση των στόχων του θέματος, αλλά και το τρέχων σενάριο, ορίζουν επακριβώς τους στόχους της εξόρυξης γνώσης. Ένα καλό σχέδιο εξόρυξης γνώσης πρέπει να είναι ιδιαίτερα λεπτομερές.

2. Κατανόηση Δεδομένων

Σε αυτή την φάση, γίνεται ο έλεγχος κατά πόσο τα δεδομένα είναι κατάλληλα για τους στόχους της εξόρυξης δεδομένων:

Αρχικά τα δεδομένα συλλέγονται από πηγές δεδομένων που είναι διαθέσιμες. Αυτές οι πηγές μπορεί να περιέχουν πολλαπλές βάσεις δεδομένων κ.α. Είναι μια πολύπλοκη διαδικασία, καθώς τα δεδομένα από πολλές πηγές, τις περισσότερες φορές, δεν ταιριάζουν μεταξύ τους. Για παράδειγμα, ο πίνακας A περιέχει μια οντότητα που ονομάζεται cust_no ενώ κάποιος άλλος πίνακας B περιέχει μια οντότητα που ονομάζεται cust_id. Αυτό έχει σαν αποτέλεσμα, να είναι εξαιρετικά δύσκολο να διασφαλιστεί ότι και τα δύο δοθέντα αντικείμενα αναφέρονται στην ίδια τιμή. Σε αυτή την περίπτωση, πρέπει να χρησιμοποιηθούν μεταδεδομένα για την μείωση των λαθών στην διαδικασία ενσωμάτωσης των δεδομένων. Επόμενο βήμα είναι η αναζήτηση ιδιοτήτων για τα δεδομένα που έχουν συλλεχθεί. Ένας καλός τρόπος για την εξερεύνηση των δεδομένων είναι η απάντηση στις ερωτήσεις εξόρυξης δεδομένων οι οποίες γράφτηκαν στην προηγούμενη φάση, χρησιμοποιώντας εργαλεία ερωτημάτων, αναφορών και οπτικοποίησης. Τέλος, βασιζόμενοι στα αποτελέσματα των ερωτημάτων, διαπιστώνεται η ποιότητα των δεδομένων.

3. Προετοιμασία Δεδομένων

Σε αυτή την φάση τα δεδομένα ετοιμάζονται για παραγωγή. Αυτή η φάση καταναλώνει περίπου το 90% του χρόνου της εργασίας. Τα δεδομένα από διαφορετικές πηγές επιλέγονται, «καθαρίζονται», μετασχηματίζονται, μορφοποιούνται και κατασκευάζονται (εάν είναι απαραίτητο). Ο «καθαρισμός» των δεδομένων είναι μια διαδικασία όπου «καθαρίζονται» τα δεδομένα με την εξομάλυνση κατεστραμμένων δεδομένων, και την συμπλήρωση ελλειπών τιμών. Για παράδειγμα, για προφίλ δημογραφικών στοιχείων πελατών, λείπουν δεδομένα σχετικά με την ηλικία. Τα δεδομένα είναι ελλιπή και πρέπει να συμπληρωθούν. Σε μερικές περιπτώσεις, θα μπορούσαν να υπάρχουν δεδομένα τα οποία είναι υπερβολικά υψηλά. Για παράδειγμα, η ηλικία έχει τιμή 300. Επιπλέον, τα δεδομένα μπορεί να είναι ασυνεπή. Για παράδειγμα, το όνομα ενός πελάτη μπορεί να είναι διαφορετικό σε διαφορετικούς πίνακες. Η φάση της προετοιμασίας δεδομένων αλλάζει τα δεδομένα και τα ετοιμάζει για την εξόρυξη δεδομένων. Σε αυτή την φάση γίνεται και η διαδικασία του μετασχηματισμού των δεδομένων. Η διαδικασία αυτή βοηθάει στην επιτυχή διαδικασία εξόρυξης δεδομένων:

Εξομάλυνση: Βοηθάει στην απομάκρυνση του θορύβου από τα δεδομένα

Συσσωμάτωση: Οι λειτουργίες συσσωμάτωσης εφαρμόζονται στα δεδομένα.

Παραδείγματος χάρη, τα δεδομένα των εβδομαδιαίων πωλήσεων συσσωματώνονται για να υπολογιστεί το μηνιαίο και το ετήσιο άθροισμα.

Γενίκευση: Σε αυτό το βήμα, τα δεδομένα χαμηλού επιπέδου αντικαθίστανται από έννοιες υψηλού επιπέδου. Σε αυτό βοηθάει και η ιεραρχίες εννοιών. Για παράδειγμα, η πόλη αντικαθίσταται από την χώρα.

Κανονικοποίηση: Η κανονικοποίηση εκτελείται όταν τα δεδομένα κλιμακώνονται πάνω ή κάτω. Παράδειγμα: Τα δεδομένα πρέπει να έχουν τιμές ανάμεσα στο 2.0 και στο -2.0 μετά την κανονικοποίηση.

Κατασκευή χαρακτηριστικών: Τα χαρακτηριστικά κατασκευάζονται και συμπεριλαμβάνονται στα δεδομένα.

Το αποτέλεσμα της διαδικασίας είναι ένα τελικό σύνολο δεδομένων, το οποίο μπορεί να χρησιμοποιηθεί στην Μοντελοποίηση.

4. Μοντελοποίηση

Σε αυτή την φάση, χρησιμοποιούνται μαθηματικά μοντέλα για να καθοριστούν τα πρότυπα δεδομένων. Βασιζόμενοι στους στόχους που έχουν καθοριστεί, επιλέγονται κατάλληλες τεχνικές μοντελοποίησης για το διαθέσιμο σύνολο δεδομένων. Στην συνέχεια γίνεται η δημιουργία σεναρίου για τον έλεγχο της ποιότητας και της εγκυρότητας των δεδομένων. Έπειτα, το μοντέλο τρέχει στο σύνολο δεδομένων. Τέλος, τα αποτελέσματα πρέπει να αξιολογηθούν από όλα τα ενδιαφερόμενα μέρη για να γίνει βέβαιο ότι το μοντέλο πληροί τους ελάχιστους στόχους.

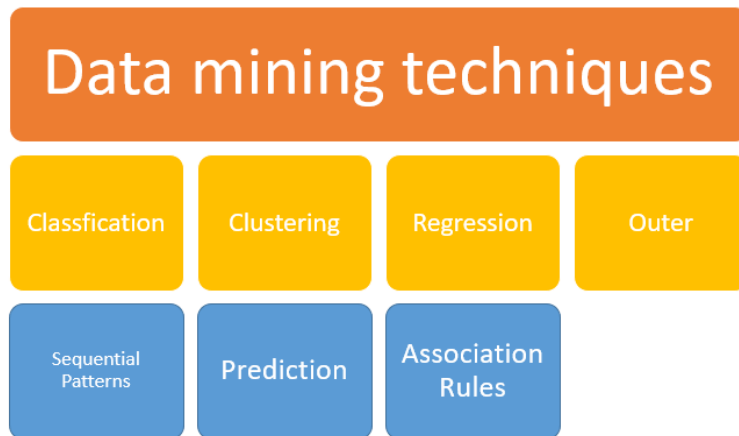
5. Αξιολόγηση

Σε αυτή την φάση, προσδιορίζονται τα πρότυπα σε σχέση με τους στόχους. Τα αποτελέσματα που παράγονται από το μοντέλο εξόρυξης δεδομένων αξιολογούνται σε σχέση με τους επιχειρησιακούς στόχους. Η απόκτηση επιχειρησιακής κατανόησης είναι μια επαναληπτική διαδικασία. Στην πραγματικότητα, κατά την διάρκεια της κατανόησης μπορεί να προκύψουν νέες επιχειρηματικές απαιτήσεις εξαιτίας της εξόρυξης δεδομένων. Σε αυτή την φάση λαμβάνεται η απόφαση εάν θα προχωρήσει το μοντέλο στην φάση της ανάπτυξης.

6. Ανάπτυξη

Στην τελευταία φάση της εξόρυξης δεδομένων, την φάση της ανάπτυξης, γίνεται η εισχώρηση των ανακαλύψεων της διαδικασίας εξόρυξης δεδομένων σε καθημερινές επιχειρηματικές δραστηριότητες. Η γνώση της πληροφορίας που ανακαλύπτεται κατά την διάρκεια της διαδικασίας πρέπει να είναι ευκολά κατανοήσιμη από τους ενδιαφερόμενους, οι οποίοι δεν έχουν τεχνικό υπόβαθρο. Δημιουργείται, ένα λεπτομερές σχέδιο ανάπτυξης για την μεταφορά, την συντήρηση και την παρακολούθηση των ανακαλύψεων της εξόρυξης γνώσης. Τέλος, δημιουργείται μια τελική έκθεση έργου με βασικές εμπειρίες κατά την διάρκεια του έργου. Αυτό βοηθά στη βελτίωση της επιχειρηματικής πολιτικής του οργανισμού.

2.1.3 Τεχνικές Εξόρυξης δεδομένων



Εικόνα 3: Τεχνικές Εξόρυξης γνώσης

1. Ταξινόμηση:

Αυτή η μέθοδος χρησιμοποιείται για την ανάκτηση σημαντικών και σχετικών πληροφοριών σχετικά με τα δεδομένα και τα μεταδεδομένα. Αυτή η μέθοδος εξόρυξης δεδομένων βοηθάει στην ταξινόμηση των δεδομένων σε διαφορετικές κλάσεις.

2. Ομαδοποίηση:

Είναι μια τεχνική εξόρυξης δεδομένων για τον εντοπισμό όμοιων δεδομένων. Αυτή η διαδικασία βοηθάει στην κατανόηση των διαφορών και των ομοιοτήτων των δεδομένων.

3. Παλινδρόμηση:

Είναι μια τεχνική εξόρυξης δεδομένων για τον εντοπισμό και την ανάλυση των σχέσεων μεταξύ των μεταβλητών. Χρησιμοποιείται για τον προσδιορισμό της πιθανότητας μιας συγκεκριμένης μεταβλητής σε σχέση με άλλες μεταβλητές.

4. Κανόνες Σύνδεσης:

Αυτή η τεχνική εξόρυξης δεδομένων βοηθάει στον εντοπισμό των συσχετίσεων μεταξύ δύο ή περισσότερων αντικειμένων. Ανακαλύπτει κρυμμένα μοτίβα στο σύνολο δεδομένων.

5. Εξωτερική ανίχνευση:

Αυτός ο τύπος τεχνική εξόρυξης δεδομένων αναφέρεται στην παρατήρηση των αντικειμένων δεδομένων στο σύνολο δεδομένων, το οποίο δεν ταιριάζει σε ένα αναμενόμενο μοτίβο ή σε κάποια αναμενόμενη συμπεριφορά.

6. Διαδοχικά Μοτίβα:

Αυτή η τεχνική εξόρυξης δεδομένων βοηθάει στον εντοπισμό όμοιων προτύπων σε δεδομένα συναλλαγών για μια συγκεκριμένη περίοδο.

7. Πρόβλεψη

Η πρόβλεψη χρησιμοποιεί έναν συνδυασμό από άλλες τεχνικές εξόρυξης πληροφορίας. Αναλύει παρελθόντα γεγονότα σε μια κατάλληλη σειρά για την πρόβλεψη μελλοντικών γεγονότων.(Data Mining Tutorial: Process, Techniques, Tools, EXAMPLES, n.d.)

2.2 Εξόρυξη Μουσικής Πληροφορίας

Στην εποχή του διαδικτύου υπάρχει μεγάλος όγκος δεδομένων που σχετίζονται με την μουσική, ο οποίος είναι εύκολα προσβάσιμος από οποιονδήποτε. Όσο η ποσότητα των διαθέσιμων δεδομένων σχετικά με την μουσική αυξάνεται, γίνεται όλο και δυσκολότερη η διαδικασία οργάνωσης και ανάλυσης αυτών των δεδομένων αυτών. Σε αυτό έρχεται να βοηθήσει η εξόρυξη γνώσης. Σε αυτή την ενότητα θα αναλύσουμε την διαδικασία της εξόρυξης μουσικής πληροφορίας.

2.2.1 Πηγές Μουσικής Πληροφορίας και ορισμός Εξόρυξης Μουσικής Πληροφορίας

Στον παρακάτω πίνακα αναφέρονται οι πηγές μουσικής πληροφορίας καθώς και παραδείγματα ιστοσελίδων όπου μπορεί κάποιος να βρει την αντίστοιχη μουσική πληροφορία:

Πηγές Δεδομένων	Σύντομη Περιγραφή
Μουσικά Μεταδεδομ ένα	Περιέχουν πολλές πληροφορίες, οι οποίες περιγράφουν συγκεκριμένα μουσικά κομμάτια. Γενικά, πολλά μουσικά αρχεία υποστηρίζουν μια δομή αποθήκευσης γνωστή ως ID3, στην οποία αποθηκεύονται μεταδεδομένα όπως το όνομα του καλλιτέχνη, ο τίτλος, περιγραφή μουσικής και το άλμπουμ
Ακουστικά	Τα ακουστικά χαρακτηριστικά της μουσικής περιλαμβάνει κάθε ακουστική

χαρακτηριστικά	ιδιότητα ενός ήχου που μπορεί να εγγραφεί και να αναλυθεί. Παραδείγματα είναι: η ένταση, το τέμπο, ο ρυθμός κ.α.
Στίχοι	Είναι ένα σύνολο από λέξεις οι οποίες φτιάχνουν ένα τραγούδι. Γενικά, η έννοια του περιεχομένου που διέπει τους στίχους μπορεί να είναι άμεση ή έμμεση. Οι περισσότεροι στίχοι έχουν συγκεκριμένες έννοιες, περιγράφοντας το συναίσθημα του καλλιτέχνη, θρησκευτικές πεποιθήσεις, ή αντιπροσωπεύει θεμάτων της εποχής, όμορφα φυσικά τοπία κ.α.
Μουσικές Κριτικές	Οι μουσικές κριτικές βοηθούν στην εξέταση των τρόπων με τους οποίους οι θαυμαστές της μουσικής περιγράφουν τις μουσικές επιπτώσεις τους και τις πιθανές επιπτώσεις αυτών των προτιμήσεων
Κοινωνικές Ετικέτες	Είναι μια συλλογή από πληροφορίες κειμένου, οι οποίες σχολιάζουν διάφορα μουσικά στοιχεία όπως άλμπουμ, τραγούδια καλλιτέχνες κ.α. Οι κοινωνικές ετικέτες δημιουργούνται από δημόσιες ετικέτες των θαυμαστών της μουσικής. Σε αυτά αποτυπώνονται πολλές πληροφορίες, συμπεριλαμβανομένων και του είδους της μουσικής, το συναίσθημα, τα όργανα και η ποιότητα
Προφίλ Χρηστών και Λίστα αναπαραγωγής	Το προφίλ των χρηστών αντιπροσωπεύει τις προτιμήσεις των χρηστών σε μουσικές πληροφορίες όπως το τι είδους τραγούδια ενδιαφέρουν κάποιον και ποιος καλλιτέχνης αρέσει σε κάποιον. Η λίστα αναπαραγωγής αναφέρεται στην λίστα μουσικών κομματιών που προτιμάει ή έχει ακούσει κάποιος
Αρχεία MIDI	Το MIDI είναι συντομογραφία του Music Instrument Digital Interface (Ψηφιακή Διεπαφή Μουσικών οργάνων) είναι ένα κριτήριο που υιοθετείται από την ηλεκτρονική μουσική βιομηχανία για τον έλεγχο των συσκευών, όπως συνθεσάιζερ και κάρτες ήχου, οι οποίες εκπέμπουν μουσική. Κατ' ελάχιστο μια αναπαράσταση του ήχου σε MIDI περιλαμβάνει τιμές για την ένταση την διάρκεια κ.α.
Μουσικές Βαθμολογίες	Αναφέρεται σε χειρόγραφή ή εκτυπωμένη μουσική σημειογραφία, η οποία χρησιμοποιεί πεντάγραμμα για να αναπαραστήσει ένα μουσικό κομμάτι. Οι μουσικές βαθμολογίες χρησιμοποιούνται για την αναπαραγωγή μουσικών

κομματιών. Για παράδειγμα ένας πιανίστας παίζει μια διάσημη μουσική

Πίνακας 1: Τύποι Μουσικών Δεδομένων

Οι παραπάνω διαφορετικοί τύποι πηγών μουσικών δεδομένων αντιπροσωπεύουν διαφορετικά χαρακτηριστικά δεδομένων μουσικής. Η εξόρυξη μουσικών δεδομένων αποσκοπεί στην ανακάλυψη χρήσιμων πληροφοριών και εγγενών χαρακτηριστικών αυτών των πηγών.

2.2.2 Επισκόπηση Εξόρυξης Μουσικής πληροφορίας

Η εξόρυξη δεδομένων αφορά δύο σημαντικά ζητήματα: σε τι είδους δεδομένα και τι είδους εργασίες. Παρομοίως, η εξόρυξη μουσικών δεδομένων εξάγει ενδιαφέροντα πρότυπα από διάφορους τύπους πηγών μουσικών δεδομένων.

Είδος Δεδομένων

Στην περίπτωση της συλλογής μουσικών δεδομένων, οι πηγές μπορεί να είναι πολλά πράγματα όπως αναφέραμε στον παραπάνω πίνακα. Διαφορετικές τεχνικές ανάλυσης και πειράματα διεξάγονται με βάση εργασίες εξόρυξης μουσικών δεδομένων.

Είδος εργασιών

Η εξόρυξη μουσικών δεδομένων συμπεριλαμβάνει μεθόδους για πολλές εργασίες. Διαφορετικές εργασίες εξόρυξης μουσικών δεδομένων εστιάζουν σε διαφορετικές πηγές δεδομένων και προσπαθούν να εξερευνήσουν διαφορετικές πτυχές των πηγών. Για παράδειγμα, η ταξινόμηση ειδών μουσικής αποσκοπεί στην αυτόματη ταξινόμηση των μουσικών σημάτων σε μια μοναδική κλάση εκμεταλλευόμενη την υπολογιστική ανάλυση των αναπαραστάσεων μουσικών χαρακτηριστικών.

2.2.3 Διαχείριση Μουσικών Δεδομένων

Μετά την εισαγωγή των πολυμέσων στους προσωπικούς υπολογιστές, γίνεται όλο και πιο σύνηθες η ψηφιοποίηση μέρους των δεδομένων πολυμέσων. Ένα σημαντικό πλεονέκτημα των ψηφιοποιημένων δεδομένων έναντι των αναλογικών, είναι ότι τα ψηφιοποιημένα δεδομένα μπορούν να χρησιμοποιηθούν από κοινού με άλλους. Μια καλή λύση για την δημιουργία μεγάλων συλλογών από μουσικά δεδομένα είναι η οργάνωσή τους σε Πολυμεσικές ψηφιακές βιβλιοθήκες. Το πρόβλημα μιας ψηφιακής βιβλιοθήκης είναι πως μπορεί να αποθηκευτεί και να οργανωθούν οι εγγραφές δεδομένων μουσικής, ώστε οι μουσικοί θαυμαστές να μπορούν να βρουν γρήγορα τους μουσικούς πόρους που τους ενδιαφέρουν.

Δημιουργία Ευρετηρίου Μουσικής

Μια πρόκληση στην διαχείριση μουσικών δεδομένων είναι ο τρόπος χρήσης τεχνικών ευρετηριασμού δεδομένων που βασίζονται σε διαφορετικές πτυχές των ίδιων των δεδομένων. Για να τα μουσικά δεδομένα, μπορούν να εφαρμοστούν διάφορα ακουστικά χαρακτηριστικά για την αποτελεσματική διαχείριση μουσικής. Για παράδειγμα η Shenetal. παρουσίασε μια νέα προσέγγιση για την δημιουργία μικρών αλλά περιεκτικών μουσικών περιγραφών για την παροχή υπηρεσιών αποδοτικής πρόσβασης και ανάκτησης των μουσικών δεδομένων. Σε αντίθεση με προσεγγίσεις που βασίζονται σε φασματικά χαρακτηριστικά χαμηλού επιπέδου προσαρμοσμένα από την τεχνολογία ανάλυσης ομιλίας, η προσέγγιση αυτή ενσωματώνει την αντίληψη της ανθρώπινης μουσικής για την βελτίωση της ακρίβειας της διαδικασίας ανάκτησης και ταξινόμησης μέσω PCA και νευρωνικών δικτύων.

2.2.4 Οπτικοποίηση Μουσικής

Σήμερα, οι άνθρωποι εκτός από τα ακουστικά ερεθίσματα λαμβάνουν και οπτικά ερεθίσματα. Αυτό έχει σαν αποτέλεσμα, όλο και περισσότερα λογισμικά αναπαραγωγής πολυμέσων να παρέχουν χαρακτηριστικά που δημιουργούν κινούμενες εικόνες με βάση ένα κομμάτι ηχογραφημένης μουσικής. Οι εικόνες συνήθως δημιουργούνται και αποδίδονται σε πραγματικό χρόνο και συγχρονίζονται με την μουσική κατά την διάρκεια της αναπαραγωγής. Στην πραγματικότητα, η οπτικοποίηση των μουσικών δεδομένων δεν περιορίζεται μόνο στην απεικόνιση ενός μόνο μουσικού εγγράφου ως κινούμενες εικόνες. Οι προσεγγίσεις για την απεικόνιση μουσικών συλλογών μπορούν να χωριστούν σε δύο κατηγορίες:

Σε εκείνους που επικεντρώνονται στην οπτικοποίηση των μεταδεδομένων ή του ακουστικού περιεχομένου μεμονωμένων μουσικών εγγράφων.

2.2.5 Εξόρυξη Συσχετίσεων

Η εξόρυξη συσχετίσεων, αναφέρεται στον εντοπισμό συσχετίσεων μεταξύ διαφορετικών στοιχείων σε ένα σύνολο δεδομένων. Ειδικά στην εξόρυξη μουσικών δεδομένων, η εξόρυξη συσχετίσεων, μπορεί να διαιρεθεί σε τρεις διαφορετικές κατηγορίες:

1. Εντοπισμός συσχετίσεων μεταξύ διαφορετικών ακουστικών χαρακτηριστικών
2. Εντοπισμός συσχετίσεων μεταξύ των μουσικών και άλλων μορφών εγγράφων.

3. Εντοπισμός συσχετίσεων μεταξύ μουσικών χαρακτηριστικών και άλλων μουσικών χαρακτηριστικών π.χ. συναίσθημα.

2.2.6 Εξόρυξη ακολουθίας

Για τα μουσικά δεδομένα, η εξόρυξη ακολουθίας στοχεύει κυρίως στην ανίχνευση δευτερευόντων ακολουθιών από διαφορετικές πτυχές ήχου, για παράδειγμα, ακολουθίες χορδών. Υπάρχουν σχετικά λίγες δημοσιεύσεις που σχετίζονται με τις εργασίες εξόρυξης μουσικών ακολουθιών. Η κύρια συμβολή της εξόρυξης ακολουθιών είναι στον τομέα της μουσικής μεταγραφής. Κατά την διάρκεια της μεταγραφής κομματιών ήχου, ενδέχεται να εισαχθούν διαφορετικοί τύποι σφαλμάτων όπως, σφάλματα τμηματοποίησης, σφάλματα υποκατάστασης κ.α. Για τον καλύτερο έλεγχο του ποσοστού σφάλματος της μετεγγραφής, οι ερευνητές προσπαθούν να διερευνήσουν την σκοπιμότητα της εφαρμογής της εξόρυξης ακολουθιών στην διαδικασία της μεταμαγνητοφώνησης μουσικών ηχογραφήσεων.

2.2.7 Ταξινόμηση

Η ταξινόμηση είναι ένα σημαντικό ζήτημα στις εργασίες εξόρυξης δεδομένων μουσικής. Κατά τις τελευταίες δεκαετίες προέκυψαν διάφορα προβλήματα ταξινόμησης. Ορισμένοι ερευνητές εστιάζουν στην κατάταξη της μουσικής από ηχητικά κομμάτια, ενώ άλλοι ασχολούνται με την κατηγοριοποίηση της μουσικής σε διαφορετικές ομάδες. Η πιο γενική ταξινόμηση επικεντρώνεται στο είδος της μουσικής. Επιπλέον, υπάρχουν και άλλες εργασίες ταξινόμησης όπως η ταξινόμηση καλλιτέχνη/τραγουδιστή, η ταξινόμηση διάθεσης η ταξινόμηση οργάνων κ.α.

- Επιγραμματικά οι κατηγορίες ταξινόμησης είναι:
- Ταξινόμηση Ήχου
- Ταξινόμηση Είδους
- Εντοπισμός Διάθεσης και Ταξινόμηση
- Αναγνώριση Οργάνων και Ταξινόμηση
- Ταξινόμηση Καλλιτέχνη
- Αναγνώριση Τραγουδιστή

2.2.8 Ομαδοποίηση

Η ομαδοποίηση είναι η διαδικασία του διαχωρισμού μιας συλλογής δεδομένων σε διαφορετικές ομάδες με βάση ορισμένα κριτήρια. Ειδικά στην περίπτωση της εξόρυξης μουσικών δεδομένων, η ομαδοποίηση στοχεύει στην διαίρεση μιας συλλογής μουσικών δεδομένων σε ομάδες παρόμοιων αντικειμένων με βάση τις ομοιότητες των ζευγών τους χωρίς προκαθορισμένες ετικέτες.

2.2.9 Σύνοψη Μουσικής

Ο τρόπος δημιουργίας μιας συνοπτικής και δημιουργικής εξαγωγής που συνοψίζει καλύτερα ένα πρότυπο ψηφιακό περιεχόμενο είναι άλλη μια πρόκληση στην μουσική εξόρυξη δεδομένων και είναι εξαιρετικά σημαντικό σε μεγάλης κλίμακας οργάνωση πληροφοριών. Πρόσφατα, τα περισσότερα από τις μουσικές συνόψεις για εμπορική χρήση δημιουργούνται χειροκίνητα από τις αρχικές ηχογραφήσεις μουσικής. Ωστόσο, δεδομένου ότι ένας μεγάλος όγκος ψηφιακού περιεχομένου έχει δημοσιοποιηθεί σε μέσα όπως για παράδειγμα το διαδίκτυο, οι αποτελεσματικές προσεγγίσεις της παραγωγής αυτόματης σύνοψης μουσικής έχουν γίνει όλο και πιο σημαντικές και απαραίτητες. Παρόμοια με την σύνοψη κειμένου, η σύνοψη μουσικής στοχεύει στο να προσδιορίσει τα πιο γενικά και κυριότερα θέματα ενός συγκεκριμένου μουσικού κομματιού που μπορεί να χρησιμοποιηθεί σαν αντιπροσωπευτικό δείγμα της μουσικής και να αναγνωριστεί με ευκολία από έναν ακροατή. Η αυτόματη σύνοψη μουσικής μπορεί να εφαρμοστεί σε ευρετηριασμό μουσικής και διανομή μουσικής με βάση το διαδίκτυο.

2.3 Ανάκτηση Μουσικής Πληροφορίας

2.3.1 Γενικά

Η ανάκτηση μουσικής πληροφορίας (Music Information Retrieval - MIR) είναι ένας αναδυόμενος ερευνητικός τομέας αφιερωμένος στην κάλυψη των αναγκών των χρηστών σε μουσική πληροφορία. Όπως θα φανεί, παρά την έμφαση στον όρο ανάκτηση, το MIR περιλαμβάνει μια σειρά διαφορετικών προσεγγίσεων που στοχεύουν στην διαχείριση της μουσικής, την εύκολη πρόσβαση και την απόλαυση. Το μεγαλύτερο μέρος των ερευνητικών εργασιών σχετικά με το MIR, τις προτεινόμενες τεχνικές και τα ανεπτυγμένα συστήματα βασίζονται στο περιεχόμενο. Η κύρια ιδέα που διέπει τις προσεγγίσεις βάση του περιεχομένου είναι ότι ένα έγγραφο μπορεί να περιγραφεί από ένα σύνολο χαρακτηριστικών που υπολογίζονται άμεσα από το περιεχόμενό του. Ωστόσο, οι βασικές τεχνικές ανάκτησης

πληροφοριών, οι οποίες βασίζονται σε στατιστικά στοιχεία και στην θεωρία πιθανοτήτων, μπορεί να χρησιμοποιούνται γενικότερα. Μια μεγάλη ποικιλία διαφορετικών μεθόδων αναζήτησης περιεχομένου σε μουσικές βαθμολογίες και ηχητικά δεδομένα έχουν προταθεί και εφαρμοστεί σε ερευνητικά πρότυπα και εμπορικά συστήματα. Πέρα από τον περιορισμένο και σαφώς καθορισμένο καθήκον της αναγνώρισης των ηχογραφήσεων, για τις οποίες λειτουργούν καλά οι τεχνικές λήψης μουσικών δακτυλικών αποτυπωμάτων, είναι δύσκολο να προσδιοριστεί ποιες μέθοδοι πρέπει να εφαρμοστούν. Αυτό υπογραμμίζει την σημασία μιας σειράς συγκρίσεων τύπου TREC για αλγορίθμους για την αναζήτηση ηχογραφήσεων. Οι ηχητικές και συμβολικές μέθοδοι είναι χρήσιμες για διαφορετικές εργασίες. Για παράδειγμα, ο προσδιορισμός των εγγραφών πρέπει να βασίζεται σε ηχητικά δεδομένα, ενώ τα έργα καλύτερα να προσδιορίζονται βάσει συμβολικής αναπαράστασης. Για τον προσδιορισμό του είδους ενός δεδομένου μουσικού κομματιού, οι προσεγγίσεις που βασίζονται στον ήχο φαίνονται ελπιδοφόρες, αλλά οι συμβολικές μέθοδοι μπορούν επίσης να λειτουργήσουν.

2.3.2 Αναζήτηση Ομοιότητας Μουσικής

Στον τομέα εξόρυξης μουσικών δεδομένων, η αναζήτηση ομοιότητας αναφέρεται στην αναζήτηση αρχείων μουσικής παρόμοιου με ένα συγκεκριμένο αρχείο μουσικής. Αρχικά, η αναζήτηση μπορεί να πραγματοποιηθεί σε οποιαδήποτε διάσταση. Για παράδειγμα, ο χρήστης θα μπορούσε να δώσει ένα παράδειγμα του χρόνου ή του ήχου που αναζητά ή να περιγράψει τη συγκεκριμένη δομή ενός τραγουδιού, και στην συνέχεια το σύστημα αναζήτησης μουσικής θα αναζητήσει παρόμοια μουσικά έργα με βάση τις πληροφορίες που του έχει δώσει ο χρήστης. Οι εργασίες αναζήτησης ομοιότητας μπορούν να διαιρεθούν σε εξαγωγή χαρακτηριστικών και επεξεργασία ερωτημάτων. Μετά την εξαγωγή χαρακτηριστικών, τα μουσικά έργα μπορούν να αναπαρασταθούν με βάση τα χαρακτηριστικά που έχουν εξαχθεί. Στο στάδιο της επεξεργασίας ερωτημάτων, η κύρια εργασία είναι η εφαρμογή ενός κατάλληλου μέτρου ομοιότητας για τον υπολογισμό της ομοιότητας μεταξύ της δεδομένης μουσικής εργασίας και των έργων της υποψήφιας μουσικής. Στο πλαίσιο αυτό, εξετάστηκαν διάφορα υφιστάμενα μέτρα ομοιότητας και συναρτήσεως απόστασης, τα οποία εκτείνονται από απλές Ευκλείδειες αποστάσεις έως θεωρητικά μέτρα. Ανεξάρτητα από το τελικό μέτρο, μια σημαντική τάση στον χώρο της ανάκτησης μουσικής ήταν η χρήση ενός μοντέλου πυκνότητας των χαρακτηριστικών. Ο στόχος της αναζήτησης ομοιότητας είναι να βρεθούν μουσικά αρχεία ήχου παρόμοια με ένα αρχείο μουσικής που δίνεται ως είσοδος. Η μουσική ταξινόμηση με βάση το είδος και το στυλ είναι

φυσικά σε μορφή κάποιας ιεραρχίας. Η ομοιότητα μπορεί να χρησιμοποιηθεί για την ομαδοποίηση ήχων μαζί σε οποιοδήποτε κόμβο της ιεραρχίας. Η χρήση ηχητικών σημάτων για ομοιότητα δικαιολογείται από μια παρατήρηση ότι τα ηχητικά σήματα (ψηφιακά ή αναλογικά) της μουσικής που ανήκουν στο ίδιο είδος, έχουν κοινά χαρακτηριστικά, επειδή αποτελούνται από παρόμοιους τύπους οργάνων. (Li, 2010)

2.4 Μηχανική Μάθηση

2.4.1 Ορισμός

Σύμφωνα με τον Arthur Samuel, έναν πρωτοπόρο στον κλάδο της τεχνητής νοημοσύνης και των παιχνιδιών υπολογιστών, **Μηχανική Μάθηση** είναι ο κλάδος που δίνει στους υπολογιστές την δυνατότητα να «μάθουν» χωρίς να έχουν προγραμματιστεί ρητά.

Με απλά λόγια, με την Μηχανική Μάθηση μπορεί να γίνει αυτοματοποίηση και βελτίωση της ικανότητας των υπολογιστών να μαθαίνουν χωρίς να έχουν προγραμματιστεί στην πραγματικότητα. Αυτό σημαίνει ότι αφότου χρησιμοποιήσουν Μηχανικοί Μάθηση οι ηλεκτρονικοί υπολογιστές δεν θα κάνουν ενέργειες επειδή τους έχει προγραμματίσει κάποιος προγραμματιστής, αλλά επειδή έχουν μάθει. Η διαδικασία ξεκινάει με το να δίνεις σαν είσοδο καλά και ποιοτικά δεδομένα. Έτσι χρησιμοποιώντας αυτά τα δεδομένα σε διαφορετικούς αλγόριθμους δημιουργούνται τα μοντέλα μηχανικής μάθησης. Η επιλογή των αλγορίθμων εξαρτάται αποκλειστικά από το είδος των δεδομένων που έχουμε στην διάθεσή μας αλλά και από το είδος της εργασίας που θέλουμε να αυτοματοποιήσουμε.

Παράδειγμα: Εκπαίδευση των μαθητών σε ένα διαγώνισμα

Κατά την διάρκεια ενός διαγωνίσματος, οι μαθητές προσπαθούν να μάθουν τα πάντα για το μάθημα. Πριν το διαγώνισμα δίνουν στο μηχάνημα(εγκέφαλο) αρκετά και ποιοτικά δεδομένα (ερωτήσεις και απαντήσεις από διαφορετικά βιβλία ή σημειώσεις από καθηγητές). Στην πραγματικότητα εκπαιδεύουν τον εγκέφαλό τους και με είσοδο και με έξοδο. Για παράδειγμα, το είδος της λογικής που θα χρησιμοποιήσουν για να λύσουν ένα πρόβλημα είναι ένας τρόπος εκπαίδευσης εξόδου. Κάθε φορά λύνουν προβλήματα και βρίσκουν το πόσο καλά τα πήγανε (ευστοχία / βαθμολογία) με το να συγκρίνουν τις απαντήσεις που έδωσαν με τις απαντήσεις των προβλημάτων. Με αυτό τον τρόπο, σταδιακά η απόδοση των ανεβαίνει και είναι πιο σίγουρο ότι θα τα πάνε καλά στο διαγώνισμα.

Αυτός είναι ο τρόπος που κατασκευάζονται τα πραγματικά μοντέλα, εκπαιδεύεται η μηχανή με δεδομένα (είσοδο και έξοδο) και όταν γίνει αυτή η διαδικασία με αποτελεσματικό τρόπο, ακολουθεί δοκιμή στα δεδομένα (μόνο είσοδος) και έτσι παράγεται η βαθμολογία του μοντέλου συγκρίνοντας την απάντηση που έδωσε το μηχάνημα με την πραγματική έξοδο η οποία δεν έχει δοθεί κατά την διάρκεια της εκπαίδευσης.

2.4.2 Διαφορά Μηχανικής Μάθησης με τον Παραδοσιακό Προγραμματισμό



Εικόνα 4: Μηχανική Μάθηση και Παραδοσιακός Προγραμματισμός

Όπως φαίνεται και στην παραπάνω εικόνα η διαφορά της μηχανικής μάθησης με τον παραδοσιακό προγραμματισμό είναι πολύ απλή. Στον παραδοσιακό προγραμματισμό με την παροχή δεδομένων αλλά και προγράμματος τα τρέχουμε στον υπολογιστή και μας παράγει κάποια έξοδο. Από την άλλη, στην μηχανική μάθηση ο υπολογιστής τροφοδοτείται με δεδομένα(είσοδο) αλλά και έξοδο, γίνεται η εκτέλεση στο μηχάνημα κατά την διάρκεια της εκπαίδευσης, και με βάση την εκπαίδευση παράγεται ένα καινούριο πρόγραμμα με την δικιά του λογική(πρόγραμμα) το οποίο μπορεί να αξιολογηθεί κατά την διάρκεια του ελέγχου.

2.4.3 Λειτουργία μάθησης για έναν υπολογιστή

Παρακάτω θα δούμε πως λειτουργεί η μηχανική μάθηση:

1. Αρχικά, συλλέγονται παλιά δεδομένα σε μορφή κατάλληλη για επεξεργασία. Όσο πιο ποιοτικά είναι τα δεδομένα τόσο πιο κατάλληλα είναι για την μοντελοποίηση.

2. Επεξεργασία Δεδομένων: Μερικές φορές τα δεδομένα είναι σε μορφή που δεν έχουν δεχτεί επεξεργασία και υπάρχει πιθανότητα να λείπουν κάποιες τιμές. Οι τιμές που λείπουν σε αριθμητικά χαρακτηριστικά, όπως είναι η τιμή του σπιτιού μπορεί να αντικατασταθεί με την μέση τιμή του χαρακτηριστικού.
3. Στην συνέχεια γίνεται διαίρεση της πληροφορίας σε εκπαιδευτικά σύνολα, σύνολα διασταυρούμενη επικύρωσης και σύνολα δοκιμών. Η αναλογία μεταξύ των επιμέρους συνόλων πρέπει να είναι 6:2:2.
4. Έπειτα, γίνεται η δημιουργία μοντέλων χρησιμοποιώντας κατάλληλους αλγορίθμους και τεχνικές στα σύνολα δοκιμών.
5. Τέλος, το μοντέλο δοκιμάζεται με δεδομένα τα οποία δεν είχαν τροφοδοτηθεί στο μοντέλο κατά την διάρκεια της εκπαίδευσης και γίνεται η αξιολόγηση της απόδοσης. (ML | What is Machine Learning ?, n.d.)

2.4.4 Κατηγοριοποίηση της Μηχανικής Μάθησης

Οι υλοποιήσεις της μηχανικής μάθησης χωρίζονται σε 3 μεγάλες κατηγορίες:

Επιτηρούμενη Μάθηση:

Όταν ένας αλγόριθμος μαθαίνει από παραδείγματα δεδομένων και σχετικές απαντήσεις, οι οποίες μπορούν να αποτελούνται από αριθμητικές τιμές ή ετικέτες όπως είναι οι κλάσεις, για να μπορεί αργότερα να προβλεφθεί η σωστή απάντηση όταν τροφοδοτείται με νέα παραδείγματα, αυτό ανήκει στην κατηγορία της επιτηρούμενης μάθησης.

Μη Επιτηρούμενη Μάθηση:

Όταν ένας αλγόριθμος μαθαίνει από απλά παραδείγματα χωρίς σχετιζόμενες απαντήσεις, αφήνουμε τον αλγόριθμο να αποφασίσει μόνος του τα πρότυπα δεδομένων. Αυτός ο τύπος αλγορίθμου τείνει να αναδομεί τα δεδομένα σε κάτι άλλο, όπως για παράδειγμα νέα χαρακτηριστικά που μπορεί να αντιπροσωπεύει μια κλάση ή μια νέα σειρά μη συσχετιζόμενων τιμών. Είναι αρκετά χρήσιμο στο να δίνει στον άνθρωπο πληροφορίες σχετικά με την σημασία των δεδομένων σε αλγόριθμους επιτηρούμενης μάθησης.

Ενίσχυση της μάθησης:

Μερικές φορές τα δεδομένα που τροφοδοτούνται στον αλγόριθμο μηχανικής μάθησης δεν συνοδεύονται από ετικέτες. Ωστόσο υπάρχει δυνατότητα μαζί ένα παράδειγμα να ακολουθείται με θετική ή αρνητική ανατροφοδότηση με βάση την λύση που προτείνει ο αλγόριθμος. Αυτή η προσέγγιση καλείται Ενίσχυση της Μάθησης και συνδέεται με εφαρμογές για τις οποίες ο αλγόριθμος πρέπει να πάρει αποφάσεις και οι αποφάσεις αυτές έχουν κάποιες συνέπειες. Στον ανθρώπινο κόσμο αυτό είναι σαν να μαθαίνεις με δοκιμή και λάθος. Τα λάθη βοηθούν στο να μαθαίνεις επειδή έχουν μια ποινή (κόστος, απώλεια χρόνου κ.λπ.). Από αυτά τα λάθη μαθαίνουμε ότι μια πορεία δράσης είναι πιο πιθανό να πετύχει σε σχέση με άλλες. Ένα τέτοιο παράδειγμα ενισχυτικής μάθησης είναι όταν οι υπολογιστές μαθαίνουν να παίζουν παιχνίδια από μόνοι τους.

Σε αυτή την περίπτωση, η εφαρμογή παρουσιάζει στον αλγόριθμο παραδείγματα συγκεκριμένων καταστάσεων. Για παράδειγμα μια τέτοια κατάσταση μπορεί να είναι όταν ένας παίκτης βρίσκεται παγιδευμένος σε έναν λαβύρινθο ενώ προσπαθεί να αποφύγει κάποιον εχθρό. Η εφαρμογή ενημερώνει τον αλγόριθμο για την έκβαση των ενεργειών στις οποίες προβαίνει και η μάθηση γίνεται όσο προσπαθεί να αποφύγει τους κινδύνους.

Ημι-επιτηρούμενη Μάθηση

Σε αυτό το είδος μηχανικής μάθησης ο αλγόριθμος τροφοδοτείται με μη ολοκληρωμένα δεδομένα. Αυτή είναι μια ιδιαίτερη περίπτωση της αρχής της Μεταγωγής όπου ολόκληρο το σύνολο των στιγμιότυπων των προβλημάτων είναι γνωστό κατά την διάρκεια της εκμάθησης, με εξαίρεση ότι λείπει το κομμάτι των στόχων.

Ένα άλλο είδος κατηγοριοποίησης εργασιών μηχανικής μάθησης υπάρχει και έχει σχέση με την επιθυμητή έξοδο ενός συστήματος μηχανικής μάθησης:

Ταξινόμηση:

Όταν οι είσοδοι χωρίζονται σε δύο ή περισσότερες κατηγορίες και ο εκπαιδευόμενος πρέπει να δώσει ένα μοντέλο, το οποίο δίνει εισόδους που δεν έχουν προβλεφθεί σε μια ή περισσότερες κλάσεις. Αυτό συνήθως γίνεται με βάση την επιτηρούμενη μάθηση. Το φιλτράρισμα ανεπιθύμητων είναι ένα τέτοιο παράδειγμα όπου η είσοδος είναι ένα email και οι κλάσεις είναι «Ανεπιθύμητο» ή «Μη ανεπιθύμητο».

Παλινδρόμηση

Είναι και αυτή μια περίπτωση επιτηρούμενης μάθησης. Σε αυτή την περίπτωση η έξοδος είναι συνεχής και όχι διακριτή.

Συσταδοποίηση:

Είναι μια τεχνική όπου ένα σύνολο εισόδων διαιρείται σε ομάδες. Σε αντίθεση με την ταξινόμηση, οι ομάδες δεν είναι γνωστές εκ των προτέρων και αυτό έχει σαν αποτέλεσμα αυτές οι εργασίες είναι μη επιτηρούμενες. (An introduction to Machine Learning, n.d.)

2.4.5 Τα δεδομένα στην μηχανική μάθηση

Δεδομένα: Μπορεί να είναι οτιδήποτε δεν έχει υποστεί επεξεργασία όπως είναι τιμές, κείμενο, ήχο ή εικόνα που δεν έχει ερμηνευτεί και αναλυθεί. Τα δεδομένα είναι το πιο σημαντικό στην ανάλυση δεδομένων, την μηχανική μάθηση και την τεχνητή νοημοσύνη. Χωρίς τα δεδομένα, δεν μπορεί να γίνει εκπαίδευση των μοντέλων και όλη η έρευνα που έχει γίνει πάει στράφι. Μεγάλες εταιρίες ξοδεύουν πολλά χρήματα μόνο για να συγκεντρώνουν μεγάλες ποσότητες συγκεκριμένων δεδομένων. Παράδειγμα: Το Facebook απέκτησε το WhatsApp πληρώνοντας 19 εκατομμύρια. Αυτό έγινε διότι η εταιρεία ήθελε να έχει πρόσβαση σε δεδομένα χρηστών που μπορεί η Facebook να μην είχε πρόσβαση αλλά η WhatsApp να είχε. Αυτά τα δεδομένα είναι σημαντικά για την Facebook διότι με την συγκέντρωσή τους η εταιρεία έχει την δυνατότητα να βελτιώσει τις υπηρεσίες της.

Πληροφορίες: Οι πληροφορίες που έχουν ερμηνευτεί και διαχειριστεί και έχουν κάποια σημασία για τους χρήστες.

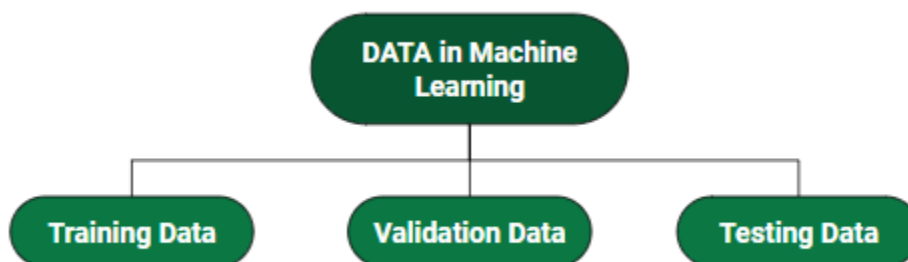
Γνώση: Συνδυασμός πληροφοριών, εμπειριών και μάθησης. Έχει σαν αποτέλεσμα την ευαισθητοποίηση ή την δημιουργία ιδεών για ένα άτομο ή οργανισμό.



Εικόνα 5: Δεδομένα Πληροφορία, Γνώση

Δεδομένα και Μηχανική Μάθηση

- Δεδομένα Εκπαίδευσης: Τα δεδομένα που χρησιμοποιούμε για να εκπαιδεύσουμε το μοντέλο μας. Αυτά είναι τα δεδομένα όπου βλέπει το μοντέλο μας.
- Δεδομένα επικύρωσης: Τα δεδομένα που χρησιμοποιούνται για να γίνεται μια συχνή αξιολόγηση του μοντέλου.
- Δεδομένα δοκιμής: Όταν το μοντέλο που έχει δημιουργηθεί εκπαιδευτεί πλήρως, τα δεδομένα δοκιμής παρέχουν αμερόληπτη αξιολόγηση. Όταν τροφοδοτούνται οι είσοδοι με δεδομένα δοκιμής, το διαθέσιμο μοντέλο προβλέπει κάποιες τιμές (χωρίς να βλέπει την πραγματική έξοδο). Μετά την πρόβλεψη, γίνεται αξιολόγηση του μοντέλου με το να συγκρίνονται με την πραγματική έξοδο. Με αυτόν τον τρόπο αξιολογούμε το μοντέλο και βλέπουμε πόσο έχει μάθει από τις εμπειρίες τις οποίες έχει λάβει σαν εκπαιδευτικά δεδομένα κατά την διάρκεια της εκπαίδευσης.



Εικόνα 6: Δεδομένα στην μηχανική μάθηση

Παράδειγμα:

Υπάρχει ο ιδιοκτήτης ενός καταστήματος, ο οποίος έκανε μια έρευνα για την οποία έχει μια μεγάλη λίστα από ερωτήσεις και απαντήσεις τις οποίες ρωτήθηκε από τους πελάτες. Η λίστα με τις ερωτήσεις και απαντήσεις είναι τα **δεδομένα**. Κάθε φορά που θέλει να βγάλει ένα συμπέρασμα, δεν είναι δυνατόν να ανατρέχει κάθε μια ερώτηση από τις χιλιάδες που υπάρχουν διαθέσιμες για να βρει κάτι σχετικό, διότι αυτό αποτελεί μια χρονοβόρα διαδικασία και καθόλου χρήσιμη. Για να λύσει αυτό το πρόβλημα και να κάνει αυτή την δουλειά πιο εύκολη, τα δεδομένα μετατρέπονται μέσω λογισμικού, υπολογισμών, γραφημάτων κ.α. Αυτή η παρεμβολή από υποβληθείσα δεδομένα είναι η **πληροφορία**. Έτσι τα δεδομένα είναι απαραίτητα για την

πληροφορία. Η γνώση παίζει ρόλο στην διαφοροποίηση δύο ατόμων που έχουν τις ίδιες πληροφορίες. Στην πραγματικότητα δεν αποτελεί τεχνικό όρο, αλλά συνδέεται με την διαδικασία της ανθρώπινης σκέψης.

Ιδιότητες Δεδομένων

- Όγκος: Κλίμακα δεδομένων. Με τον παγκόσμιο πληθυσμό να αυξάνεται και την τεχνολογική έκθεση, τεράστια δεδομένα δημιουργούνται κάθε δευτερόλεπτο.
- Ποικιλία: Διαφορετικές μορφές δεδομένων – υγεία, εικόνες, βίντεο και ήχος.
- Ταχύτητα: Ρυθμός ροής και δημιουργίας δεδομένων.
- Τιμή: Σημασία των δεδομένων όσον αφορά τις πληροφορίες που μπορούν να συνάγουν οι ερευνητές από αυτά.
- Ειλικρίνεια: Ασφάλεια και ορθότητα των δεδομένων στα οποία εργαζόμαστε. (ML | Introduction to Data in Machine Learning, n.d.)

Κεφάλαιο 3: Αλγόριθμοι

3.1 Κατηγοριοποίηση

3.1.2 Ορισμός

Η κατηγοριοποίηση είναι όλες εκείνες οι διαδικασίες εύρεσης ενός μοντέλου, το οποίο περιγράφει και διακρίνει τις κλάσεις και τις έννοιες των δεδομένων. Πιο συγκεκριμένα η κατηγοριοποίηση είναι το πρόβλημα του προσδιορισμού σε ποιο σύνολο κατηγοριών (ΥΠΟ πληθυσμών) ανήκει μια νέα παρατήρηση, βάσει ενός συνόλου από “Training data”, τα οποία περιέχουν παρατηρήσεις και των οποίων η σύνθεση είναι γνωστή.

Παράδειγμα: Πριν από την έναρξη οποιουδήποτε έργου, πρέπει να ελεγχθεί εάν η υλοποίηση του έργου είναι εφικτή. Σε αυτή την περίπτωση, το άτομο που αναλαμβάνει την κατηγοριοποίηση, απαιτείται να προβλέπει ετικέτες κλάσης όπως «ασφαλή» και «επικίνδυνο» με σκοπό την έγκριση του έργου.

3.1.2 Βήματα Κατηγοριοποίησης

Πρόκειται για μια διαδικασία δύο σταδίων:

- **Βήμα εκμάθησης (Φάση εκπαίδευσης):** Αυτή περιλαμβάνει την κατασκευή του μοντέλου κατηγοριοποίησης. Χρησιμοποιούνται διαφορετικοί αλγόριθμοι για την κατασκευή ενός κατηγοριοποιητή. Αυτό γίνεται με την εκπαίδευση του μοντέλου χρησιμοποιώντας κάποιο διαθέσιμο σύνολο εκπαίδευσης. Το μοντέλο πρέπει να είναι εκπαιδευμένο για την ακριβή πρόβλεψη αποτελεσμάτων.
- **Βήμα κατηγοριοποίησης:** Το μοντέλο χρησιμοποιείται για την πρόβλεψη των ετικετών κλάσης και τη δοκιμή του κατασκευασμένου μοντέλου σε δεδομένα δοκιμών, και ως εκ τούτου, για την εκτίμηση της ακρίβειας των κανόνων κατηγοριοποίησης.

3.1.3 Εκπαίδευση και δοκιμή

Ας υποθέσουμε ότι υπάρχει ένα άτομο που κάθεται κάτω από έναν ανεμιστήρα, και ο ανεμιστήρας αρχίζει να πέφτει πάνω του. Έτσι, θα πρέπει να κάνει στην άκρη για να μην πληγωθεί. Αυτό είναι το εκπαιδευτικό μέρος. Κατά την διάρκεια των δοκιμών, αν το άτομο δει ένα βαρύ αντικείμενο να πέφτει επάνω του και μετακινηθεί στην άκρη, τότε το σύστημα έχει

περάσει τις δοκιμές θετικά. Από την άλλη εάν το άτομο δεν μετακινηθεί στην άκρη τότε το σύστημα δεν έχει περάσει τις δοκιμές.

Το ίδιο ισχύει και στην περίπτωση των δεδομένων. Το μοντέλο θα πρέπει να εκπαιδεύεται ώστε να επιτυγχάνει τα ακριβή και καλύτερα αποτελέσματα. Υπάρχουν συγκεκριμένοι τύποι δεδομένων, οι οποίοι σχετίζονται με την εξόρυξη δεδομένων και μας ενημερώνουν για τη μορφή του αρχείου (είτε είναι σε μορφή κειμένου είτε σε αριθμητική μορφή).

Χαρακτηριστικά – Αντιπροσωπεύουν διαφορετικές δυνατότητες ενός αντικειμένου. Οι διάφοροι τύποι χαρακτηριστικών είναι:

1. Δυαδικά: Διαθέτει μόνο δύο τιμές, δηλαδή Αληθής ή Ψευδής
Παράδειγμα: Ας υποθέσουμε ότι υπάρχει μια έρευνα για την αξιολόγηση κάποιου προϊόντος. Πρέπει να ελέγξουμε εάν το προϊόν είναι χρήσιμο ή όχι. Επομένως, ο πελάτης πρέπει να απαντήσει είτε Ναι είτε Όχι.
2. Συμμετρικά: Και οι δύο τιμές είναι εξίσου σημαντικές
3. Ασύμμετρα: Όταν και οι δύο τιμές μπορεί να μην είναι σημαντικές
4. Ονομαστικά: Όταν υπάρχει η πιθανότητα για περισσότερα από δύο αποτελέσματα. Είναι σε αλφαβητική μορφή και όχι σε κάποια ακέραιη μορφή
Παράδειγμα: Κάποιος πρέπει να επιλέξει κάποιο υλικό αλλά με διαφορετικό χρώμα. Το χρώμα μπορεί να είναι: Κίτρινο, Πράσινο, Μαύρο, Κόκκινο. Διαφορετικά χρώματα: Κόκκινο, Πράσινο, Μαύρο, Κίτρινο
5. Τακτικά Αριθμητικά: Τιμές που πρέπει να έχουν κάποια σειρά που βγάζει νόημα.
Παράδειγμα: Ας υποθέσουμε ότι υπάρχουν φύλλα βαθμολογίας μαθητών που μπορεί να περιέχουν διαφορετικούς βαθμούς ανάλογα με την απόδοση τους όπως: Α, Β, Γ, Δ, Βαθμοί: Α, Β, Γ, Δ
6. Συνεχή: Μπορεί να έχει άπειρο αριθμό τιμών πραγματικού τύπου.
Παράδειγμα: Μέτρηση βάρους λίγων μαθητών με σειρά, δηλαδή 50, 51, 52, 53 Βάρος: 50, 51, 52, 53
7. Διακριτά: Πεπερασμένος αριθμός τιμών
Παράδειγμα: Βαθμοί ενός μαθητή σε μερικά μαθήματα: 60, 65, 75, 80, 90 Βαθμοί: 60, 65, 75, 80, 90

3.1.4 Συντακτικό

Μαθηματική Σημειογραφία: Η κατηγοριοποίηση βασίζεται στη δημιουργία μιας συνάρτησης, η οποία λαμβάνει το διάνυσμα εισόδου «X» και προβλέπει το αποτέλεσμα της «Y» (Ποιοτική απόκριση που λαμβάνει τιμές από ένα σύνολο C) Στην περίπτωση αυτή, χρησιμοποιείται ο κατηγοριοποιητής (ή αλλιώς το μοντέλο) που είναι μια εποπτευόμενη λειτουργία και μπορεί να σχεδιαστεί με βάση τις γνώσεις του εμπειρογνώμονα. Έχει κατασκευαστεί για την πρόβλεψη ετικετών κλάσης (Παράδειγμα: Ετικέτα «Ναι» ή «Όχι» για την έγκριση κάποιου συμβάντος).

Οι κατηγοριοποιητές μπορούν να ταξινομηθούν σε δύο κύριους τύπους:

1. **Διακρίνων:** Είναι ένας πολύ βασικός κατηγοριοποιητής και καθορίζει μόνο μιας κλάση για κάθε σειρά δεδομένων. Προσπαθεί να κάνει την μοντελοποίηση βασιζόμενος στα παρατηρούμενα δεδομένα και η λειτουργία του εξαρτάται σε μεγάλο βαθμό από την ποιότητα των δεδομένων.

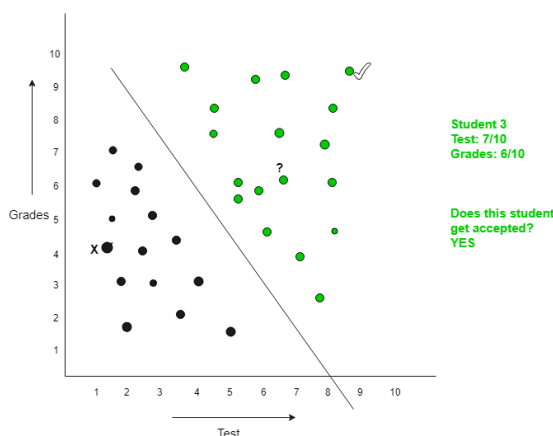
Παράδειγμα: Λογιστική Παλινδρόμηση

Αποδοχή φοιτητών σε πανεπιστήμιο (Τα διαγωνίσματα και οι βαθμοί πρέπει να ληφθούν υπόψη). Ας υποθέσουμε ότι υπάρχουν λίγοι φοιτητές και το αποτέλεσμα τους είναι το ακόλουθο:

Student 1 : Test Score: 9/10, Grades: 8/10 Result: Accepted

Student 2 : Test Score: 3/10, Grades: 4/10, Result: Rejected

Student 3 : Test Score: 7/10, Grades: 6/10, Result: To be tested



Εικόνα 5: Λογιστική Παλινδρόμηση των φοιτητών

2. **Γενετικός:** Παρουσιάζει την κατανομή των επιμέρους κατηγοριών και προσπαθεί να εκπαιδεύσει το μοντέλο που δημιουργεί τα δεδομένα στο παρασκήνιο, υπολογίζοντας τις παραδοχές και τις κατανομές του μοντέλου. Χρησιμοποιείται για να προβλέψει δεδομένα τα οποία δεν φαίνονται εξ αρχής.

Παράδειγμα: Ταξινομητής Naive Bayes

Εντοπισμός ανεπιθύμητων μηνυμάτων ηλεκτρονικού ταχυδρομείου με τον έλεγχο προηγούμενων δεδομένων. Ας υποθέσουμε ότι έχουμε 100 μηνύματα ηλεκτρονικού ταχυδρομείου και αυτά μοιράζονται ως εξής: Κλάση A: 25% ανεπιθύμητα μηνύματα ηλεκτρονικού ταχυδρομείου και Κλάση B: 75% μη ανεπιθύμητα μηνύματα ηλεκτρονικού ταχυδρομείου. Αν ο χρήστης θέλει να ελέγξει αν ένα μήνυμα ηλεκτρονικού ταχυδρομείου περιέχει την λέξη φθηνό τότε αυτό μπορεί να ονομαστεί ανεπιθύμητη αλληλογραφία. Φαίνεται ότι στην κατηγορία A (δηλαδή στο 25% των δεδομένων), 20 από τα 25 μηνύματα ηλεκτρονικού ταχυδρομείου είναι ανεπιθύμητα. Στην κατηγορία B (δηλαδή στο 75% των δεδομένων), 70 τα 75 email δεν είναι ανεπιθύμητα μηνύματα και τα υπόλοιπα είναι. Οπότε, αν το email περιέχει τη λέξη φθηνό η πιθανότητα να είναι ανεπιθύμητη αλληλογραφία είναι 80%.

3.1.5 Ταξινομητές Μηχανικής Μάθησης και εργαλεία

Ταξινομητές:

1. Δένδρα Αποφάσεων
2. Ταξινομητές Bayesian
3. Νευρωνικά Δίκτυα
4. KNN
5. Γραμμική Παλινδρόμηση
6. Λογιστική Παλινδρόμηση

Μερικούς από αυτούς τους ταξινομητές θα τους περιγράψουμε παρακάτω στο κεφάλαιο.

Εργαλεία και γλώσσες:

- Κύριες Γλώσσες προγραμματισμού που χρησιμοποιούνται: R, SAS, Python, SQL
- Κύρια Εργαλεία που χρησιμοποιούνται: RapidMiner, Orange, KNIME, Spark, Weka
- Βιβλιοθήκες που χρησιμοποιούνται: Jupyter, NumPy, Matplotlib, Pandas, ScikitLearn, NLTK, TensorFlow, Seaborn, Basemapκ.λπ.

3.1.6 Εφαρμογές

- Ανάλυση καλαθιού αγοράς: Είναι μια τεχνική μοντελοποίησης που έχει συσχετιστεί με συχνές συναλλαγές αγοράς ενός συνδυασμού ειδών.

Παράδειγμα: Το Amazon και πολλοί άλλοι πωλητές λιανικής χρησιμοποιούν αυτή την τεχνική. Κατά την προβολή κάποιων προϊόντων, εμφανίζονται ορισμένες προτάσεις για βασικά προϊόντα που έχουν αγοράσει κάποιοι στο παρελθόν.

- Πρόβλεψη Καιρού: Τα μεταβαλλόμενα μοτίβα στις καιρικές συνθήκες πρέπει να παρατηρούνται με βάση παραμέτρους όπως η θερμοκρασία, η υγρασία, η κατεύθυνση του ανέμου. Αυτή η έντονη παρατήρηση απαιτεί επίσης τη χρήση προηγούμενων εγγραφών προκειμένου να προβλεφθεί με ακρίβεια.

Γενικά οι ταξινομητές εφαρμόζονται στους ακόλουθους τομείς:

- Μάρκετινγκ και λιανικό εμπόριο
- Κατασκευαστικός τομέας
- Τηλεπικοινωνίες
- Εντοπισμός Εισβολέα
- Εκπαιδευτικό σύστημα
- Εντοπισμός απάτης

3.1.7 Πλεονεκτήματα και Μειονεκτήματα

Πλεονεκτήματα

- Οι μέθοδοι που βασίζονται στην εξόρυξη είναι οικονομικά αποδοτικοί και αποτελεσματικοί
- Βοηθά στον εντοπισμό υπόπτων για εγκλήματα
- Βοηθά στην πρόβλεψη του κινδύνου ασθενειών
- Βοηθά τις τράπεζες και τα χρηματοπιστωτικά ιδρύματα να εντοπίζουν αθέτηση έτσι ώστε να μπορούν να εγκρίνουν κάρτες, δάνεια κ.λπ.

Μειονεκτήματα

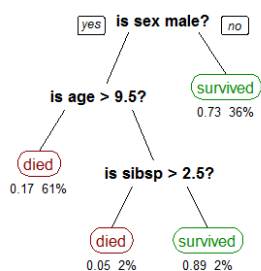
- Προστασία Προσωπικών δεδομένων: Υπάρχει η πιθανότητα, μια εταιρεία να δώσει ορισμένες πληροφορίες σχετικά με τους πελάτες της σε άλλους προμηθευτές ή να χρησιμοποιήσει αυτές τις πληροφορίες προς όφελός τους.
- Πρόβλημα ακρίβειας: Η επιλογή του μοντέλου ακριβείας πρέπει να υπάρχει, για να υπάρχει η βέλτιστη ακρίβεια και αποτέλεσμα. (Basic Concept of Classification (Data Mining), n.d.)

3.2 Δένδρα Αποφάσεων

3.2.1 Βασικά Στοιχεία

Ένα δέντρο έχει πολλές αναλογίες στην πραγματική ζωή, και φαίνεται ότι έχει επηρεάσει ένα ευρύ πεδίο μηχανικής μάθησης, καλύπτοντας τόσο την ταξινόμηση όσο και την οπισθοδρόμηση. Στην ανάλυση αποφάσεων, ένα **δένδρο αποφάσεων** μπορεί να χρησιμοποιηθεί για την οπτική και ρητή παρουσίαση των αποφάσεων αλλά και της λήψης αποφάσεων. Όπως λέει και το όνομα, χρησιμοποιεί ένα τρισδιάστατο μοντέλο αποφάσεων. Αν και είναι ένα κοινώς χρησιμοποιούμενο εργαλείο στην εξόρυξη δεδομένων για την εξαγωγή μιας στρατηγικής για την επίτευξη ενός συγκεκριμένου στόχου, χρησιμοποιείται επίσης ευρέως στον τομέα της μηχανικής μάθησης.

Αναπαράσταση Αλγορίθμου σαν δένδρο



Εικόνα 6: Δένδρο απόφασης σχετικά με την επιβίωση επιβάτη

Ας εξετάσουμε ένα πολύ βασικό παράδειγμα που χρησιμοποιεί ένα σύνολο δεδομένων του Τιτανικού για να προβλέψει εάν ένας επιβάτης θα επιβιώσει. Το μοντέλο που ακολουθεί χρησιμοποιεί 3 χαρακτηριστικά από το σύνολο δεδομένων, δηλαδή το φύλλο, την ηλικία και το sibsp (αριθμός συζύγων ή παιδιών).

Ένα δένδρο αποφάσεων είναι αναποδογυρισμένο με την ρίζα του να είναι στην κορυφή. Στην εικόνα 6, το έντονο μαύρο κείμενο αντιπροσωπεύει μια συνθήκη / εσωτερικό κόμβο, βάσει της οποίας το

δένδρο χωρίζεται σε κλαδιά/ άκρα. Το τέλος του κλαδιού που δεν χωρίζεται πλέον είναι η απόφαση / φύλλο και στην περίπτωση αυτή είτε ο επιβάτης πέθανε είτε επέζησε, κάτι που αντιπροσωπεύεται ως κόκκινο και πράσινο κείμενο αντίστοιχα. Αν και ένα πραγματικό σύνολο δεδομένων θα έχει πολύ περισσότερα χαρακτηριστικά και αυτό θα είναι απλά ένα κλαδί σε ένα

πολύ μεγαλύτερο δένδρο, αλλά δεν μπορεί να αγνοηθεί η απλότητα αυτού του αλγορίθμου. Αυτή η μεθοδολογία είναι γνωστή ως δένδρο αποφάσεων μάθησης από τα δεδομένα και το δένδρο της εικόνας 6 καλείται δένδρο ταξινόμησης, καθώς στόχος είναι να ταξινομηθούν οι επιβάτες ως ζωντανοί ή νεκροί. Τα δένδρα παλινδρόμησης, αναπαρίστανται με τον ίδιο τρόπο, με την μόνη διαφορά ότι προβλέπουν συνεχείς τιμές όπως, για παράδειγμα, την τιμή ενός σπιτιού.

Γενικά οι αλγόριθμοι του δένδρου αποφάσεων αναφέρονται ως CART (Classification and Ratio Tree). Ας δούμε λοιπόν τι συμβαίνει παρασκηνιακά: Η ανάπτυξη ενός τέτοιου δένδρου περιλαμβάνει τη λήψη αποφάσεων για τα χαρακτηριστικά που πρέπει να επιλεγούν και τις συνθήκες που είναι αναγκαίο να χρησιμοποιηθούν για τον διαχωρισμό, καθώς και για το πότε πρέπει να σταματήσει. Σε αυτό το σημείο κρίνεται σκόπιμο να αναφερθούμε σε μια κοινή τεχνική που χρησιμοποιείται για τον διαχωρισμό και καλείται Αναδρομική Δυναδική Διαίρεση.

Αναδρομική Δυναδική Διαίρεση

Σε αυτή την τεχνική εξετάζονται όλα τα χαρακτηριστικά και δοκιμάζονται διαφορετικά σημεία διαχωρισμού με τη χρήση μιας συνάρτησης κόστους. Επιλέγεται η διαίρεση με το καλύτερο (χαμηλότερο) κόστος. Ας δούμε ξανά το προηγούμενο παράδειγμα δένδρου που έχει εκπαιδευτεί από το σύνολο δεδομένων του Τιτανικού. Στην πρώτη διαίρεση ή στην ρίζα, εξετάζονται όλα τα χαρακτηριστικά και τα δεδομένα εκπαίδευσης χωρίζονται σε ομάδες με βάση αυτή την διαίρεση. Έχουμε 3 χαρακτηριστικά, έτσι θα έχουμε 3 πιθανές διαιρέσεις. Τώρα θα υπολογίσουμε πόση ακρίβεια θα μας κοστίσει κάθε διαχωρισμός, χρησιμοποιώντας μια συνάρτηση. Επιλέγεται η διαίρεση που κοστίζει λιγότερο, η οποία στο παράδειγμά μας είναι το φύλλο του επιβάτη. Αυτός ο αλγόριθμος είναι αναδρομικός στην φύση του, καθώς οι ομάδες που σχηματίζονται μπορούν να υποδιαιρεθούν με την ίδια στρατηγική. Λόγω αυτής της διαδικασίας, αυτός ο αλγόριθμος είναι επίσης γνωστός ως άπληστος αλγόριθμος, καθώς υπάρχει υπερβολική επιθυμία για μείωση του κόστους. Αυτό καθιστά τον κόμβο-ρίζα ως τον καλύτερο κατηγοριοποιητή.

Κόστος διαίρεσης

Ας ρίξουμε μια προσεκτικότερη ματιά στις συναρτήσεις κόστους που χρησιμοποιούνται για την ταξινόμηση και την παλινδρόμηση. Και στις δύο περιπτώσεις οι συναρτήσεις κόστους προσπαθούν να βρουν τα πιο ομογενή κλαδιά ή κλαδιά που έχουν ομάδες με παρόμοιες

αποκρίσεις. Αυτό είναι λογικό, διότι μπορούμε να είμαστε σίγουροι ότι τα δεδομένα δοκιμής θα ακολουθήσουν μια συγκεκριμένη διαδρομή.

Παλινδρόμηση: άθροισμα($y - \text{πρόβλεψη}$)²

Ας υποθέσουμε ότι προσπαθούμε να προβλέψουμε τις τιμές των σπιτιών. Τώρα το δένδρο αποφάσεων θα αρχίσει να διαιρείται λαμβάνοντας υπόψη κάθε χαρακτηριστικό στα δεδομένα εκπαίδευσης. Ο μέσος όρος των απαντήσεων των εισόδων δεδομένων εκπαίδευσης μιας συγκεκριμένης ομάδας, θεωρείται πρόβλεψη για την εν λόγω ομάδα. Η παραπάνω συνάρτηση εφαρμόζεται σε όλα τα σημεία δεδομένων και το κόστος υπολογίζεται για όλες τις διαιρέσεις. Και σε αυτή την περίπτωση επιλέγεται ο διαχωρισμός με το χαμηλότερο κόστος. Μια άλλη συνάρτηση κόστους περιλαμβάνει την μείωση της τυπικής απόκλισης.

Ταξινόμηση: $G = \text{άθροισμα}(pk(1 - pk))$

Η βαθμολογία Gini δίνει μια ιδέα για το πόσο καλός είναι ο διαχωρισμός από το πόσο μικτές είναι οι κλάσεις απόκρισης στις ομάδες που δημιουργούνται από το διαχωρισμό. Στην περίπτωση αυτή, η pk είναι η αναλογία των ιδίων των κλάσεων εισόδου που υπάρχουν σε μια συγκεκριμένη ομάδα. Η απόλυτη «καθαρότητα» της κατηγορίας προκύπτει όταν μια ομάδα περιέχει όλες τις εισόδους από την ίδια κατηγορία, οπότε το pk είναι είτε 1 είτε 0 και $G = 0$, όπου ως κόμβος με διαίρεση κλάσεων 50-50 σε μια ομάδα έχει τη χειρότερη καθαρότητα, έτσι για δυαδική ταξινόμηση θα έχει $pk = 0.5$ και $G = 0.5$.

Πότε σταματάει η διαίρεση

Αν κάποιος αναρωτιέται πότε πρέπει να σταματάει να μεγαλώνει ένα δένδρο, η απάντηση είναι η εξής: Καθώς ένα πρόβλημα έχει συνήθως πολλά χαρακτηριστικά, αυτό έχει σαν αποτέλεσμα μεγάλο αριθμό διαχωρισμών, γεγονός που με την σειρά του δίνει ένα τεράστιο δένδρο. Αυτά τα δένδρα είναι σύνθετα και μπορούν να οδηγήσουν σε υπερπροσαρμογή. Ένας τρόπος για να γνωρίζει κάποιος πότε πρέπει να σταματήσει να αναπτύσσει περαιτέρω το δένδρο, είναι να ορίσει ένας ελάχιστος αριθμός εισόδων εκπαίδευσης που θα χρησιμοποιούνται σε κάθε φύλλο. Για παράδειγμα, μπορούν να χρησιμοποιηθούν τουλάχιστον 10 επιβάτες για να ληφθεί μια απόφαση (πέθανε ή επέζησε) και να αγνοηθεί κάθε φύλλο, το οποίο δέχεται λιγότερους από 10 επιβάτες. Ένας άλλος τρόπος είναι να οριστεί το μέγιστο βάθος του μοντέλου. Το μέγιστο βάθος αναφέρεται στο μήκος της μεγαλύτερης διαδρομής από την ρίζα σε φύλλο.

Περικοπή

Η απόδοση ενός δένδρου μπορεί να αυξηθεί περεταίρω με περικοπή. Περιλαμβάνει την αφαίρεση των κλαδιών που χρησιμοποιούν χαρακτηριστικά, τα οποία δεν είναι σημαντικά. Με αυτόν τον τρόπο, μειώνουμε την πολυπλοκότητα του δένδρου και, ως εκ τούτου, αυξάνουμε την προβλεπόμενη ισχύ του. Η περικοπή μπορεί να ξεκινήσει είτε από τη ρίζα είτε από τα φύλλα. Η απλούστερη μέθοδος περικοπής ξεκινά από τα φύλλα και αφαιρεί κάθε κόμβο με την πιο δημοφιλή κλάση σε αυτό το φύλλο, αυτή η αλλαγή διατηρείται αν δεν χειροτερέψει την ακρίβεια. Επίσης, αποκαλείται και περικοπή σφαλμάτων. Μπορούν να χρησιμοποιηθούν πιο εξελιγμένες μέθοδοι περικοπής, όπως η περικοπή της πολυπλοκότητας του κόστους, όπου χρησιμοποιείται μιας παράμετρος εκμάθησης (άλφα) για να σταθμιστεί εάν είναι δυνατή η αφαίρεση κόμβων με βάση το μέγεθος του υποδένδρου. Αυτό είναι επίσης ως περικοπή πιο αδύναμων συνδέσμων.

Πλεονεκτήματα CART(Classification and Ratio Tree)

- Απλό στην κατανόηση, την ερμηνεία και την οπτικοποίηση.
- Τα δένδρα αποφάσεων εκτελούν στο παρασκήνιο τον έλεγχο μεταβλητής ή την επιλογή χαρακτηριστικών.
- Μπορεί να χειριστεί αριθμητικά δεδομένα και δεδομένα κατηγοριών. Μπορεί επίσης να χειριστεί προβλήματα πολλαπλών εξόδων.
- Τα δένδρα αποφάσεων απαιτούν σχετικά μικρή προσπάθεια από τους χρήστες για την προετοιμασία των δεδομένων.
- Οι μη γραμμικές σχέσεις μεταξύ των παραμέτρων δεν επηρεάζουν την απόδοση του δένδρου.

Μειονεκτήματα CART

- Οι κατασκευαστές δένδρων αποφάσεων μπορεί να δημιουργήσουν υπερβολικά σύνθετα δένδρα που δεν γενικεύουν σωστά τα δεδομένα.
- Τα δένδρα αποφάσεων μπορεί να είναι ασταθή, επειδή μικρές διακυμάνσεις στα δεδομένα μπορεί να έχουν ως αποτέλεσμα τη δημιουργία ενός εντελώς διαφορετικού δένδρου.

- Οι άπληστοι αλγόριθμοι δεν μπορούν να εγγυηθούν την παραγωγή του βέλτιστου δένδρου αποφάσεων. Αυτό μπορεί να μετριαστεί με την εκπαίδευση πολλών δένδρων , όπου τα χαρακτηριστικά και τα δείγματα λαμβάνονται τυχαία με αντικατάσταση.
- Οι δημιουργοί δένδρων απόφασης δημιουργούν μεροληπτικά δένδρα εάν κυριαρχούν κάποιες κλάσεις. Επομένως, συνιστάται η εξισορρόπηση του συνόλου δεδομένων πριν από την προσαρμογή με το δένδρο αποφάσεων. (Gupta, 2017)

3.2.2 Αλγόριθμος ID3

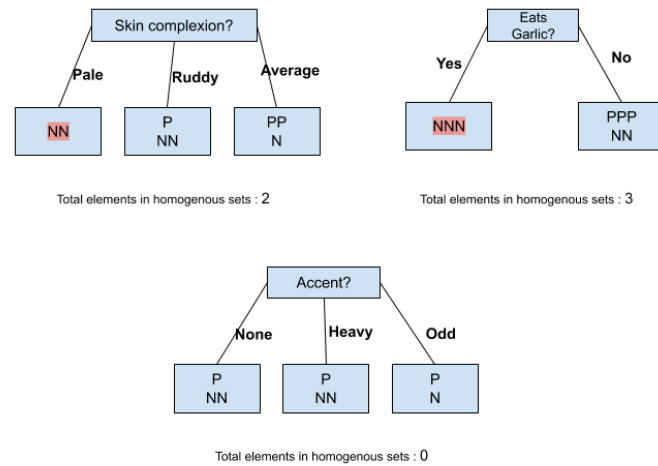
Έστω ότι έχουμε τις παρατηρήσεις της εικόνας 7. Όπως βλέπουμε το ζητούμενο είναι να φτιαχτεί ένα δένδρο απόφασης σχετικά με το αν ένας άνθρωπος είναι βρικόλακας ή όχι. Ο πίνακας έχει ένα σύνολο από διάφορα χαρακτηριστικά (αν έχει σκιά, αν τρώει σκόρδο, τι επιδερμίδα έχει και τι προφορά έχει).

#	Casts shadow?	Eats garlic?	Skin complexion?	accent?	Vampire?
1	? (don't know)	Yes	Pale	None	<u>No</u>
2	Yes	Yes	Ruddy	None	<u>No</u>
3	? (don't know)	No	Ruddy	None	<u>Yes</u>
4	No	No	Average	Heavy	<u>Yes</u>
5	? (don't know)	No	Average	Odd	<u>Yes</u>
6	Yes	No	Pale	Heavy	<u>No</u>
7	Yes	No	Average	Heavy	<u>No</u>
8	? (don't know)	Yes	Ruddy	Odd	<u>No</u>

Εικόνα 7: Παρατηρήσεις για τον αν κάποιος είναι βρικόλακας (τα ? δείχνουν ότι για το συγκεκριμένο δείγμα δεν έχει αποφασιστεί ακόμα η τιμή)

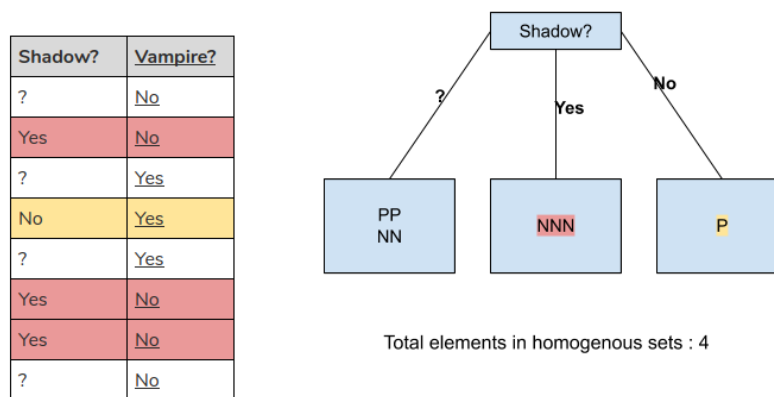
Αρχικά, για να φτιάξουμε το δένδρο κοιτάμε το σύνολο δεδομένων που έχουμε διαθέσιμο και αποφασίζουμε ποιο χαρακτηριστικό θα χρησιμοποιήσουμε για την ρίζα του δένδρου. Αυτή είναι μια δυαδική ταξινόμηση, έτσι στο τέλος του δένδρου αποφάσεων θα έχουμε δύο πιθανά αποτελέσματα (είτε είναι βαμπίρ είτε όχι), έτσι κάθε παράδειγμα εισαγωγής θα ταξινομηθεί ως αληθές (θετικό παράδειγμα) και ψευδές (αρνητικό παράδειγμα). Εδώ το «P» που φαίνεται στην παρακάτω εικόνα αναφέρεται στο θετικό, που σημαίνει ότι ένα άτομο είναι βαμπίρ, και το «N» αναφέρεται στο αρνητικό, που σημαίνει ότι το άτομο δεν είναι βαμπίρ. Θέλουμε ιδιότητες, οι οποίες διαχωρίζουν περισσότερα δεδομένα σε ομοιογενή σύνολα, το οποίο σημαίνει ότι θέλουμε τέτοια σύνολα στα οποία υπάρχει μόνο το Π ή μόνο το N, διότι με αυτό τον τρόπο, μπορούμε να

απαντήσουμε με σιγουριά αν είναι βαμπίρ ή όχι. Επομένως αυτοί θα είναι τα φύλλα του δένδρου.



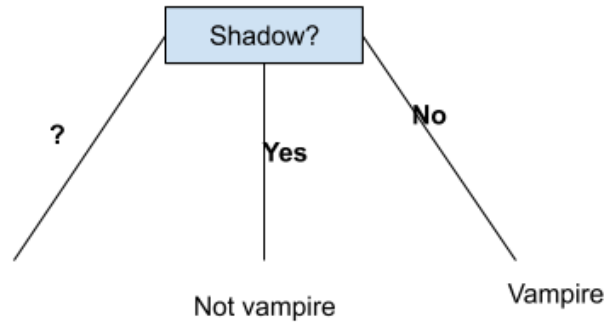
Εικόνα 8: Πιθανά Φύλλα του δένδρου αποφάσεων

Ελέγχουμε το κάθε χαρακτηριστικό και βλέπουμε ποιο έχει το μεγαλύτερο αριθμό στοιχείων στο ομοιογενές σύνολο. Στην συγκεκριμένη περίπτωση βλέπουμε ότι το χαρακτηριστικό «Σκιά» έχει την υψηλότερη καταμέτρηση στοιχείων στο ομοιογενές σύνολο, έτσι επιλέγουμε αυτό το σύνολο.



Εικόνα 9: Το χαρακτηριστικό «Σκιά» έχει την υψηλότερη καταμέτρηση στοιχείων στο ομοιογενές σύνολο

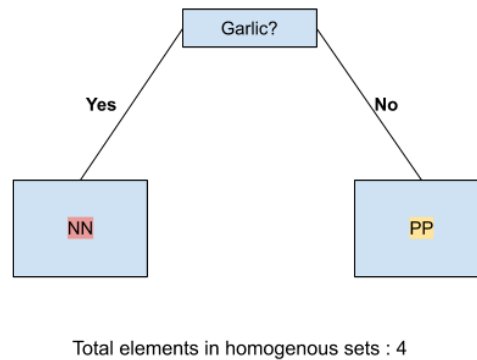
Έτσι μέχρι αυτό το σημείο έχουμε το παρακάτω δένδρο:



Εικόνα 10: Η ρίζα του δένδρου απόφασης

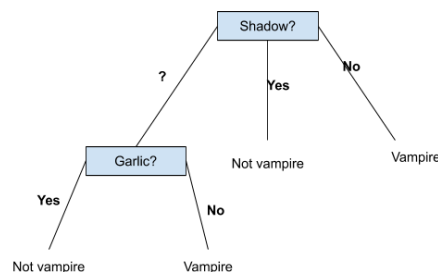
Για το χαρακτηριστικό της σκιάς «Ναι» και «Όχι » μπορούμε να αποφασίσουμε εάν ένα άτομο είναι βαμπίρ ή όχι, αλλά σε περίπτωση που συναντήσουμε το ? δεν γνωρίζουμε με ακρίβεια και πρέπει να αποφασίσουμε ποιο χαρακτηριστικό διαιρεί τα δεδομένα με καλό τρόπο, όταν η σκιά έχει το χαρακτηριστικό ?. Επομένως ας αναλύσουμε ένα άλλο χαρακτηριστικό ενώ η σκιά είναι άγνωστη.

Garlic?	Shadow?	Vampire?
Yes	?	No
Yes	Yes	No
No	?	Yes
No	No	Yes
No	?	Yes
No	Yes	No
No	Yes	No
Yes	?	No



Στην διπλανή εικόνα βλέπουμε ότι το χαρακτηριστικό «σκόρδο» διαιρεί τα μέγιστα στοιχεία, και μάλιστα, όλα τα στοιχεία, σε ομοιογενή σύνολα. Έτσι το δένδρο μας θα μοιάζει κάπως έτσι:

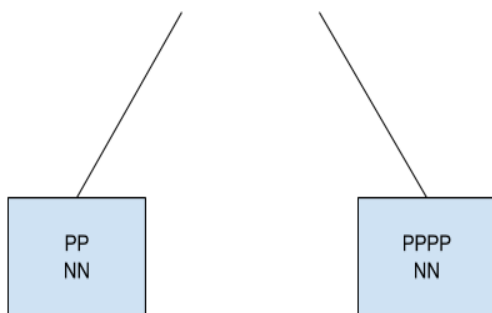
Εικόνα 11: Το χαρακτηριστικό "Σκόρδο"



Εικόνα 12: Το δένδρο απόφασης μετά την προσθήκη του χαρακτηριστικού "Σκόρδο"

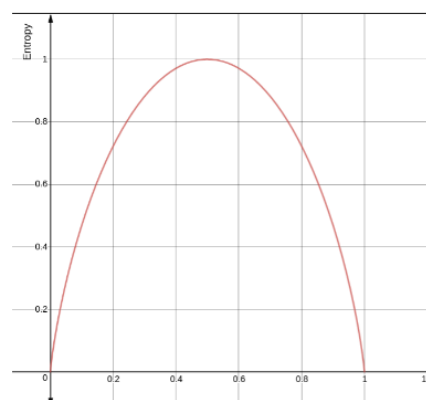
Σε ένα μεγάλο σύνολο δεδομένων όμως, ο παραπάνω τρόπος δεν είναι σωστός, διότι σε μια τέτοια περίπτωση δεν είναι απαραίτητο ότι θα έχουμε χαρακτηριστικά που διαχωρίζονται στο ομοιογενές σύνολο. Αυτό σημαίνει ότι μπορεί να βρούμε για όλα τα χαρακτηριστικά του συνόλου, στοιχεία τα οποία στο ομοιογενές σύνολο είναι 0.

Και εδώ έρχεται η αναγκαιότητα της χρήσης του αλγορίθμου ID3 για την δημιουργία δένδρων αποφάσεων, η οποία χρησιμοποιεί την έννοια της απόκτησης πληροφοριών, η οποία ορίζεται από την πλευρά της εντροπίας, τη θεμελιώδη ποσότητα στην θεωρία πληροφορίας. Εντροπία πληροφοριών είναι ο μέσος ρυθμός με τον οποίο παράγονται οι πληροφορίες από μια πηγή δεδομένων. Στην παρακάτω εικόνα βλέπουμε αυτές τις δύο διαιρέσεις ενός χαρακτηριστικού:



Παρατηρούμε ότι στα αριστερά έχει τον ίδιο αριθμό από «P» «N», έτσι ώστε να μην μας δίνεται καμία υπόνοια για την απόφαση. Από την άλλη ο κόμβος στα δεξιά έχει περισσότερα «P» από «N» και έτσι μπορεί να μας κατευθύνει κάπως προς το «P», έτσι ανάμεσα στους κόμβους μπορεί να

θεωρήσουμε σωστό τον δεξιό. Έτσι αντί να τους δώσουμε και στους δύο την τιμή 0 θα δοκιμάσουμε έναν άλλο τρόπο. Ας πούμε ότι ένας κόμβος, ο οποίος έχει το ίδιο πλήθος από «P» και «N», έχει την υψηλότερη εντροπία (1) και κάποιος που έχει μόνο «P» ή «N», έχει τη χαμηλότερη εντροπία (0). Μπορούμε να έχουμε κάτι σαν αυτό $P/(P+N)$ ενάντια στο γράφημα της εντροπίας.



Εικόνα 13: Γράφημα εντροπίας μιας πηγής πληροφορίας

Έτσι, όταν το $P = N$, τότε $P/(P+N) = 0.5$ και η εντροπία $= 1$. Επίσης αν $P = k$ (οποιοσδήποτε ακέραιος) και $N = 0$ τότε η εντροπία $= 0$. Μοιάζει με ένα αρκετά κατάλληλο γράφημα για να πετύχουμε αυτό που θέλουμε οπότε υπάρχει κάποιος μαθηματικός τρόπος για να πάρουμε αυτό το γράφημα. Η παραπάνω καμπύλη μπορεί να επιτευχθεί με την ακόλουθη εξίσωση:

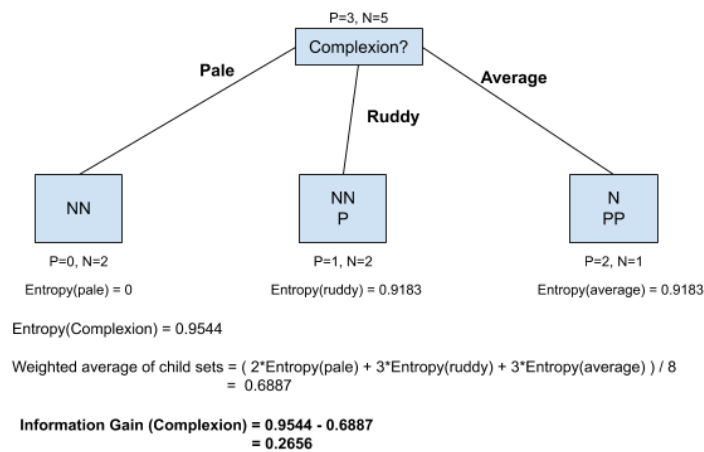
$$y = -x \log_2(x) - (1-x) \log_2(1-x)$$

Η παραπάνω εξίσωση μπορεί να γραφτεί στην μορφή $P/(P+N)$ αλλά και της εντροπίας αντικαθιστώντας: $x = P/(P+N)$ και $y = \text{Εντροπία}$. Έτσι έχουμε:

$$\begin{aligned} \text{Entropy} &= -\frac{P}{P+N} \log_2\left(\frac{P}{P+N}\right) - \left(1 - \frac{P}{P+N}\right) \log_2\left(1 - \frac{P}{P+N}\right) \\ &= -\frac{P}{P+N} \log_2\left(\frac{P}{P+N}\right) - \left(\frac{N}{P+N}\right) \log_2\left(\frac{N}{P+N}\right) \\ \text{Entropy} &= -\frac{P}{P+N} \log_2\left(\frac{P}{P+N}\right) - \left(\frac{N}{P+N}\right) \log_2\left(\frac{N}{P+N}\right) \end{aligned}$$

Εικόνα 14: Η εξίσωση στην μορφή $P/(P+N)$ και εντροπίας

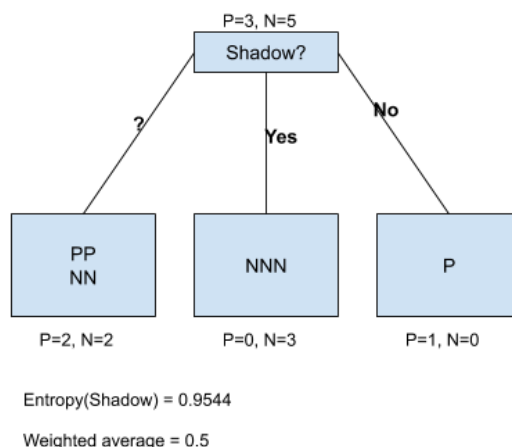
Όπου P και N το πλήθος των P και N ενός χαρακτηριστικού. Θέλουμε να βρούμε πληροφορίες για το χαρακτηριστικό, το οποίο ορίζεται ως IG (Information Gain – κέρδος πληροφοριών από κάποιο χαρακτηριστικό A είναι η αναμενόμενη μείωση της εντροπίας). $IG(\text{Χαρακτηριστικό}) = \text{εντροπία χαρακτηριστικού} - \text{Σταθμισμένος μέσος όρος εντροπίας κάθε συνόλου παιδιών}$. Για παράδειγμα:



Εικόνα 15: Εντροπία και κέρδος πληροφοριών για το Χαρακτηριστικό "Επιδερμίδα"

Μιας και έχουμε μια ιδέα για την εντροπία και το κέρδος πληροφοριών, ας χτίσουμε εκ νέου το δένδρο αποφάσεων εκ του μηδενός χρησιμοποιώντας μια νέα προσέγγιση:

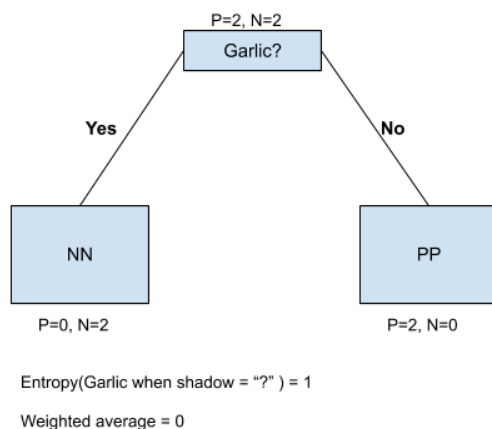
Shadow?	Vampire?
?	<u>No</u>
Yes	<u>No</u>
?	<u>Yes</u>
No	<u>Yes</u>
?	<u>Yes</u>
Yes	<u>No</u>
Yes	<u>No</u>
?	<u>No</u>



$$\text{Information Gain} = 0.9544 - 0.5 = 0.4544$$

Παρατηρούμε ότι λαμβάνουμε το μέγιστο κέρδος πληροφοριών από το χαρακτηριστικό σκιά. Έτσι το επιλέγουμε σαν ρίζα του δένδρου. Έτσι η ρίζα του δένδρου είναι ίδια με την εικόνα 10. Πρέπει να βρούμε άλλο ένα χαρακτηριστικό για Σκιά = ?.

Garlic?	Shadow?	Vampire?
Yes	?	<u>No</u>
Yes	Yes	<u>No</u>
No	?	<u>Yes</u>
No	No	<u>Yes</u>
No	?	<u>Yes</u>
No	Yes	<u>No</u>
No	Yes	<u>No</u>
Yes	?	<u>No</u>



$$\text{Information Gain} = 1 - 0 = 1$$

Λαμβάνουμε το μέγιστο κέρδος πληροφορίας από το σκόρδο. Έτσι το δένδρο μας θα μοιάζει με εκείνο της εικόνας 12. Αυτό είναι ακριβώς το ίδιο με την προηγούμενη προσέγγιση, διότι σε κάθε βήμα καταφέραμε να βρούμε κάποια χαρακτηριστικά που χωρίζονται σε ένα ομοιογενές σύνολο, αλλά η δεύτερη προσέγγιση είναι πιο ισχυρή και μπορεί να εφαρμοστεί για να δημιουργηθεί ένα δένδρο αποφάσεων από ένα μεγάλο σύνολο δεδομένων. (Maurya, 2019)

3.3 Ταξινομητές Bayes

3.1.1 Βασικά Στοιχεία

Οι ταξινομητές Naive Bayes είναι μια συλλογή αλγορίθμων ταξινόμησης που βασίζονται στο θεώρημα Bayes. Δεν πρόκειται για έναν μόνο αλγόριθμο, αλλά για μια οικογένεια αλγορίθμων όπου όλοι έχουν μια κοινή αρχή, δηλαδή κάθε ζεύγος χαρακτηριστικών που κατηγοριοποιούνται είναι ανεξάρτητο το ένα από το άλλο. Καταρχάς, ας εξετάσουμε ένα σύνολο δεδομένων.

Σκεφτείτε ένα φανταστικό σύνολο δεδομένων, το οποίο περιγράφει τις καιρικές συνθήκες για ένα παιχνίδι γκολφ. Δεδομένων των καιρικών συνθηκών, κάθε γραμμή ταξινομεί τις συνθήκες ανάλογα με το αν είναι κατάλληλος ο καιρός για παιχνίδι («Ναι») ή είναι ακατάλληλος («Όχι»).

Ακολουθεί μια απεικόνιση σε πίνακα του συνόλου των δεδομένων μας:

	OUTLOOK	TEMPERATURE	HUMIDITY	WINDY	PLAY GOLF
0	Rainy	Hot	High	False	No
1	Rainy	Hot	High	True	No
2	Overcast	Hot	High	False	Yes
3	Sunny	Mild	High	False	Yes
4	Sunny	Cool	Normal	False	Yes
5	Sunny	Cool	Normal	True	No
6	Overcast	Cool	Normal	True	Yes
7	Rainy	Mild	High	False	No
8	Rainy	Cool	Normal	False	Yes
9	Sunny	Mild	Normal	False	Yes
10	Rainy	Mild	Normal	True	Yes
11	Overcast	Mild	High	True	Yes
12	Overcast	Hot	Normal	False	Yes
13	Sunny	Mild	High	True	No

Εικόνα 16: Το σύνολο των δεδομένων σε μορφή πίνακα

Το σύνολο δεδομένων διαιρείται σε δύο μέρη, δηλαδή τον πίνακα των χαρακτηριστικών και το διάλυμα απόκρισης.

- Ο πίνακας χαρακτηριστικών περιέχει όλες τις γραμμές του συνόλου των δεδομένων στα οποία κάθε γραμμή αποτελείται από την τιμή των εξαρτημένων χαρακτηριστικών. Στο

παραπάνω σύνολο χαρακτηριστικών τα χαρακτηριστικά είναι: “Outlook”, “Temperature”, “Humidity” και “Windy”.

- Το διάνυσμα απόκρισης περιέχει την τιμή της μεταβλητής κλάσης (πρόβλεψη ή έξοδος) για κάθε γραμμή του πίνακα χαρακτηριστικών. Στον παραπάνω πίνακα, το όνομα της μεταβλητής κλάσης είναι «Play Golf».

Υπόθεση:

Η βασική υπόθεση του naïve Bayes είναι ότι κάθε χαρακτηριστικό είναι:

- Ανεξάρτητο
- Ίσο
- Έχει συμβολή στο αποτέλεσμα

Σε σχέση με το σύνολο δεδομένων μας, η έννοια αυτή μπορεί να γίνει αντιληπτή ως εξής:

- Υποθέτουμε ότι δύο χαρακτηριστικά δεν είναι εξαρτώμενα. Για παράδειγμα η θερμοκρασία «Hot» δεν έχει καμία σχέση με την υγρασία ή το αν ο καιρός είναι «Rainy» δεν έχει καμία επίπτωση στους ανέμους. Ως εκ τούτου, θεωρούμε ότι τα χαρακτηριστικά είναι ανεξάρτητα.
- Κάθε χαρακτηριστικό έχει την ίδια βαρύτητα ή σημασία. Για παράδειγμα, η γνώση μόνο της θερμοκρασίας και της υγρασίας δεν μπορεί να προβλέψει το αποτέλεσμα με ακρίβεια. Κανένα από τα χαρακτηριστικά αυτά δεν είναι άσχετο και θεωρείται ότι συμβάλει εξίσου στο αποτέλεσμα.

Σημείωση: Οι παραδοχές που γίνονται από τους Naive Bayes δεν είναι γενικά ορθές σε πραγματικές καταστάσεις. Στην πραγματικότητα η υπόθεση της ανεξαρτησίας δεν είναι ποτέ σωστή αλλά συχνά λειτουργεί καλά.

Τώρα πριν προχωρήσουμε στις λεπτομέρειες για το Naive Bayes, είναι σημαντικό να γνωρίζουμε το Θεώρημα του Bayes.

3.3.2 Θεώρημα Bayes

Το θεώρημα του Bayes βρίσκει την πιθανότητα να συμβεί ένα συμβάν, δεδομένης της πιθανότητας ενός άλλου συμβάντος που έχει ήδη συμβεί. Το Θεώρημα του Bayes αναφέρεται

μαθηματικά σαν την ακόλουθη εξίσωση: $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

Όπου τα A και B είναι τα γεγονότα και το P(B) δεν είναι 0.

Βασικά προσπαθούμε να βρούμε την πιθανότητα ενός γεγονότος A, δεδομένου ότι το γεγονός B είναι αλήθεια. Το γεγονός B αναφέρεται επίσης και ως αποδεικτικό στοιχείο. P(A) είναι η πιθανότητα να συμβεί το A πριν από την εμφάνιση των αποδεικτικών στοιχείων. Η απόδειξη είναι μια τιμή χαρακτηριστικού ενός άγνωστου στιγμιότυπου (εδώ είναι το γεγονός B). Το P(A|B) είναι η πιθανότητα γεγονότος μετά την εμφάνιση των αποδεικτικών στοιχείων. Τώρα, Τώρα, όσον αφορά το σύνολο των δεδομένων μας, μπορούμε να εφαρμόσουμε το θεώρημα του Bayes με τον ακόλουθο τρόπο: $P(y|X) = \frac{P(X|y)P(y)}{P(X)}$ όπου y είναι μεταβλητή κλάσης και X είναι ένα εξαρτημένο διάνυσμα χαρακτηριστικών όπου: $X = (x_1, x_2, x_3, \dots, x_n)$.

Ένα παράδειγμα ενός διανύσματος χαρακτηριστικών και μιας αντίστοιχης μεταβλητής κλάσης μπορεί να είναι: $X = (\text{Rainy}, \text{Hot}, \text{High}, \text{False})$, $y = \text{No}$. Οπότε, το $P(X|y)$ εδώ σημαίνει, την πιθανότητα «No playing golf» δεδομένου ότι οι καιρικές συνθήκες είναι «Rainy», «Hot», «High», «False».

3.3.3 Αφελής Υπόθεση

Τώρα, ας κάνουμε μια αφελή παραδοχή στο θεώρημα του Bayes, η οποία είναι η ανεξαρτησία μεταξύ των χαρακτηριστικών. Οπότε τώρα μοιράζουμε τα στοιχεία στα ανεξάρτητα μέρη. Τώρα αν δύο γεγονότα A και B είναι ανεξάρτητα, τότε: $P(A,B) = P(A)P(B)$

Ως εκ τούτου καταλήγουμε στο αποτέλεσμα:

$$P(y|x_1, \dots, x_n) = \frac{P(x_1|y)P(x_2|y) \dots P(x_n|y)P(y)}{P(x_1)P(x_2) \dots P(x_n)}$$

Το οποίο μπορεί να γραφτεί ως:

$$P(y|x_1, \dots, x_n) = \frac{P(y) \prod_{i=1}^n P(x_i|y)}{P(x_1)P(x_2) \dots P(x_n)}$$

Τώρα, καθώς ο παρονομαστής παραμένει σταθερός για μια δεδομένη είσοδο, μπορούμε να αφαιρέσουμε αυτόν τον όρο:

$$P(y|x_1, \dots, x_n) \propto P(y) \prod_{i=1}^n P(x_i|y)$$

Τώρα πρέπει να δημιουργήσουμε ένα μοντέλο ταξινομητή. Για να γίνει αυτό βρίσκουμε την πιθανότητα ενός δοθέντος συνόλου εισόδων για όλες τις πιθανές τιμές της μεταβλητής κλάσης y και παίρνουμε την έξοδο με την μεγαλύτερη πιθανότητα, Αυτό μπορεί να εκφραστεί μαθηματικά ως εξής:

$$y = \operatorname{argmax}_y P(y) \prod_{i=1}^n P(x_i|y)$$

Outlook				
	Yes	No	P(yes)	P(no)
Sunny	2	3	2/9	3/5
Overcast	4	0	4/9	0/5
Rainy	3	2	3/9	2/5
Total	9	5	100%	100%

Temperature				
	Yes	No	P(yes)	P(no)
Hot	2	2	2/9	2/5
Mild	4	2	4/9	2/5
Cool	3	1	3/9	1/5
Total	9	5	100%	100%

Humidity				
	Yes	No	P(yes)	P(no)
High	3	4	3/9	4/5
Normal	6	1	6/9	1/5
Total	9	5	100%	100%

Wind				
	Yes	No	P(yes)	P(no)
False	6	2	6/9	2/5
True	3	3	3/9	3/5
Total	9	5	100%	100%

Play		P(Yes)/P(No)
Yes	9	9/14
No	5	5/14
Total	14	100%

Τέλος, μας μένει το έργο του υπολογισμού του $P(y)$ και του $P(x_i|y)$. Σημειώστε ότι το $P(y)$ ονομάζεται επίσης πιθανότητα τάξης και το $P(x_i|y)$ λέγεται υπό όρους πιθανότητα. Οι διάφοροι αφελείς ταξινομητές Bayes διαφέρουν κυρίως από τις παραδοχές που γίνονται σχετικά με την κατανομή του $P(x_i|y)$. Ας προσπαθήσουμε να εφαρμόσουμε τον παραπάνω τύπο στο σύνολο δεδομένων για τις καιρικές συνθήκες. Για αυτό, θα πρέπει να γίνουν κάποιες πράξεις στο σύνολο των δεδομένων. Πρέπει να υπολογίσουμε το $P(x_i|y_i)$ για κάθε x_i και y_j . Όλοι οι υπολογισμοί φαίνονται στην εικόνα δίπλα. Έτσι, όπως φαίνεται στην εικόνα

υπολογίσαμε το $P(x_i|y_i)$ για κάθε x_i και y_j . Έτσι ο Ταξινομητής είναι έτοιμος. (Kumar, n.d.)

3.3.3 Πλεονεκτήματα και Μειονεκτήματα

Πλεονεκτήματα

- Εύκολο στην υλοποίηση
- Γρήγορος
- Εάν η υπόθεση ανεξαρτησίας ισχύει, τότε λειτουργεί πιο αποτελεσματικά από άλλους αλγόριθμους
- Απαιτεί λιγότερα δεδομένα εκπαίδευσης
- Έχει μεγάλη δυνατότητα κλιμάκωσης
- Μπορεί να κάνει πιθανές προβλέψεις
- Δυνατότητα διαχείρισης συνεχών και διακριτών δεδομένων
- Μπορεί να λειτουργήσει εύκολα με τιμές που λείπουν
- Εύκολη ενημέρωση κατά την άφιξη νέων δεδομένων
- Κατάλληλο για προβλήματα ταξινόμησης κειμένου

Μειονεκτήματα

- Η ισχυρή υπόθεση σχετικά με τα χαρακτηριστικά της ανεξαρτησίας, κάτι που δεν ισχύει στην πραγματική ζωή
- Έλλειψη δεδομένων
- Πιθανότητες απώλειας της ακρίβειας
- Μηδενική συχνότητα. Αυτό σημαίνει ότι εάν η κατηγορία οποιασδήποτε κατηγορικής μεταβλητής δεν εμφανίζεται στο σύνολο των δεδομένων εκπαίδευσης, τότε το μοντέλο δίνει μηδενική πιθανότητα στην εν λόγω κατηγορία. (What is Naive Bayes Algorithm?, n.d.)

3.4 Δίκτυα Bayesian

Τα δίκτυα Bayesian είναι ένα τύπος πιθανοτικού γραφικού μοντέλου που χρησιμοποιεί συμπεράσματα Bayesian, για υπολογισμό πιθανοτήτων. Τα δίκτυα Bayesian, στοχεύουν στη μοντελοποίηση των εξαρτήσεων υπό όρους, και επομένως στην αιτιώδη συνάφεια, αντιπροσωπεύοντας τις εξαρτήσεις υπό όρους των κορυφών σε ένα κατευθυνόμενο γράφημα. Μέσω αυτών των σχέσεων, μπορεί κάποιος να βρει τις τυχαίες μεταβλητές στο γράφημα μέσω της χρήσης παραγόντων.

4.4.1 Πιθανότητες

Πριν περιγράψουμε την λειτουργία του δικτύου Bayesian, είναι χρήσιμο να δούμε κάποια πράγματα από την θεωρία πιθανοτήτων.

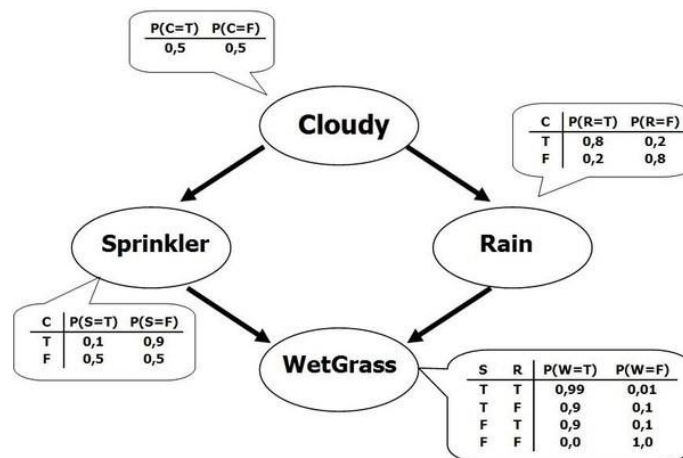
Αρχικά, πρέπει να θυμόμαστε ότι η κοινή κατανομή των πιθανοτήτων των τυχαίων μεταβλητών $A_0, A_1, \dots, A_n$, οι οποίες σημειώνονται ως $P(A_0, A_1, \dots, A_n)$, ισούται με $P(A_1 | A_2, \dots, A_n) * P(A_2 | A_3, \dots, A_n) * \dots * P(A_n)$ με βάση τον κανόνα της αλυσίδας των πιθανοτήτων. Μπορούμε να θεωρήσουμε ότι πρόκειται για μια τεκμηριωμένη αναπαράσταση της κατανομής, καθώς είναι προϊόν N παραγόντων, οι οποίες είναι τοπικές πιθανότητες.

$$P\left(\bigcap_{k=1}^n A_k\right) = \prod_{k=1}^n P(A_k \mid \bigcap_{j=1}^{k-1} A_j)$$

Στην συνέχεια, πρέπει να θυμηθούμε ότι υπό όρους ανεξαρτησία μεταξύ δύο τυχαίων μεταβλητών, των A και B , και με δεδομένη μιας άλλης τυχαία μεταβλητή C , ισοδυναμεί με ικανοποίηση της ακόλουθης ιδιότητας: $P(A, B | C) = P(A | C) * P(B | C)$. Με άλλα λόγια, εφόσον η τιμή του C είναι γνωστή τότε τα A και B είναι ανεξάρτητα. Ένας άλλος τρόπος για να το δηλώσουμε αυτό, τον οποίο θα χρησιμοποιήσουμε αργότερα, είναι το $P(A | B, C) = P(A | C)$.

4.4.2 Περιγραφή του δικτύου

Χρησιμοποιώντας τις σχέσεις που προσδιορίζονται από το Bayesian δίκτυο, μπορούμε να αποκτήσουμε μια τεκμηριωμένη αναπαράσταση της κοινής κατανομής πιθανότητας, εκμεταλλευόμενοι την υπό όρους ανεξαρτησία.



Εικόνα 17: Παράδειγμα

Ένα Bayesian δίκτυο, είναι ένα κατευθυνόμενο κυκλικό γράφημα, στο οποίο κάθε κορυφή αντιστοιχεί σε μια εξαρτώμενη εξάρτηση υπό όρους και κάθε κόμβος αντιστοιχεί σε μια μοναδική τυχαία μεταβλητή. Επίσημα, εάν υπάρχει ένα άκρο (A, B) στο γράφημα που συνδέει τις τυχαίες μεταβλητές A και B, αυτό σημαίνει ότι το $P(B|A)$ είναι ένας παράγοντας στην κοινή κατανομή πιθανοτήτων και επομένως πρέπει να γνωρίζουμε το $P(B|A)$ για όλες τις τιμές B και A προκειμένου να βγάλουμε συμπέρασμα.

Στο παραπάνω παράδειγμα, καθώς το Rain έχει ένα άκρο που πηγαίνει στο WetGrass, σημαίνει ότι το $P(\text{WetGrass}|\text{Rain})$ θα αποτελέσει παράγοντα, του οποίου οι τιμές πιθανότητας καθορίζονται δίπλα στον κόμβο WetGrass σε έναν πίνακα πιθανοτήτων υπό όρους. Τα δίκτυα Bayesian, ικανοποιούν την τοπική ιδιότητα του Markov, η οποία δηλώνει ότι ένας κόμβος είναι υπό όρους ανεξάρτητος από τους κόμβους, οι οποίοι δεν είναι απόγονοί του, δοθέντων των γονιών του. Στο παραπάνω παράδειγμα, αυτό σημαίνει ότι το $P(\text{Sprinkler}|\text{Cloudy}, \text{Rain}) = P(\text{Sprinkler}|\text{Cloudy})$ καθώς ο κόμβος Sprinkler είναι υπό όρους ανεξάρτητος από τον κόμβο Rain, δοσμένου του κόμβου γονέα Cloudy. Αυτή η ιδιότητα μας επιτρέπει να απλοποιήσουμε την κοινή διανομή, η οποία αποκτήθηκε στο προηγούμενο τμήμα χρησιμοποιώντας τον κανόνα της αλυσίδας, σε μικρότερη μορφή. Μετά την απλοποίηση, η κοινή διανομή για ένα δίκτυο Bayesian είναι ίση με: $P(\text{κόμβος}|\text{γονείς}(\text{κόμβος}))$ για όλους τους κόμβους, με βάση τον παρακάτω τύπο:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | \text{Γονέας}(X_i))$$

Σε μεγαλύτερα δίκτυα, αυτή η ιδιότητα μας επιτρέπει να μειώσουμε σημαντικά την ποσότητα των απαιτούμενων υπολογισμών, καθώς οι περισσότεροι κόμβοι θα έχουν λίγους γονείς σε σχέση με το συνολικό μέγεθος του δικτύου.

4.4.3 Συμπερασμοί

Το συμπέρασμα για ένα δίκτυο Bayesian μπορεί να έχει δύο μορφές. Το πρώτο είναι απλώς η αξιολόγηση της κοινής πιθανότητας μιας συγκεκριμένης εκχώρησης τιμών για κάθε μεταβλητή ή υποσύνολο στο δίκτυο. Για αυτό, έχουμε ήδη μιας παραγοντοποιημένη μορφή κοινής διανομής, οπότε απλά θα το αξιολογήσουμε χρησιμοποιώντας υπό όρους πιθανότητες. Εάν μας ενδιαφέρει μόνο ένα υποσύνολο μεταβλητών, θα πρέπει να μην απαλείψουμε αυτές που δεν μας

ενδιαφέρουν. Σε πολλές περιπτώσεις, αυτό μπορεί να έχει σαν αποτέλεσμα την υποχείλιση, επομένως είναι συνηθισμένο να λαμβάνεται ο λογάριθμος αυτού, ο οποίος ισοδυναμεί με την προσθήκη των μεμονωμένων λογαρίθμων κάθε όρου.

Η δεύτερη και πιο ενδιαφέρουσα εργασία συμπεράσματος, είναι η εύρεση του $P(x|e)$ ή η εύρεση της πιθανότητας μιας ανάθεσης ενός υποσυνόλου των μεταβλητών (x), που έχουν εκχωρηθεί σε άλλες μεταβλητές. Στο παραπάνω παράδειγμα, ένα παράδειγμα αυτού θα μπορούσε να είναι η εύρεση του $P(\text{Sprinkler}, \text{WetGrass} | \text{Cloudy})$ όπου το $\{\text{Sprinkler}, \text{WetGrass}\}$ είναι το x και το $\{\text{Cloudy}\}$ είναι το e . Για να το υπολογίσουμε αυτό, θα χρησιμοποιήσουμε το γεγονός ότι $P(x|e) = P(x, e) / P(e) = \alpha P(x, e)$, όπου το α είναι μια σταθερά κανονικοποίησης που θα υπολογίσουμε στο τέλος, έτσι ώστε $P(x|e) + P(\neg x | e) = 1$. Προκειμένου να υπολογιστεί το $P(x, e)$, πρέπει να περιθωριοποιηθεί η κοινή κατανομή των πιθανοτήτων πάνω στις μεταβλητές που δεν εμφανίζονται στο x ή στο e , τις οποίες θα δηλώσουμε ως Y .

$$P(x|e) = \alpha \sum_{y \in Y} P(x, e, Y)$$

Για το συγκεκριμένο παράδειγμα, μπορούμε να υπολογίσουμε το $P(\text{Sprinkler}, \text{WetGrass} | \text{Cloudy})$ ως εξής:

$$\begin{aligned} &P(\text{Sprinkler}, \text{WetGrass} | \text{Cloudy}) \\ &= \alpha \sum_{\text{Rain}} P(\text{WetGrass} | \text{Sprinkler}, \text{Rain}) P(\text{Sprinkler} | \text{Cloudy}) P(\text{Rain} | \text{Cloudy}) P(\text{Cloudy}) \\ &= \alpha P(\text{WetGrass} | \text{Sprinkler}, \text{Rain}) P(\text{Sprinkler} | \text{Cloudy}) P(\text{Rain} | \text{Cloudy}) P(\text{Cloudy}) \\ &\quad + \alpha P(\text{WetGrass} | \text{Sprinkler}, \neg \text{Rain}) P(\text{Sprinkler} | \text{Cloudy}) P(\neg \text{Rain} | \text{Cloudy}) P(\text{Cloudy}) \end{aligned}$$

Θα μπορούσαμε να υπολογίσουμε το $P(\neg x | e)$ κατά τον ίδιο τρόπο, απλώς με το να βάζαμε την τιμή των μεταβλητών στο x σε Ψευδή αντί για αληθή. Όταν υπολογιστούν και τα δύο $P(x | e)$ and $P(\neg x | e)$, μπορούμε να το λύσουμε για α , το οποίο ισούται με $1 / (P(x | e) + P(\neg x | e))$. (Soni, 2018)

4.4.4 Πλεονεκτήματα και Μειονεκτήματα

Πλεονεκτήματα

- Ικανότητα ενσωμάτωσης τόσο των άμεσων αποδεικτικών στοιχείων όσο και των έμμεσων αποδεικτικών στοιχείων.

- Δυνατότητα παραγωγής αποτελεσμάτων για όλες τις συγκρίσεις εντός συνδεδεμένου δικτύου
- Δυνατότητα άμεσου υπολογισμού της πιθανότητας κάθε φάρμακο να είναι η καλύτερη θεραπεία σε μια ασθένεια
- Δυνατότητα προσαρμογής συσχετίσεων στο πλαίσιο δοκιμών πολλαπλών σκελών
- Δυνατότητα ενσωμάτωσης της μετα-παλινδρόμησης για την αξιολόγηση της ετερογένειας εντός του μοντέλου
- Φαίνεται να παράγει έγκυρα, ακριβή αποτελέσματα για μοτίβα δικτύου αστέρα

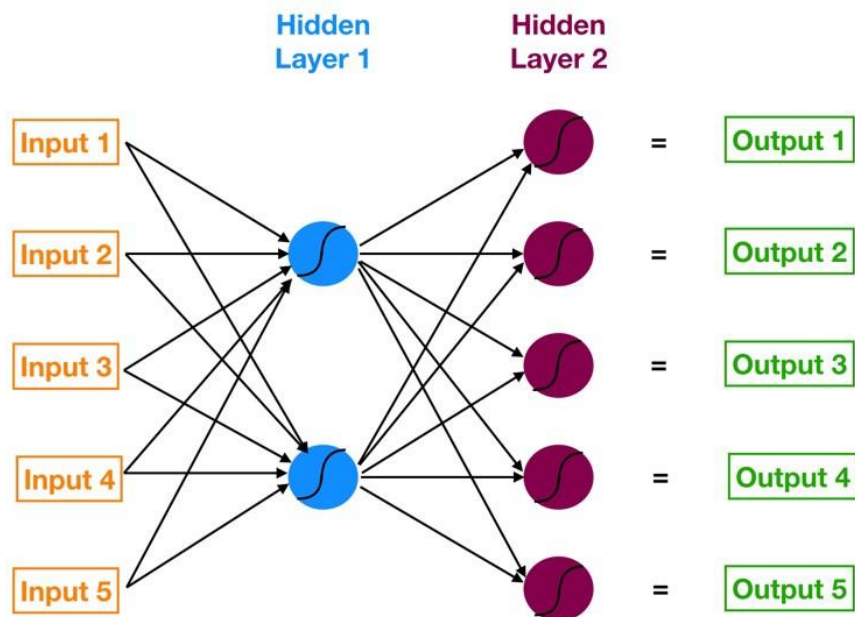
Μειονεκτήματα

- Λιγότερο προσβάσιμες από πολλούς ερευνητές, επειδή οι αναλύσεις εκτελούνται συνήθως χρησιμοποιώντας λογισμικό που δεν είναι γνωστό
- Απαιτεί μεγαλύτερη στατιστική εμπειρογνωμοσύνη από ορισμένες άλλες μεθόδους
- Απαιτείται περαιτέρω έρευνα για την αξιολόγηση των επιδόσεων σε συγκεκριμένα σενάρια
- Ενδέχεται να μην παράγει ακριβή αποτελέσματα για ένα δίκτυο κλειστού βρόχου
- Πιθανώς ευαίσθητα στις προηγούμενες πιθανότητες που έχουν επιλεγεί

3.5 Νευρωνικά Δίκτυα

3.5.1 Βασικά

Τα **Νευρωνικά Δίκτυα**, είναι δίκτυα πολλαπλών στρωμάτων νευρώνων(οι μπλε και οι μωβ κόμβοι στην παρακάτω εικόνα) που χρησιμοποιούνται για ταξινόμηση πραγμάτων, πραγματοποίηση προβλέψεων κ.α. Στην παρακάτω εικόνα παραθέτουμε το διάγραμμα ενός απλού νευρωνικού δικτύου αποτελούμενου από 5 εισόδους, 5 εξόδους και δύο κρυμμένα επίπεδα νευρώνων.



Εικόνα 18: Η δομή ενός απλού νευρωνικού δικτύου

Βλέποντας την εικόνα 17 και ξεκινώντας από την αριστερή πλευρά έχουμε:

1. Το επίπεδο εισόδου του μοντέλου, το οποίο είναι σε πορτοκαλί χρώμα
2. Το πρώτο κρυμμένο στρώμα νευρώνων, το οποίο είναι σε μπλε χρώμα
3. Το δεύτερο κρυμμένο στρώμα νευρώνων, το οποίο είναι σε μωβ χρώμα
4. Το επίπεδο εξόδου ή επίπεδο πρόβλεψης του μοντέλου, το οποίο είναι σε πράσινο χρώμα

Τα βέλη που συνδέουν τους κόμβους δείχνουν πως διασυνδέονται όλοι οι νευρώνες και πως τα δεδομένα μεταφέρονται από το επίπεδο εισόδου μέχρι το επίπεδο εξόδου. Αργότερα στο κεφάλαιο θα υπολογίσουμε αναλυτικά κάθε τιμή εξόδου. Επίσης, θα παρακολουθήσουμε το νευρωνικό δίκτυο να μαθαίνει από τα λάθη του χρησιμοποιώντας μια διαδικασία η οποία είναι γνωστή ως αντίστροφη μετάδοση.

3.5.2 Βασικά συστατικά ενός Νευρωνικού Δικτύου

Σε αυτό το σημείο, κρίνεται καλό να γίνει αναφορά στην εργασία που προσπαθεί να κάνει ένα νευρωνικό δίκτυο. Με απλά λόγια, ένα νευρωνικό δίκτυο, όπως κάθε άλλο μοντέλο, προσπαθεί να κάνει μια καλή πρόβλεψη. Έχουμε ένα σύνολο εισόδων και ένα σύνολο τιμών-στόχων. Με βάση αυτά προσπαθούμε να πάρουμε προβλέψεις που ταιριάζουν με τις δοθείσες τιμές-στόχους όσο το δυνατόν πιο πολύ.



Εικόνα 19: Ένα απλούστατο μοντέλο νευρωνικού δικτύου

Στην εικόνα 18 βλέπουμε μια και μοναδική λογιστική παλινδρόμηση (δίνουμε στο μοντέλο μόνο μια μεταβλητή X) που εκφράζεται μέσω ενός νευρωνικού δικτύου. Για να δούμε πως συνδέονται ξαναγράφουμε την εξίσωση της λογιστικής παλινδρόμησης χρησιμοποιώντας τα χρώματα που χρησιμοποιήσαμε στο νευρωνικό δίκτυο:

$$\text{Sigmoid} (B1 * X + B0) = \text{Predicted Probability}$$

Εικόνα 20: Εξίσωση Λογιστικής παλινδρόμησης

Ας εξετάσουμε το κάθε στοιχείο:

1. Το X (πορτοκαλί χρώμα) είναι η είσοδος, το μοναδικό χαρακτηριστικό που δίνουμε στο μοντέλο για να υπολογίσουμε μια πρόβλεψη.
2. Το $B1$ (γαλάζιο χρώμα) είναι η εκτιμώμενη παράμετρος κλίσης της λογιστικής μας παλινδρόμησης (βάρος). Το $B1$ μας ενημερώνει πόσο αλλάζει το Log_Odds όσο αλλάζει το X . Στην εικόνα 18, παρατηρούμε ότι το $B1$ υπάρχει στην γαλάζια γραμμή, η οποία συνδέει την είσοδο X με τον μπλε νευρώνα στο κρυφό επίπεδο 1.
3. Το $B0$ (μπλε χρώμα) είναι η πόλωση.
4. Ο μπλε νευρώνας περιλαμβάνει επίσης μια σιγμοειδή συνάρτηση ενεργοποίησης (υποδηλώνεται από την καμπύλη γραμμή μέσα στον μπλε κύκλο). Η σιγμοειδής συνάρτηση είναι αυτή που χρησιμοποιούμε για να πάμε από τις αποτιμήσεις μέχρι την πιθανότητα.
5. Τέλος, έχουμε την προβλεπόμενη πιθανότητα εφαρμόζοντας την σιγμοειδή συνάρτηση στην ποσότητα $(B1 * X + B0)$.

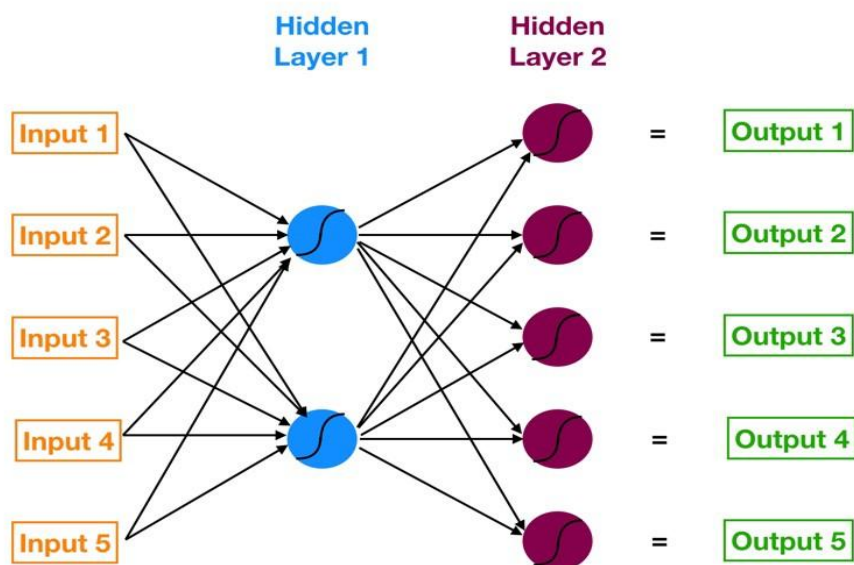
Ας ανακεφαλαιώσουμε. Ένα εξαιρετικά απλό νευρωνικό δίκτυο αποτελείται μόνο από τα ακόλουθα συστατικά μέρη:

- Μια σύνδεση (αν και στην πράξη, θα υπάρχουν γενικά πολλαπλές συνδέσεις, η καθεμία με το δικό της βάρος, που θα πηγαίνει σε δικό της νευρώνα), με ένα βάρος, που μετατρέπει την είσοδο(χρησιμοποιώντας το $B1$) και την δίνει στον νευρώνα.
- Έναν νευρώνα, ο οποίος περιλαμβάνει ένα έναν όρο πόλωσης ($B0$) και μια συνάρτηση ενεργοποίησης (την σιγμοειδή συνάρτηση στην περίπτωση μας).

Αυτά τα δύο αντικείμενα είναι τα βασικά συστατικά στοιχεία του νευρωνικού δικτύου. Πιο σύνθετα νευρωνικά δίκτυα είναι μοντέλα με κρυμμένα επίπεδα και αυτό έχει σαν αποτέλεσμα περισσότερους νευρώνες και περισσότερες συνδέσεις μεταξύ νευρώνων. Αυτό το πιο περίπλοκο δίκτυο συνδέσεων (και βαρών και πολώσεων) είναι αυτό που επιτρέπει στο νευρωνικό «μάθει» τις περίπλοκες σχέσεις που κρύβονται στα δεδομένα.

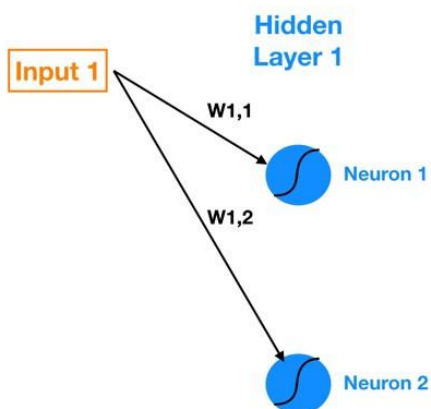
3.5.3 Από την είσοδο στην έξοδο

Τώρα που έχουμε την βασική δομή, ας επανέλθουμε στο ελαφρώς πιο σύνθετο δίκτυο και ας δούμε πως από την είσοδο παράγει την έξοδο. Για να γίνει αυτό ας δούμε την εικόνα 17 την οποία παραθέτουμε ξανά παρακάτω για λόγους ευκολίας:



Το πρώτο κρυμμένο επίπεδο αποτελείται από δύο νευρώνες. Για να συνδέσουμε και τις πέντε εισόδους στους νευρώνες του κρυφού επιπέδου 1, χρειαζόμαστε δέκα συνδέσεις. Η παρακάτω

εικόνα εμφανίζει μόνο τις συνδέσεις μεταξύ εισόδου 1(Input 1) και κρυφού επιπέδου 1 (Hidden Layer 1):



Εικόνα 21: Οι συνδέσεις μεταξύ της εισόδου 1 και του κρυφού επιπέδου 1

Ας δούμε τα βάρη που υπάρχουν στις συνδέσεις. Το $W_{1,1}$ είναι το βάρος που υπάρχει στην σύνδεση μεταξύ της Εισόδου 1 (Input 1) και του Νευρώνα 1(Neuron 1) και $W_{1,2}$ είναι το βάρος που υπάρχει στην σύνδεση μεταξύ της εισόδου 1(Input 1) και νευρώνα 2 (Neuron 2). Η γενική σημειογραφία που θα ακολουθήσουμε είναι $W_{a,b}$ που υποδηλώνει το βάρος στην σύνδεση μεταξύ του Input a και Neuron b. Τώρα, ας υπολογίσουμε τις εξόδους (Outputs) κάθε νευρώνα στο κρυφό επίπεδο 1(γνωστές και ως ενεργοποιήσεις). Αυτό γίνεται χρησιμοποιώντας τους ακόλουθους τύπους (W σημαίνει βάρος και In σημαίνει είσοδος)

$$Z1 = W1*In1 + W2*In2 + W3*In3 + W4*In4 + W5*In5 + Bias_Neuron1$$

$$\text{Ενεργοποίηση Νευρώνα 1} = \text{Σιγμοειδής}(Z1)$$

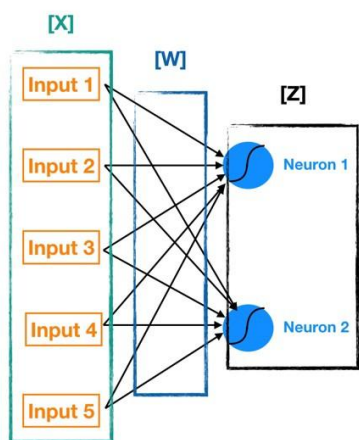
Μπορούμε να χρησιμοποιήσουμε μαθηματικά πινάκων για να συνοψίσουμε αυτόν τον υπολογισμό (πρέπει να θυμόμαστε την σημειογραφία που χρησιμοποιούμε. Παράδειγμα: το $W_{4,2}$ είναι το βάρος που υπάρχει στην σύνδεση μεταξύ της εισόδου 4 και του νευρώνα 2):

$$\begin{bmatrix} W_{1,1} & W_{2,1} & W_{3,1} & W_{4,1} & W_{5,1} \\ W_{1,2} & W_{2,2} & W_{3,2} & W_{4,2} & W_{5,2} \end{bmatrix} \times \begin{bmatrix} X1 \\ X2 \\ X3 \\ X4 \\ X5 \end{bmatrix} + \begin{bmatrix} Bias1 \\ Bias2 \end{bmatrix} = \begin{bmatrix} Z1 \\ Z2 \end{bmatrix}$$

Εικόνα 22: Υπολογισμός ενεργοποίησης

Για οποιοδήποτε επίπεδο ενός νευρωνικού δικτύου όπου το προηγούμενο επίπεδο έχει βάθος έχει m στοιχεία βάθος και το τρέχον επίπεδο έχει βάθος n, αυτό γενικεύεται ως εξής:

$$[W] [X] + [Bias] = [Z]$$



Εικόνα 23: Σχεδιάγραμμα της σχέσης που περιγράψαμε παραπάνω

Όπου $[W]$ είναι ο πίνακας βαρών n επί m (οι συνδέσεις μεταξύ του προηγούμενου και του τρέχοντος επιπέδου), $[X]$ είναι ο πίνακας m επί 1 με τις εισόδους ή τις ενεργοποιήσεις από το προηγούμενο επίπεδο, $[Bias]$ είναι ο πίνακας n επί 1 με τις πολώσεις του νευρώνα, και $[Z]$ είναι ο πίνακας n επί 1 με τις ενδιαμέσες εξόδους. Μόλις αποκτήσουμε το $[Z]$, μπορούμε να εφαρμόσουμε την συνάρτηση ενεργοποίησης (σιγμοειδής στην περίπτωση μας) σε κάθε στοιχείο το $[Z]$ και αυτό μας δίνει τις

εξόδους τους νευρώνα μας (ενεργοποιήσεις) για το τρέχον επίπεδο. Το σχεδιάγραμμα της σχέσης που περιγράψαμε

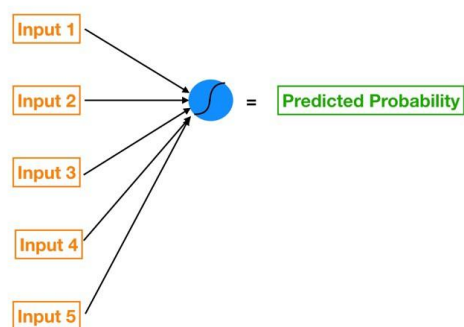
φαίνεται στην εικόνα 22. Με την επαναλαμβανόμενο υπολογισμό του $[Z]$ και εφαρμόζοντας την συνάρτηση ενεργοποίησης σε αυτό για κάθε διαδοχικό επίπεδο, μπορούμε να πάμε από την είσοδο στην έξοδο. Αυτή η διαδικασία είναι γνωστή ως μετάδοση προώθησης. Τώρα που ξέρουμε πως υπολογίζονται οι εξόδοι, είναι καιρός να αρχίσουμε να αξιολογούμε την ποιότητα των εξόδων και να εκπαιδεύσουμε το νευρωνικό μας δίκτυο.

3.5.4 Εκπαίδευση Νευρωνικού Δικτύου

Τώρα που γνωρίζουμε πως υπολογίζονται οι τιμές εξόδου ενός νευρωνικού δικτύου, ας δούμε πως μπορούμε να το εκπαιδεύσουμε. Η διαδικασία εκπαίδευσης ενός νευρωνικού δικτύου, σε υψηλό επίπεδο, είναι σας τις διαδικασίες εκπαίδευσης άλλων μοντέλων της επιστήμης δεδομένων. Αυτές περιλαμβάνουν τον ορισμό μιας συνάρτησης κόστους και χρήση αυτής για την βελτιστοποίηση της κλίσης διαβάθμισης, με σκοπό την ελαχιστοποίηση της. Πρώτα ας σκεφτούμε τι μπορούμε να κάνουμε για να ελαχιστοποιήσουμε την συνάρτηση κόστους. Στην παραδοσιακή γραμμική η λογιστική παλινδρόμηση αναζητούμε συντελεστές βήτα (B_0, B_1, B_2 κ.λπ.), οι οποίοι ελαχιστοποιούν την συνάρτηση κόστους. Για ένα νευρωνικό δίκτυο, ακολουθούμε την ίδια διαδικασία αλλά σε πολύ μεγαλύτερη και πιο περίπλοκη κλίμακα. Στην παραδοσιακή παλινδρόμηση, μπορούμε να αλλάξουμε κάθε βήτα μεμονωμένα χωρίς να επηρεάσουμε τους άλλους συντελεστές. Έτσι, εφαρμόζοντας μικρές απομονωμένες αλλαγές σε κάθε συντελεστή βήτα και μετρώντας την επίδρασή του στην συνάρτηση κόστους, είναι σχετικά

εύκολο να υπολογίσουμε προς ποια κατεύθυνση πρέπει να κινηθούμε για να μειώσουμε και τελικά να ελαχιστοποιήσουμε την συνάρτηση κόστους.

Σε ένα νευρωνικό δίκτυο, η αλλαγή του βάρους οποιασδήποτε σύνδεσης (ή της πόλωσης ενός νευρώνα) έχει αναστρέψιμη επίδραση σε όλους τους άλλους νευρώνες και τις ενεργοποιήσεις



Εικόνα 24: Λογιστική παλινδρόμηση πέντε χαρακτηριστικών μέσω ενός νευρωνικού δικτύου

τους στα επόμενα επίπεδα. Αυτό συμβαίνει επειδή κάθε νευρώνας σε ένα νευρωνικό δίκτυο είναι σαν το δικό του μικρό μοντέλο. Για παράδειγμα, αν θέλαμε μια λογιστική παλινδρόμηση πέντε χαρακτηριστικών, θα μπορούσαμε να την εκφράσουμε μέσω ενός νευρωνικού δικτύου, όπως αυτό στα αριστερά, χρησιμοποιώντας έναν και

μοναδικό νευρώνα. Έτσι, κάθε κρυφό επίπεδο είναι βασικά μια στοίβα μοντέλων (κάθε μεμονωμένος

νευρώνας στο επίπεδο λειτουργεί σαν να είναι ξεχωριστό μοντέλο), των οποίων οι έξοδοι τροφοδοτούνται σε ακόμα περισσότερα μοντέλα πιο κάτω (κάθε διαδοχικό κρυφό επίπεδο νευρωνικού δικτύου έχει ακόμα περισσότερους νευρώνες).

Η συνάρτηση κόστους

Δεδομένης αυτής της πολυπλοκότητας πρέπει να ακολουθήσουμε κάποια βήματα. Πριν όμως από αυτό πρέπει να ορίσουμε σαφώς τον στόχο. Δεδομένων των εκπαιδευτικών εισόδων, και των αποτελεσμάτων (ο στόχος που προσπαθούμε να προβλέψουμε): Θέλουμε να βρούμε το σύνολο των βαρών και των πολώσεων, τα οποία ελαχιστοποιούν την συνάρτηση κόστους. Η συνάρτηση κόστους στην ουσία είναι μια προσέγγιση του πόσο λάθος είναι οι προβλέψεις σε σχέση με το αποτέλεσμα του στόχου. Για την εκπαίδευση του νευρωνικού δικτύου, θα χρησιμοποιήσουμε το μέσω Τετραγωνικό Σφάλμα (MSE - Mean Squared Error) ως συνάρτηση κόστους:

$$MSE = \text{Άθροισμα} [(\text{Πρόβλεψη} - \text{Πραγματική})^2] * (1 / \text{αριθμός παρατηρήσεων})$$

Το MSE ενός μοντέλου μας λέει κατά πόσο λάθος είμαστε. Οι συναρτήσεις κόστους της γραμμικής και λογιστικής παλινδρόμησης λειτουργούν με παρόμοιο τρόπο. Μέχρι τώρα έχουμε μια συνάρτηση κόστους να ελαχιστοποιήσουμε. Για να χρησιμοποιήσουμε την κλίση διαβάθμισης, πρέπει να ξέρουμε την διαβάθμιση της συνάρτησης κόστους μας, δηλαδή το

διάνυσμα το οποίο δείχνει προς την κατεύθυνση του μέρους όπου η κλίση είναι πιο απότομη (θέλουμε να κάνουμε επανειλημμένα βήματα προς την αντίθετη κατεύθυνση της κλίσης για να φτάσουμε στην ελάχιστη τιμή). Σε ένα νευρωνικό δίκτυο έχουμε πολλά μεταβλητά βάρη και πολώσεις, τα οποία είναι διασυνδεδεμένα. Η διαβάθμιση όλων αυτών θα γίνει με την μέθοδο της αντίστροφης διάδοσης την οποία θα αναλύσουμε παρακάτω.

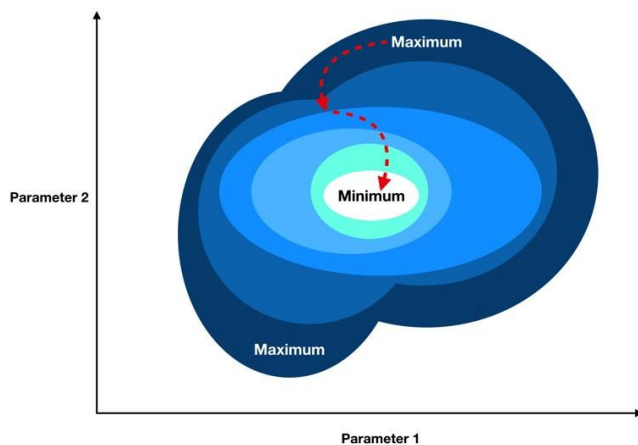
Ας συνοψίσουμε την λειτουργία της της καθόδου βασισμένη στην κλίση:

Η διαβάθμιση μιας συνάρτησης είναι το διάνυσμα, τα στοιχεία του οποίου είναι μερικά παράγωγά του σε σχέση με κάθε παράμετρο. Για παράδειγμα, αν προσπαθήσουμε να ελαχιστοποιήσουμε μια συνάρτηση κόστους $C(B0, B1)$, με μόνο δύο μεταβλητές $B0$ και $B1$, η διαβάθμιση θα ήταν:

$$\text{Διαβάθμιση } C(B0, B1) = [[dC / dB0], [dC / dB1]]$$

Έτσι, κάθε στοιχείο της διαβάθμισης μας λέει πως θα αλλάξει η συνάρτηση κόστους αν εφαρμόσουμε μια μικρή αλλαγή στην συγκεκριμένη παράμετρο, οπότε θα ξέρουμε τι να αλλάξουμε και πόσο. Για να γίνει αυτό θα

Συνοψίζοντας, μπορούμε να κινηθούμε προς το ελάχιστο ακολουθώντας τα εξής βήματα:



Εικόνα 25: Απεικόνιση της κλίσης διαβάθμισης

1. Υπολογισμός της διαβάθμισης της «τρέχουσας τοποθεσίας» (υπολογίστε την διαβάθμιση χρησιμοποιώντας τις τρέχουσες τιμές των παραμέτρων).
2. Τροποποιήστε κάθε παράμετρο κατά ένα ποσό, το οποίο είναι ανάλογο με το στοιχείο διαβάθμισης και έχει την αντίθετη κατεύθυνση από αυτό το στοιχείο. Για

παράδειγμα, αν η μερική παράγωγος

της συνάρτησης κόστους σε σχέση με το $B0$ είναι θετική αλλά πολύ μικρή και η μερική παράγωγος σε σχέση με το $B1$ είναι αρνητική και μεγάλο, τότε θέλουμε να μειώσουμε το $B0$ κατά ένα πολύ μικρό ποσό και να αυξήσουμε το $B1$ κατά ένα μεγάλο ποσό για να μειώσουμε την συνάρτηση κόστους.

3. Υπολογίζουμε ξανά την διαβάθμιση χρησιμοποιώντας τις νέες τιμές των παραμέτρων και επαναλαμβάνουμε τα προηγούμενα βήματα μέχρι να φτάσει στο ελάχιστο.

Αντίστροφη Μετάδοση

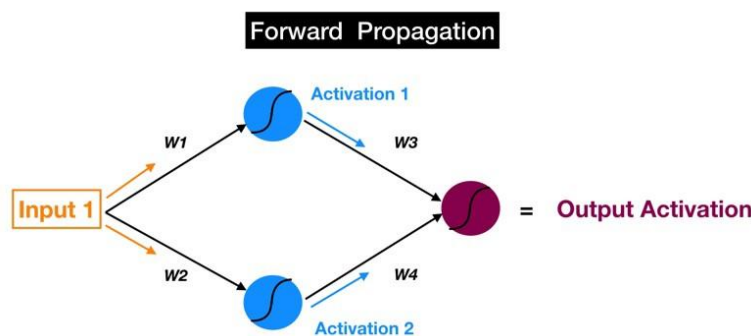
Όπως θυμόμαστε, η μετάδοση προς τα εμπρός είναι η διαδικασία μετακίνησης προς τα εμπρός μέσω του νευρωνικού δικτύου (από εισόδους προς την τελική έξοδο ή πρόβλεψη). Η αντίστροφη μετάδοση είναι το αντίθετο. Αντί να πηγαίνουμε από την είσοδο στην έξοδο, μετακινούμε το λάθος προς τα πίσω μέσω του μοντέλου. Παρακάτω είναι μια εικόνα ενός νευρωνικού δικτύου καθώς πηγαίνει από την είσοδο στην έξοδο. Η διαδικασία μπορεί να συνοψιστεί με τα ακόλουθα βήματα:

- Οι εισοδοί τροφοδοτούνται στο μπλε επίπεδο νευρώνων και τροποποιούνται από τα βάρη, την πόλωση και την σιγμοειδή συνάρτηση σε κάθε νευρώνα, ούτως ώστε να γίνουν οι ενεργοποιήσεις. Για παράδειγμα:

$$\text{Ενεργοποίηση}_1 = \text{Σιγμοειδής}(\text{Πόλωση}_1 + W1 * \text{Είσοδος}_1)$$

- Η ενεργοποίηση 1 και η ενεργοποίηση 2, που προέρχονται από το μπλε στρώμα, τροφοδοτούνται στον μωβ νευρώνα, ο οποίος τα χρησιμοποιεί για να παράγει την τελική ενεργοποίηση εξόδου.

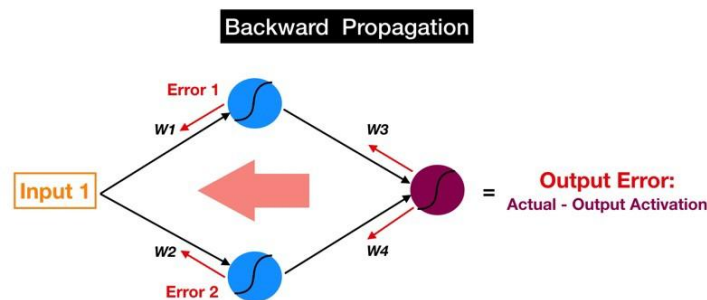
Ο στόχος της προώθησης είναι να υπολογιστούν οι ενεργοποιήσεις σε κάθε νευρώνα για κάθε διαδοχικό επίπεδο μέχρι να φτάσουμε στην έξοδο.



Εικόνα 26: Η μετάδοση προς τα μπροστά σε ένα νευρωνικό δίκτυο

Τώρα ας το αντιστρέψουμε. Εάν ακολουθήσουμε τα κόκκινα βέλη στην παρακάτω εικόνα, θα παρατηρήσουμε ότι αυτή την φορά αρχίζουμε με την έξοδο του μωβ νευρώνα. Αυτή είναι η ενεργοποίηση της εξόδου, την οποία χρησιμοποιούμε για να κάνουμε την πρόβλεψή μας, και η

τελική πηγή σφάλματος στο μοντέλο μας. Στην συνέχεια, μετακινούμε αυτό το σφάλμα προς τα πίσω μέσω του μοντέλου μας, χρησιμοποιώντας τα ίδια βάρη και συνδέσεις που χρησιμοποιήσαμε την προώθηση (έτσι αντί για ενεργοποίηση 1, τώρα έχουμε το Σφάλμα 1, το χαρακτηριστικό σφάλματος στον μπλε νευρώνα). Όπως αναφέραμε καινωρίτερα, ο στόχος της μετάδοσης προς τα μπροστά είναι να υπολογίσουμε τις ενεργοποιήσεις του νευρώνα ανά επίπεδο μέχρι να φτάσουμε στην έξοδο. Τώρα, μπορούμε να δηλώσουμε με παρόμοιο τρόπο τον στόχο της αντίστροφης μετάδοσης: Θέλουμε να υπολογίσουμε στο σφάλμα που δίνεται σε κάθε νευρώνα, ξεκινώντας από το επίπεδο που βρίσκεται πιο κοντά στην έξοδο, μέχρι την έξοδο του μοντέλου.



Εικόνα 27: Η αντίστροφη μετάδοση σε ένα νευρωνικό δίκτυο

Σε αυτό το σημείο είναι καλό να δούμε τον λόγο που μας ενδιαφέρει το λάθος για κάθε νευρώνα. Αν θυμόμαστε, τα δύο δομικά στοιχεία ενός νευρωνικού δικτύου είναι οι συνδέσεις, οι οποίες μεταφέρουν σήματα σε έναν συγκεκριμένο νευρώνα (με βάρη που υπάρχουν σε κάθε σύνδεση) αλλά και τον ίδιο τον νευρώνα (με μια πόλωση).

Το μέγεθος του σφάλματος ενός συγκεκριμένου νευρώνα (σε σχέση με τα σφάλματα όλων των άλλων νευρώνων) είναι ευθέως ανάλογο της επίδρασης της εξόδου του εν λόγω νευρώνα (αλλιώς ενεργοποίηση) στην συνάρτηση κόστους. Έτσι το σφάλμα κάθε νευρώνα είναι ένας διαμεσολαβητής για την μερική παράγωγο της συνάρτησης κόστους σε σχέση με τις εισόδους του νευρώνα. Αυτό είναι λογικό: αν ένας συγκεκριμένος νευρώνας έχει πολύ μεγαλύτερο σφάλμα από όλους τους άλλους, τότε η αλλαγή του βάρους και της πόλωσης του προβληματικού νευρώνα θα έχει μεγαλύτερη επίδραση στο συνολικό σφάλμα του μοντέλου. Και οι μερικές παράγωγοι σε σχέση με κάθε βάρος και πόλωση είναι τα μεμονωμένα στοιχεία που συνθέτουν την διανυσματική διαβάθμιση της συνάρτησης κόστους. Έτσι η αντίστροφη μετάδοση, επιτρέπει τον υπολογισμό σφάλματος, το οποίο μπορεί να αποδοθεί σε κάθε νευρώνα και αυτό με τη σειρά

του μας επιτρέπει να υπολογίσουμε τις μερικές παραγώγους και την διαβάθμιση, έτσι ώστε να μπορούμε να υπολογίσουμε την κλίση διαβάθμισης. (Yiu, 2019)

3.5.5 Πλεονεκτήματα και Μειονεκτήματα Νευρωνικών Δικτύων

Αφότου είδαμε πως λειτουργεί γενικά ένα τεχνητό νευρωνικό δίκτυο καλό θα είναι να δούμε τα πλεονεκτήματα αλλά και τα μειονεκτήματα χρήσης τεχνητών νευρωνικών δικτύων.

Πλεονεκτήματα

- **Αποθήκευση πληροφοριών για ολόκληρο το δίκτυο:** Πληροφορίες όπως στον παραδοσιακό προγραμματισμό αποθηκεύονται σε ολόκληρο το δίκτυο και όχι σε βάση δεδομένων. Η εξαφάνιση ορισμένων στοιχείων πληροφοριών σε ένα μέρος δεν εμποδίζει την λειτουργία του δικτύου.
- **Ικανότητα εργασίας με ελλιπείς γνώσεις:** Μετά την εκπαίδευση του δικτύου, τα δεδομένα μπορούν να παράγουν αποτελέσματα ακόμη και με ελλιπείς πληροφορίες. Η απώλεια απόδοσης δεν εξαρτάται από την σημασία των πληροφοριών που λείπουν.
- **Ανοχή σφαλμάτων:** Η καταστροφή ενός ή περισσότερων νευρώνων δεν εμποδίζει την δημιουργία της εξόδου. Αυτό το χαρακτηριστικό καθιστά το δίκτυο ανεκτικό σε σφάλματα.
- **Κατανεμημένη μνήμη:** Για να μπορεί ένα νευρωνικό δίκτυο να μαθαίνει, είναι απαραίτητο να καθοριστούν τα παραδείγματα και να διδαχθεί το δίκτυο σύμφωνα με το επιθυμητό αποτέλεσμα, παρουσιάζοντας αυτά τα παραδείγματα στο δίκτυο.
- **Σταδιακή φθορά:** Ένα δίκτυο επιβραδύνεται με την πάροδο του χρόνου και υφίσταται σχετική υποβάθμιση. Το δίκτυο δεν διαβρώνεται αμέσως.
- **Δυνατότητα παράλληλης επεξεργασίας:** Τα τεχνητά νευρωνικά δίκτυα έχουν αριθμητική δύναμη που μπορεί να εκτελέσει περισσότερες από μια εργασίες ταυτόχρονα.

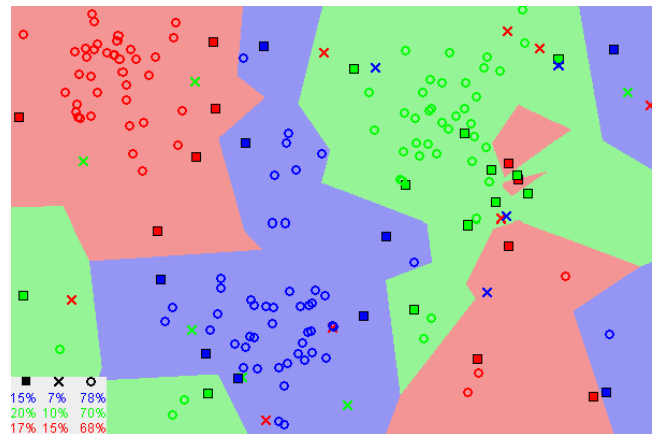
Μειονεκτήματα

- **Έχει εξάρτηση από το υλικό:** Τα τεχνητά νευρωνικά δίκτυα απαιτούν επεξεργαστές με παράλληλη επεξεργαστική ισχύ, ανάλογα με τη δομή τους.
- **Ανεξήγητη Συμπεριφορά δικτύου:** Αυτό είναι το σημαντικότερο πρόβλημα των τεχνητών νευρωνικών δικτύων. Όταν το δίκτυο παράγει μια δοκιμαστική λύση, δεν υπάρχει καμία ένδειξη για το γιατί και το πως. Αυτό μειώνει την εμπιστοσύνη στο δίκτυο.

- **Προσδιορισμός της κατάλληλης δομής δικτύου:** Δεν υπάρχει ειδικός κανόνας για τον καθορισμό της δομής των τεχνητών νευρωνικών δικτύων. Η κατάλληλη δομή του δικτύου επιτυγχάνεται μέσω της εμπειρίας, της δοκιμής και του σφάλματος.
- **Δυσκολία ερμηνείας του προβλήματος στο δίκτυο:** Τα νευρωνικά δίκτυα μπορούν να εργαστούν με αριθμητικές πληροφορίες. Τα προβλήματα πρέπει να μεταφράζονται σε αριθμητικές τιμές πριν εισαχθούν στο δίκτυο. Ο μηχανισμός απεικόνισης που θα καθοριστεί, θα επηρεάσει άμεσα τις επιδόσεις του δικτύου. Αυτό εξαρτάται από την ικανότητα του χρήστη. (Mijwil, 2018)

3.6 Αλγόριθμος KNN (K-Nearest Neighbors)

Ο αλγόριθμος KNN υποθέτει ότι παρόμοια πράγματα υπάρχουν σε κοντινή απόσταση.



Εικόνα 28: Εικόνα που δείχνει πως παρόμοια σημεία δεδομένων υπάρχουν κοντά το ένα στο άλλο

Παρατηρούμε στην παραπάνω εικόνα ότι τις περισσότερες φορές παρόμοια σημεία δεδομένων είναι κοντά το ένα στο άλλο. Ο αλγόριθμος KNN εξαρτάται από αυτή την παραδοχή, ούτως ώστε να είναι χρήσιμος. Ο KNN χρησιμοποιεί την ιδέα της ομοιότητας με τον υπολογισμό της απόστασης μεταξύ δύο σημείων ενός γραφήματος.

3.6.1 Ο Αλγόριθμος

Φόρτωση των δεδομένων

1. Αρχικοποίηση του K στον επιλεγμένο αριθμό γειτόνων
2. Για κάθε παράδειγμα στα δεδομένα:
 - 2.5 Υπολογισμός της απόστασης μεταξύ του παραδείγματος ερωτήματος και του τρέχοντος παραδείγματος από τα δεδομένα

- 2.6 Προσθήκη της απόστασης και του ευρετηρίου του παραδείγματος σε μια ταξινομημένη συλλογή
3. Ταξινόμηση της συλλογής αποστάσεων και δεικτών από τη μικρότερο στη μεγαλύτερη απόσταση
4. Επιλογή των πρώτων K καταχωρήσεων από την συλλογή
5. Λήψη των ετικετών των K επιλεγμένων καταχωρήσεων
6. Εάν έχουμε παλινδρόμηση, επέστρεψε τον μέσο όρο των K ετικετών
7. Εάν έχουμε κατηγοριοποίηση, επέστρεψε την λειτουργία των K ετικετών

3.6.2 Επιλογή του κατάλληλου K

Για την επιλογή του K που είναι κατάλληλο για τα δεδομένα που έχουμε, εκτελούμε τον αλγόριθμο KNN αρκετές φορές με διαφορετικές τιμές του K και επιλέγουμε το K που μειώνει τον αριθμό των σφαλμάτων που συναντούμε, διατηρώντας παράλληλα την ικανότητα του αλγορίθμου να κάνει με ακρίβεια προβλέψεις όταν πρόκειται για δεδομένα που δεν έχει ξαναδεί.

Ακολουθούν κάποια πράγματα που πρέπει να λαμβάνουμε υπόψη:

1. Καθώς μειώνουμε την τιμή του K στο 1, οι προβλέψεις γίνονται λιγότερο σταθερές. Απλά ας σκεφτούμε για ένα λεπτό και να θέσουμε νοερά το $K = 1$ και έχουμε ένα σημείο ερωτήματος περιτριγυρισμένο από αρκετές κόκκινες και μια πράσινη (με βάση το χρωματιστό σχέδιο της εικόνας 28). Λογικά πιστεύουμε, ότι το σημείο ερωτήματος είναι πιθανότερα κόκκινο, αλλά επειδή $K = 1$, ο KNN προβλέπει λανθασμένα ότι το σημείο του ερωτήματος είναι πράσινο.
2. Αντιστρόφως, καθώς αυξάνουμε την αξία του K , οι προβλέψεις μας γίνονται πιο σταθερές λόγω της πλειοψηφικής ψηφοφορίας / μέσου όρου, και επομένως, είναι πιο πιθανό να γίνουν ακριβείς προβλέψεις. Τελικά, αρχίζουμε να γινόμαστε μάρτυρες ενός αυξανόμενου αριθμού σφαλμάτων. Σε αυτό το σημείο γνωρίζουμε ότι έχουμε αυξήσει το K πάρα πολύ.
3. Στις περιπτώσεις όπου έχουμε μια πλειοψηφική ψήφο (π.χ. επιλέγοντας τον τρόπο λειτουργίας σε ένα πρόβλημα κατηγοριοποίησης) μεταξύ ετικετών, συνήθως κάνουμε τον K μονό για να έχουμε έναν διαχωριστή. (Harrison, 2018)

3.6.3 Πλεονεκτήματα και Μειονεκτήματα

Πλεονεκτήματα

- Ο αλγόριθμος είναι απλός και εύκολος στην υλοποίηση
- Δεν υπάρχει ανάγκη κατασκευής μοντέλου, συντονισμού αρκετών παραμέτρων ή να κάνετε πρόσθετες υποθέσεις
- Ο αλγόριθμος είναι ευέλικτος. Μπορεί να χρησιμοποιηθεί για ταξινόμηση, παλινδρόμηση και αναζήτηση

Μειονεκτήματα

- Ο αλγόριθμος είναι σημαντικά πιο αργός καθώς αυξάνεται ο αριθμός παραδειγμάτων ή των προγνωστικών/ανεξάρτητων μεταβλητών

Κεφάλαιο 4: Εξόρυξη γνώσης από Μουσικά Δεδομένα

4.1 Αναπαράσταση Μουσικής Πληροφορίας

4.1.1 Εισαγωγικά Στοιχεία

Κατά τις τελευταίες δεκαετίες η ψηφιακή μουσική έχει γίνει πιο προσβάσιμη και οι μουσικές πηγές έχουν γίνει πολύ παραγωγικές. Οι διακομιστές Web για τη μουσική εξερεύνηση και υποβολή προτάσεων περιέχουν τεράστια repositories μουσικών αντικειμένων. Επομένως, απαιτείται έξυπνη αυτοματοποίηση για την παραγωγή καλών προτάσεων και την αποτελεσματική αναζήτηση μουσικής. Δύο από τις πιο χρήσιμες διεπαφές για ένα χρήστη για τη λήψη προτάσεων μουσικής είναι το ερώτημα ανά ετικέτα (QbT) και το ερώτημα ανά παράδειγμα (QbE). Στο ερώτημα ανά ετικέτα, το σύστημα ταξινομεί τα μουσικά αντικείμενα ανάλογα με τη συνάφειά τους με μια λέξη ετικέτας, περιγράφοντας κάποια σημασιολογική σημασία της επιθυμητής μουσικής (συναισθηματικό περιεχόμενο, συγκεκριμένα όργανα, μουσικό στυλ κ.λπ.). Σε ένα ερώτημα ανα παράδειγμα, το σύστημα ταξινομεί τα μουσικά αντικείμενα ανάλογα με τη συνάφεια ή την ομοιότητα με ένα δεδομένο μουσικό παράδειγμα (ένα τραγούδι που αρέσει ήδη στο χρήστη). Αυτό μπορεί να γίνει με τη μορφή ηλεκτρονικής ραδιοπρόσβασης, δημιουργώντας αυτόματα μια λίστα αναπαραγωγής για το χρήστη και τελικά με ένα περιβάλλον εργασίας που επιτρέπει στο χρήστη να ανεβάζει μουσικά αποσπάσματα άγνωστα στο σύστημα και να βρίσκει παρόμοια μουσική. Και για τις δύο αυτές διεπαφές αναζήτησης απαιτείται σχόλια ή ευρετηριασμός των τραγουδιών στο repository. Προϋπάρχοντα μετα-δεδομένα της μουσικής (π.χ. τίτλος, καλλιτέχνης, στίχοι, είδος, όργανα) είναι μια πηγή τέτοιων σχόλιων και μπορεί να βοηθήσει στην ανάκτηση των επιθυμητών αντικειμένων. Τέτοιες πληροφορίες μπορούν να δοθούν με τα αρχεία των μουσικών κομματιών καθώς προστίθενται στο χώρο αποθήκευσης (τίτλος, διάρκεια κομματιού, καλλιτέχνης κ.λπ.), που συλλέγονται από ειδικούς (όπως έγινε από το Music Genome Project¹, όπου οι ειδικοί της μουσικής προσλήφθηκαν για να ακούσουν τα τραγούδια και να τα σχολιάσουν χειροκίνητα με σχετικές ετικέτες) ή που συγκεντρώθηκαν από χρήστες της υπηρεσίας Web (Last.FM²).

Ενώ η εμπειρική μέθοδος για τη συλλογή μετα-δεδομένων είναι εντατική και δαπανηρή, η μέθοδος "χρήστης" είναι λιγότερο αξιόπιστη και επιρρεπής σε ασυνεπείς περιγραφές. Μια άλλη πηγή χρήσιμων γνώσεων είναι τα προηγούμενα αρχεία προτιμήσεων χρηστών, είτε συγκεκριμένων χρηστών, για λόγους εξατομίκευσης, είτε πλήθους χρηστών, για γενικές

προτάσεις. Μια τέτοια προσέγγιση ονομάζεται συνεργατικό φιλτράρισμα και αξιοποιεί τη προ πολλών χρηστών. Για παράδειγμα, αν πολλοί χρήστες όπως ο καλλιτέχνης Α και ο καλλιτέχνης Β, και ένας νέος χρήστης θέλουν να ακούν τον καλλιτέχνη Α, το σύστημα θα προτείνει τον καλλιτέχνη Β σε αυτόν το χρήστη. Η προσέγγιση συνεργατικού φιλτραρίσματος εφαρμόζεται μόνο όταν υπάρχει μεγάλο ιστορικό χρήσης (αναπαραγωγή) από πολλούς χρήστες. Ένα σύστημα προτάσεων που βασίζεται αποκλειστικά σε αυτή την προσέγγιση δεν θα προτείνει ποτέ τραγούδια από νέους, άγνωστους καλλιτέχνες, παρόλο που είναι δυνητικά κατάλληλα για ορισμένους χρήστες.

Καθώς η διαθεσιμότητα των χρήσιμων μεταδεδομένων και των δεδομένων προτιμήσεων χρηστών είναι περιορισμένη, τα μεγάλης κλίμακας μουσικά repositories πρέπει να βασίζονται σε συστήματα που βασίζονται στο περιεχόμενο για την εκτέλεση αποδοτικών αυτόματων προτάσεων μουσικής. Τέτοια συστήματα θα πρέπει να είναι "μουσικά ευφυή", πράγμα που σημαίνει ότι θα πρέπει να αναλύουν ψηφιακά ηχητικά σήματα μουσικής και να εξάγουν σημαντικές πληροφορίες. Κατά την τελευταία δεκαετία, πολλή έρευνα αφιερώθηκε στην κατασκευή συστημάτων με βάση το περιεχόμενο για εργασίες ανάκτησης μουσικών πληροφοριών (MIR), όπως ταξινόμηση μουσικής (σε καλλιτέχνες, είδος κ.λπ.), σημασιολογική σημείωση (αυτόματη σήμανση) και ανάκτηση (QbT) και ομοιότητα μουσικής για πρόταση τραγουδιού-προς-τραγούδι (QbE). Η εστίαση ήταν κυρίως σε αλγόριθμους μηχανικής μάθησης που χρησιμοποιούν βασικά χαρακτηριστικά ήχου για την εκτέλεση της εργασίας. Σε αυτή τη δουλειά χρησιμοποιούνται απλά συστήματα ανάκτησης και επικεντρώνονται στη σύγκριση διαφορετικών χαρακτηριστικών ήχου και αναπαραστάσεων. Πριν γίνει η εξέταση νέων χαρακτηριστικών ήχου χαμηλού επιπέδου, προσπαθούμε να αξιοποιήσουμε στο έπακρο τα παραδοσιακά χαρακτηριστικά, με βάση κλιμακούμενα φάσματα σύντομων χρονικών πλαισίων. Προσθέτουμε ένα στάδιο κωδικοποίησης αυτών των διανυσμάτων χαρακτηριστικών πλαισίων με ένα προ-υπολογισμένο βιβλίο κώδικα και ένα στάδιο συνένωσης των κωδικοποιημένων πλαισίων (χρονική ενσωμάτωση) για να έχουμε μια συνοπτική απεικόνιση σταθερής διάστασης ενός ολόκληρου τραγουδιού. Η κωδικοποίηση εντοπίζει τα τοπικά μοτίβα και αναπαριστά τα πλαίσια σε υψηλότερο επίπεδο. Το στάδιο της συγκέντρωσης καθιστά την αναπαράσταση ενός ολόκληρου τραγουδιού συμπαγή και εύκολο στην εργασία με (χαμηλή αποθήκευση, γρήγορος υπολογισμός και επικοινωνία) και δημιουργεί μια αναπαράσταση που έχει την ίδια διάσταση για όλα τα τραγούδια, ανεξάρτητα από τη διάρκειά τους. Δείχνουμε πώς μπορεί να είναι χρήσιμη η

ίδια συνοπτική αναπαράσταση τόσο για την ανάκτηση ερωτήματος ανά ετικέτα όσο και για την ανάκτηση ανά ερώτημα.

4.1.2 Αναπαράσταση Τραγουδιών

Εξετάζουμε το σχέδιο κωδικοποίησης για να έχουμε μια συμπαγή αναπαράσταση κάθε τραγουδιού (ή μουσικού κομματιού). Το σχέδιο αποτελείται από τρία στάδια:

- 1. Χαρακτηριστικά σύντομης χρονικής περιόδου:** Κάθε τραγούδι υποβάλλεται σε επεξεργασία με μια σειρά διανυσμάτων χαρακτηριστικών χαμηλού επιπέδου $X \in R^{d \times T}$ (Τ χρονικά πλαίσια. Από κάθε τέτοιο πλαίσιο εξάγεται ένα d-διαστατό διάνυσμα χαρακτηριστικών).
- 2. Κωδικοποίηση:** Κάθε διάνυσμα χαρακτηριστικών $x_t \in R^d$ κωδικοποιείται στην συνέχεια σε ένα διάνυσμα κώδικα $c_t \in R$, χρησιμοποιώντας ένα προ-υπολογισμένο βιβλίο κώδικα $D \in R^{d \times k}$ με κ «βασικά διανύσματα» d διαστάσεων. Παίρνουμε το κωδικοποιημένο τραγούδι $C \in R^{k \times T}$.
- 3. Ομαδοποίηση:** Τα κωδικοποιημένα διανύσματα πλαισίου συγκεντρώνονται μαζί δημιουργώντας ένα μοναδικό διάνυσμα $v \in R^k$.

Η προσέγγιση αυτή είναι επίσης γνωστή ως προσέγγιση του συνόλου χαρακτηριστικών (Bag Of Features BoF): τα χαρακτηριστικά συλλέγονται από διαφορετικές επιδιορθώσεις ενός αντικειμένου (μικρές δισδιάστατες επιδιορθώσεις μιας εικόνας, σύντομα χρονικά πλαίσια ενός τραγουδιού κ.λπ.) για να διαμορφώσουν ένα σύνολο χαρακτηριστικών μεταβλητού μεγέθους που έχουν εντοπιστεί. Το στάδιο της συγκέντρωσης μας επιτρέπει να έχουμε μια ενοποιημένη διάσταση στις αναπαραστάσεις όλων των τραγουδιών, ανεξάρτητα από τη διάρκεια τους. Ένας κοινός τρόπος για να ενωθούν τα διανύσματα των πλαισίων χαμηλού επιπέδου είναι να πάρουμε κάποια στατιστικά στοιχεία από αυτά, συνήθως τον μέσο όρο τους. Για ένα μονότονο, σύντομο τραγούδι, μια τέτοια στατιστική μπορεί να είναι καλός εκπρόσωπος των ιδιοτήτων του τραγουδιού. Ωστόσο, ένα τυπικό τραγούδι είναι επιρρεπές σε αλλαγές στο φασματικό περιεχόμενο, και μια απλή στατιστική συνάρτηση συνένωσης πάνω στα πλαίσια χαμηλών χαρακτηριστικών μπορεί να μην το αντιπροσωπεύει καλά. Για το λόγο αυτό εισήχθη το δεύτερο στάδιο (κωδικοποίηση). Σε ένα κωδικοποιημένο διάνυσμα, κάθε καταχώριση κωδικοποιεί την παρουσία/απουσία/προβολή ενός συγκεκριμένου μοτίβου σε αυτό το πλαίσιο. Το προ-υπολογισμένο βιβλίο περιέχει κώδικες (μοτίβα) που υποτίθεται ότι αντιπροσωπεύουν περίπου

την ποικιλία των εξεχόντων μοτίβων στα τραγούδια. Η χρήση της αραιότητας στην κωδικοποίηση (με ελάχιστα διανύσματα βάσης να είναι ενεργά σε κάθε πλαίσιο), προωθεί την επιλογή κωδίκων που αντιπροσωπεύουν τυπικά ολόκληρα μοτίβα ήχου (που αποτελούνται πιθανώς από πολλές ζώνες συχνότητας). Η ομαδοποίηση αυτών των κωδικοποιημένων διανυσμάτων έχει νόημα: η χρήση της μέσης ομαδοποίησης δίνει μια απεικόνιση ιστογράμματος, δηλώνοντας τη συχνότητα εμφάνισης κάθε ηχητικού προτύπου, ενώ η χρήση της ομαδοποίησης max-abs (μέγιστη απόλυτη τιμή) παρέχει περισσότερες ενδείξεις αναπαράστασης. για κάθε ηχητικό μοτίβο που εμφανίστηκε οποιαδήποτε στιγμή στο τραγούδι και σε ποια δύναμη. Για ορισμένες μεθόδους κωδικοποίησης, είναι σκόπιμο να λαμβάνεται η απόλυτη τιμή και να αντιμετωπίζονται οι αρνητικές τιμές μακριά από το μηδέν ως ισχυρές τιμές. (Yonatan Vaizman, 2013)

4.2 Εξόρυξη Μουσικής Πληροφορίας

4.2.1 Βαθιά Μάθηση για ανάλυση μουσικού περιεχομένου

Η βαθιά μάθηση επηρεάζει τον ερευνητικό τομέα της ανάλυσης μουσικού περιεχομένου και της ανάκτησης μουσικών πληροφοριών (MIR), αλλά οι πρόσφατες εξελίξεις δείχνουν ότι η υψηλή απόδοση αυτών των συστημάτων δεν αντικατοπτρίζει το πόσο καλά έχουν μάθει να επιλύουν υψηλού επιπέδου προβλήματα της ακρόασης μουσικής. Το MIR έχει ως στόχο την παραγωγή συστημάτων που θα διευκολύνουν την εύρεση "μουσικής, ή πληροφοριών σχετικά με τη μουσική". Αυτό έχει πρωταρχική σημασία για την αντιμετώπιση της τεράστιας ποσότητας μουσικών δεδομένων που υπάρχουν και συνεχίζουν να δημιουργούνται. Οι συσκευές ακρόασης που μπορούν να παράγουν με ευελιξία ακριβείς, ουσιαστικές περιγραφές της μουσικής, μπορούν να μειώσουν σημαντικά το κόστος επεξεργασίας των δεδομένων της μουσικής και να διευκολύνουν την ποικιλία των εφαρμογών. Αυτές εκτείνονται από την αναγνώριση μουσικής, την απόδοση συγγραφέα, την σύσταση, την μεταγραφή και τη δημιουργία λίστας αναπαραγωγής μέχρι την εξαγωγή σημασιολογικών περιγραφών όπως είδος και διάθεση στην υπολογιστική μουσική, ακόμα και στην σύνθεση μουσικής.

Πρόσφατες έρευνες στον τομέα της βαθιάς μάθησης καταγράφουν εντυπωσιακά αποτελέσματα για αρκετά προβλήματα συγκριτικής αξιολόγησης. Εκτός από αυτές τις μεγάλες επιτυχίες, οι μέθοδοι βαθιάς μάθησης είναι πολύ ελκυστικές για τρεις άλλους λόγους: υπάρχουν τώρα αποτελεσματικοί και αποδοτικοί αλγόριθμοι εκπαίδευσης για τη βαθιά μάθηση, για να μην

αναφέρουμε τις πλήρως ελεύθερες και ανοικτές υλοποιήσεις σε όλες τις πλατφόρμες, π.χ. Theano. Αυτό έχει σαν αποτέλεσμα την από κοινού βελτιστοποίηση της εκμάθησης και της ταξινόμησης χαρακτηριστικών, επιτρέποντας έτσι σε κάποιον να μην ασχοληθεί με πολλές δυσκολίες που είναι εγγενείς στην επίσημη κωδικοποίηση των γνώσεων των εμπειρογνομόνων σε ένα μηχανήμα. Επίσης, οι πολυεπίπεδες δομές τους φαίνεται να ευνοούν τις ιεραρχικές αναπαραστάσεις των δομών στα δομών δεδομένων. Ωστόσο, ένα αρνητικό σε όλη αυτή την διαδικασία είναι ότι αυτές οι μέθοδοι απαιτούν πολλά δεδομένα προκειμένου να εκτιμηθούν οι παράμετροι. (Corey Kereliuk, 2015)

4.2.2 Deep Belief Networks

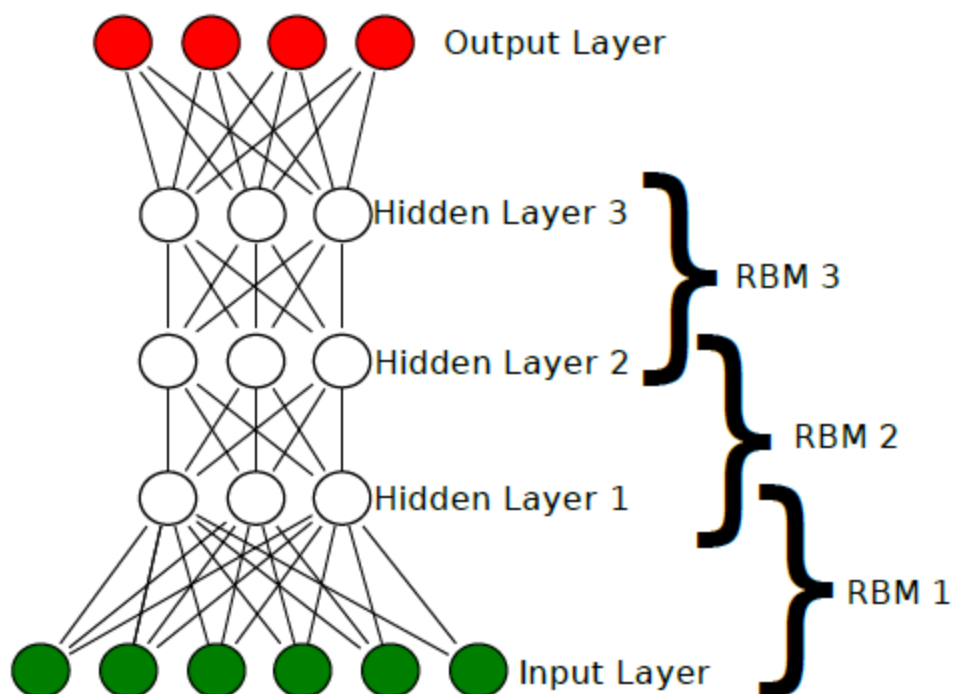
Η αυτόματη ταξινόμηση ειδών μουσικής αποτελείται κυρίως από δύο μέρη: εξαγωγή χαρακτηριστικών και ταξινόμηση μουσικής με ταξινομητές. Στην αρχή, διάφορα διακριτικά χαρακτηριστικά εξάγονται μέσω ανάλυσης του μουσικού σήματος βάσει περιεχομένου και μπορούν να ταξινομηθούν σε βραχυπρόθεσμα και μακροπρόθεσμα χαρακτηριστικά. Τα βραχυπρόθεσμα χαρακτηριστικά αντιπροσωπεύουν κυρίως τα χαρακτηριστικά του μουσικού φάσματος και υπολογίζονται σε κάθε σύντομο χρονικό διάστημα ή πλαίσιο που δηλώνεται ως "παράθυρο ανάλυσης". Τα γενικά χρησιμοποιούμενα βραχυπρόθεσμα χαρακτηριστικά περιλαμβάνουν φασματική φυγοειδή, φασματική ροή, μηδενικές διασταυρώσεις, αισθητικούς συντελεστές Mel-frequency (MFCC) και φασματική αντίθεση βάσει οκτάβας (OSC). Τα μακροπρόθεσμα χαρακτηριστικά, τα οποία συνήθως περιγράφουν την παραλλαγή του φασματικού σχήματος και υπολογίζονται σε μακρά χρονική περίοδο, δηλαδή "παράθυρο υφής", περιλαμβάνουν χαμηλή ενέργεια, ιστογράμματα συντελεστών κυματομορφής (DWCHs), φασματική αντίθεση διαμόρφωσης βάσει οκτάβας (OMSC), μέτρηση φασματικής κορυφής διαμόρφωσης (MSCM), μέτρηση φασματικής ισοτιμίας διαμόρφωσης (MSM) fm) και φασματική ανάλυση διαμόρφωσης φασματικών και αντιληπτικών χαρακτηριστικών. Μετά την εξαγωγή χαρακτηριστικών, μπορούν να χρησιμοποιηθούν πολλοί ταξινομητές ή μέθοδοι ταξινόμησης για την αναγνώριση των ειδών των δοσμένων μουσικών κομματιών.

Σε αυτή την παράγραφο θα δούμε μια προσέγγιση για την ταξινόμηση των μουσικών κομματιών με τη χρήση Deep Belief Networks (DBN). Ένα DBN αποτελείται από δίκτυα κατευθυνόμενων πεποιθήσεων με πολλά επίπεδα κρυφών μεταβλητών. Τα δίκτυα με βαθιά πεποίθηση έχουν προσελκύσει μεγάλη προσοχή και έχουν εφαρμοστεί σε πολλούς τομείς, όπως η ανάκτηση

εγγράφων, η ταξινόμηση ήχου και η ταξινόμηση εικόνων. Οι παράμετροι του μοντέλου καθορίζονται αρχικά μετά την εκπαίδευση του Deep Belief Network από άπληστο αλγόριθμο μάθησης με διανυσματικά χαρακτηριστικά που αποτελούνται από βραχυπρόθεσμα και μακροπρόθεσμα χαρακτηριστικά. Στη συνέχεια, οι παράμετροι ρυθμίζονται με ακρίβεια στο τοπικό βέλτιστο σύμφωνα με τον αλγόριθμο αντίστροφης μετάδοσης.

Ο στόχος της βαθιάς μάθησης είναι η εκμάθηση πιο αφηρημένων αναπαραστάσεων των δεδομένων εισόδου κατά τρόπο επίπεδο χρησιμοποιώντας μάθηση χωρίς επίβλεψη. Αυτές οι διδασκόμενες αναπαραστάσεις μπορούν να χρησιμοποιηθούν ως συνεισφορά για επιτηρούμενη μάθηση σε καθήκοντα όπως η ταξινόμηση και η παλινδρόμηση. Τα τυπικά νευρωνικά δίκτυα είχαν σκοπό να μάθουν τέτοιες βαθιές αναπαραστάσεις. Ωστόσο, τα βαθιά νευρωνικά δίκτυα (δηλαδή τα δίκτυα που έχουν πολλά κρυφά επίπεδα) είναι δύσκολο ή αδύνατο να εκπαιδευτούν χρησιμοποιώντας την κλίση. Το DBN παρακάμπτει αυτό το πρόβλημα εκτελώντας μια άπληστη φάση προ εκπαίδευσης χωρίς επίβλεψη. Έχει αποδειχθεί ότι αυτή η μη επιτηρούμενη προ εκπαίδευση δημιουργεί μια αναπαράσταση από την οποία είναι δυνατή η επιτυχή επιτηρούμενη μάθηση μέσω της "ρύθμισης των λεπτομερειών" των βαρών που προκύπτουν με τη χρήση της μάθησης με κλίση. Με άλλα λόγια, το μη επιτηρούμενο στάδιο ορίζει ότι τα βάρη του δικτύου πρέπει να βρίσκονται πιο κοντά σε μια καλή λύση και να μην παράγονται με τυχαίο τρόπο, αποφεύγοντας έτσι τα τοπικά ελάχιστα όταν χρησιμοποιείται επιτηρούμενη κλίση.

Το Δίκτυο Βαθιών Πεποιθήσεων (Deep Belief Network - DBN) είναι ένα νευρωνικό δίκτυο που κατασκευάζεται από πολλά επίπεδα Περιοριστικών Μηχανημάτων Boltzmann (RBMs). Η σχηματική απεικόνιση απεικονίζεται στην εικόνα 29. Ένα RBM είναι δομημένος ως δύο επίπεδα νευρώνων: ένα ορατό επίπεδο και ένα κρυφό επίπεδο. Κάθε νευρώνας είναι πλήρως συνδεδεμένος με τους νευρώνες του άλλου επιπέδου, αλλά δεν υπάρχει σύνδεση μεταξύ νευρώνων του ίδιου επιπέδου. Ο ρόλος ενός RBM είναι να παρουσιάζει την κατανομή των εισόδων του. Μπορούμε να στοιβάξουμε πολλά RBM το ένα πάνω στο άλλο συνδέοντας το κρυφό επίπεδο ενός RBM με το ορατό επίπεδο του επόμενου RBM.



Εικόνα 29: Σχηματική αναπαράσταση ενός DBN

Εκπαίδευση του Δικτύου

Για να δούμε πως λειτουργεί το συγκεκριμένο είδος δικτύου, θα χρησιμοποιήσουμε το σύνολο δεδομένων Τζανετάκης για αναγνώριση είδους μουσικής. Αυτό το σύνολο δεδομένων αποτελείται από 1000 ηχητικά αποσπάσματα 30 δευτερολέπτων που έχουν αντιστοιχιστεί σε ένα από τα 10 μουσικά είδη. Το σύνολο δεδομένων εξισορροπείται ώστε να έχει 100 αποσπάσματα για κάθε είδος.

Τα δεδομένα θα χρησιμοποιηθούν με τον ακόλουθο τρόπο: 50% για εκπαίδευση, 20% για επικύρωση και 30% για έλεγχο. Ο ήχος διαιρέθηκε σε μικρά πλαίσια των 46,44 ms (1024 δείγματα με ρυθμό δειγματοληψίας 22050 Hz). Για καθένα από αυτά τα πλαίσια, υπολογιστικό ο διακριτός μετασχηματισμός Fourier (DFT). Διατηρήθηκαν μόνο οι απόλυτες τιμές των DFT, και λαμβάνοντας υπόψη τη συμμετρία στο DFT, οι ερευνητές κατέληξαν με εισόδους μεγέθους 513.

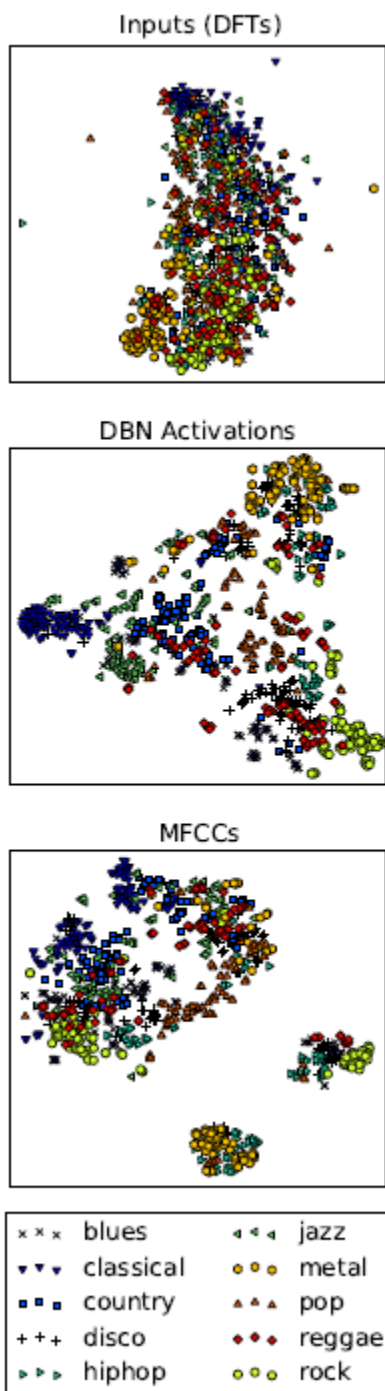
Οι DBN είχαν αρχικά προεκπαιδευτεί με την εκπαίδευση που είχε οριστεί χωρίς επίβλεψη. Στη συνέχεια έγινε η επιτηρούμενη ρύθμιση των λεπτομερειών χρησιμοποιώντας το ίδιο σύνολο εκπαίδευσης και χρησιμοποιώντας το σύνολο επικύρωσης για να πραγματοποιηθεί πρόωρη

διακοπή. Το επιτηρούμενο βήμα χρησιμοποίησε κλίση για να μάθει ένα σταθμισμένο μείγμα ενεργοποιήσεων στο βαθύτερο επίπεδο για να προβλέψει ένα από τα 10 είδη. Τόσο το softmax όσο και το κόστος διασταυρούμενης εντροπίας ελαχιστοποιήθηκαν με συγκρίσιμα

αποτελέσματα.

Δοκιμάστηκαν περίπου 200 διαφορετικούς συνδυασμοί υπερ-παραμέτρων και επιλέχθηκε το μοντέλο με το καλύτερο σφάλμα επικύρωσης σε επίπεδο πλαισίου. Το επιλεγμένο μοντέλο DBN περιγράφεται στον παρακάτω πίνακα:

Αριθμός Κρυφών Επιπέδων	3
Μονάδες ανά επίπεδο	50
Εποπτευόμενο ποσοστό μάθησης	0.1
Μη Εποπτευόμενο ποσοστό μάθησης	0.001
Αριθμός των μη εποπτευόμενων epochs	5
Αριθμός των εποπτευόμενων epochs	474
Συνολικός Χρόνος Εκπαίδευσης (Ώρες)	104
Ακρίβεια Ταξινόμησης	0.737



Εικόνα 30: 2-Διαστάσεις προβολές διαφορετικών απεικονίσεων του ήχου ως προς το είδος τους.

Ο κατηγοριοποιητής που εκπαιδεύτηκε από το τελευταίο επίπεδο του DBN δίνει μια πρόβλεψη του είδους για κάθε καρέ. Στην συνέχεια υπολογίζεται ο μέσος όρος σε όλες τις προβλέψεις για ένα τραγούδι και επιλέγεται η υψηλότερη βαθμολογία ως νικητήρια πρόβλεψη. Αυτό έδωσε μια ακρίβεια πρόβλεψης 73,7%. Αφού εκπαιδευτεί το μοντέλο μπορούν να χρησιμοποιηθούν οι ενεργοποιήσεις των κρυφών

μονάδων DBN ως μια μαθημένη αναπαράσταση του ήχου εισόδου. Έγινε ανάλυση της απόδοσης κάθε επιπέδου του δικτύου ξεχωριστά, καθώς και όλα τα επίπεδα μαζί. Για να δειχθεί τι έμαθε το DBN, στο σχήμα 2 σχεδιάστηκε μια δισδιάστατη προβολή ορισμένων από τις αναπαραστάσεις που χρησιμοποιήθηκαν. (Eck, 2010)

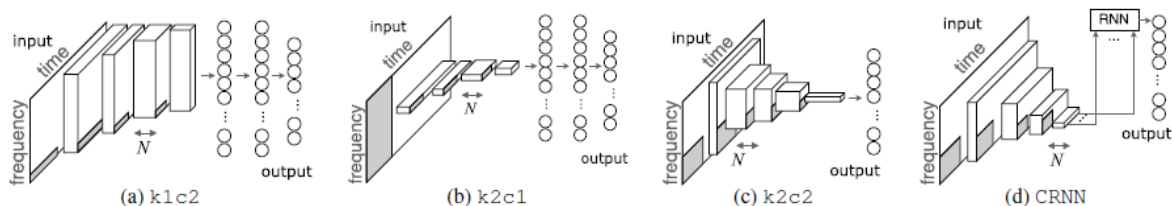
4.2.3 Convolutional Neural Networks – CNN

Τα συνελκτικά νευρωνικά δίκτυα (Convolutional Neural Networks - CNN) έχουν χρησιμοποιηθεί ενεργά για διάφορες εργασίες ταξινόμησης μουσικής, όπως η προσθήκη μουσικών ετικετών, η ταξινόμηση κατά είδος και η πρόβλεψη λανθάνοντος χαρακτηριστικού για τα στοιχεία χρήστη. Τα CNN λαμβάνουν χαρακτηριστικά που βρίσκονται σε διαφορετικά επίπεδα ιεραρχίας και μπορούν να εξαχθούν από συνελκτικούς πυρήνες. Τα ιεραρχικά χαρακτηριστικά μαθαίνουν να επιτυγχάνουν ένα συγκεκριμένο καθήκον κατά τη διάρκεια της επιτηρούμενης εκπαίδευσης. Για παράδειγμα, τα μαθημένα χαρακτηριστικά από ένα CNN που είναι εκπαιδευμένο για την ταξινόμηση ειδών, παρουσιάζουν χαρακτηριστικά χαμηλού επιπέδου έως χαρακτηριστικά υψηλού επιπέδου (π.χ., κρουστικά μοτίβα οργάνων). Πρόσφατα, τα CNN συνδυάστηκαν με επαναλαμβανόμενα νευρωνικά δίκτυα (Recurrent Neural Networks - RNNs), τα οποία συχνά χρησιμοποιούνται για τη μοντελοποίηση διαδοχικών δεδομένων, όπως ηχητικά σήματα ή ακολουθίες λέξεων. Αυτό το υβριδικό μοντέλο ονομάζεται ένα συνελκτικό επαναλαμβανόμενο νευρωνικό δίκτυο (CRNN). Ένα CRNN μπορεί να περιγραφεί ως τροποποιημένο CNN αντικαθιστώντας τα τελευταία επίπεδα συνέλιξης με ένα RNN. Στα CRN, τα CNN και τα RNN διαδραματίζουν τους ρόλους του εξαγωγέα χαρακτηριστικών και του χρονικού περιληπτικού, αντίστοιχα. Η υιοθέτηση ενός RNN για τη συγκέντρωση των χαρακτηριστικών επιτρέπει στα δίκτυα να λαμβάνουν υπόψη την παγκόσμια δομή, ενώ τα τοπικά χαρακτηριστικά εξορύσσονται από τα υπόλοιπα επίπεδα της συνέλιξης. Η δομή αυτή προτάθηκε αρχικά για την ταξινόμηση εγγράφων και αργότερα εφαρμόστηκε στην ταξινόμηση εικόνων και στη μεταγραφή μουσικής. Τα CRN ταιριάζουν καλά στην εργασία προσθήκης ετικετών μουσικής. Τα RNN είναι πιο ευέλικτα κατά την επιλογή του τρόπου σύνοψης των τοπικών χαρακτηριστικών από τα CNN που είναι μάλλον στατικά με τη χρήση του σταθμισμένου μέσου όρου (συγκέντρωση) και της υπό-δειγματοληψίας. Αυτή η ευελιξία μπορεί να είναι χρήσιμη, επειδή ορισμένες από τις ετικέτες (π.χ. ετικέτα διάθεσης) ενδέχεται να επηρεάζονται από την καθολική δομή, ενώ άλλες ετικέτες, όπως τα μουσικά όργανα, μπορούν

να επηρεαστούν από πληροφορίες τοπικών και σύντομων τμημάτων. Σε αυτή την παράγραφο, παρουσιάζουμε τα CRNs για τη σήμανση μουσικής.

Μοντέλα

Συγκρίνουμε το CRNN με τα k1c2, k2c1 και k2c2, τα οποία απεικονίζονται στο σχήμα 1. Τα τρία δίκτυα συνέλιξης ονομάζονται για να καθορίσουν το σχήμα του πυρήνα τους (π.χ. k1 για πυρήνες 1D) και τη διάσταση της συνέλιξης (π.χ. c2 για συναθροίσεις 2D). Οι προδιαγραφές παρουσιάζονται στον πίνακα 1. Για όλα τα δίκτυα, η είσοδος θεωρείται ότι έχει μέγεθος 96x1366 (εύρος ζώνης συχνοτήτων x χρονικό πλαίσιο) και ένα κανάλι. Οι συναρτήσεις Sigmoid χρησιμοποιούνται ως ενεργοποίηση στους κόμβους εξόδου, επειδή η προσθήκη ετικετών μουσικής είναι εργασία ταξινόμησης πολλαπλών ετικετών. Όλα τα επίπεδα είναι εξοπλισμένα με πανομοιότυπες τεχνικές βελτιστοποίησης και λειτουργίες ενεργοποίησης - ομαλοποίηση και λειτουργίας ενεργοποίησης ELU. Πρόκειται για σωστή σύγκριση, δεδομένου ότι οι τεχνικές βελτιστοποίησης βελτιώνουν σημαντικά τις επιδόσεις των δικτύων που έχουν ουσιαστικά την ίδια δομή. Κατ' εξαίρεση, το CRNN έχει ασθενή διακοπή (0.1) μεταξύ επιπέδων συνέλιξης για να αποτραπεί η υπερπροσαρμογή των επιπέδων RNN.



Εικόνα 31: Διαγράμματα τμημάτων των k1c2, k2c1, k2c2 και CRNN. Οι γκρι περιοχές απεικονίζουν τους πυρήνες της συνέλιξης. Ο αριθμός N αναφέρεται στον αριθμό των χαρτών χαρακτηριστικών των επιπέδων συνέλιξης.

- CNN - k1c2

Το k1c2 στην εικόνα 31α δημιουργήθηκε για την ταξινόμηση των ειδών. Το δίκτυο αποτελείται από 4 επίπεδα συνέλιξης που ακολουθούνται από 2 επίπεδα πλήρους σύνδεσης. Ένα μονοδιάστατο επίπεδο συνέλιξης (1 x 4 για όλα, δηλ. συνέλιξη κατά μήκος άξονα χρόνου) και μέγιστα επίπεδα ομαδοποίησης (1 x 4)-(1 x 5)-(1 x 8)-(1 x 8)-(1 x 8) εναλλακτικά. Κάθε στοιχείο της τελευταίας αντιστοίχισης χαρακτηριστικών (η έξοδος του 4-ου επιπέδου υπό-δευματοληψίας) κωδικοποιεί ένα χαρακτηριστικό για κάθε ζώνη. Τροφοδοτούνται σε ένα πλήρως συνδεδεμένο επίπεδο, το οποίο λειτουργεί ως ο κατηγοριοποιητής.

- CNN - k2c1

το k2c1 στην εικόνα 31 β δημιουργήθηκε για την προσθήκη ετικετών μουσικής και την ταξινόμηση του είδους. Το δίκτυο αποτελείται από 5 επίπεδα συνέλιξης που ακολουθούνται από 2 επίπεδα πλήρους σύνδεσης. Το πρώτο επίπεδο συνέλιξης (96×4) μαθαίνει τους πυρήνες 2D που εφαρμόζονται σε ολόκληρη τη ζώνη συχνοτήτων. Στη συνέχεια, εναλλάσσονται τα μονοδιάστατα στρώματα συνέλιξης (1×4 για όλα, δηλαδή η συνέλιξη κατά μήκος του άξονα χρόνου) και τα μέγιστα επίπεδα ομαδοποίησης (1×4) ή (1×5). Τα αποτελέσματα τροφοδοτούνται σε ένα πλήρως συνδεδεμένο επίπεδο. Αυτό το μοντέλο συμπιέζει τις πληροφορίες ολόκληρης της περιοχής συχνοτήτων σε μία ζώνη στο πρώτο επίπεδο συνέλιξης και αυτό συμβάλλει στη σημαντική μείωση της πολυπλοκότητας των υπολογισμών.

- CNN - k2c2

Κατασκευές CNN με συνέλιξη 2D έχουν χρησιμοποιηθεί σε μουσικές ετικέτες και φωνητική/οργανική ταξινόμηση. Το k2c2 αποτελείται από πέντε επίπεδα συνέλιξης 3×3 πυρήνων και μέγιστων επιπέδων ομαδοποίησης ($(2 \times 4)-(2 \times 4)-(2 \times 4)-(3 \times 5)-(4 \times 4)$) όπως απεικονίζεται στην εικόνα 31c. Το δίκτυο μειώνει το μέγεθος των χαρτών χαρακτηριστικών σε 1×1 στο τελικό επίπεδο, όπου κάθε χαρακτηριστικό καλύπτει ολόκληρη την είσοδο και όχι κάθε ζώνη συχνοτήτων όπως στα k1c1 και k2c1. Αυτό το μοντέλο επιτρέπει χρονικές σταθερότητες αλλά και σταθερότητα στην συχνότητα και σε διαφορετική κλίμακα με βαθμιαία υποδείγματα 2D. Επίσης, η χρήση υπό-δειγματοληψίας 2D επιτρέπει στο δίκτυο να είναι πλήρως συνελκτικό, γεγονός που τελικά οδηγεί σε λιγότερες παραμέτρους.

- CRNN

Το CRNN χρησιμοποιεί ένα RNN 2 επιπέδων με ελεγχόμενες επαναλαμβανόμενες μονάδες (GRU) για να συνοψίσει τα προσωρινά μοτίβα στην κορυφή των δισδιάστατων CNN 4 στρωμάτων, όπως φαίνεται εικόνα 31d. Η υπόθεση στην οποία βασίζεται αυτό το μοντέλο είναι ότι το χρονικό μοτίβο μπορεί να αθροιστεί καλύτερα με τα RNN και όχι με τα CNN, ενώ βασίζεται στα CNN στην πλευρά εισόδου για την εξαγωγή τοπικών χαρακτηριστικών. Στο CRNN, τα RNN χρησιμοποιούνται για τη συγκέντρωση των χρονικών μοτίβων αντί, για παράδειγμα, του μέσου όρου των αποτελεσμάτων από τα μικρότερα τμήματα ή τη συγκέντρωση και την υπό-δειγματοληψία. Στην υποδομή του

CNN, τα μεγέθη των επιπέδων συμφόρησης και των επιπέδων μέγιστης συγκέντρωσης είναι 3×3 και (2×2) - (3×3) - (4×4) - (4×4) . Αυτή η υπό-δειγματοληψία έχει ως αποτέλεσμα μέγεθος χάρτη δυνατοτήτων $N \times 1 \times 15$ (αριθμός χαρτών χαρακτηριστικών $\times$ συχνότητα $\times$ χρόνος). Στη συνέχεια τροφοδοτούνται σε ένα RNN 2 επιπέδων, από το οποίο η τελευταία κρυφή κατάσταση συνδέεται με την έξοδο του δικτύου. (Keunwoo Choi, 2016)

4.2.4 Βαθιά Νευρωνικά Δίκτυα (DNN – Deep Neural Networks)

Οι περισσότερες εργασίες ταξινόμησης του MIR χρησιμοποιούν μια κοινή διοχέτευση η οποία αποτελείται από εξαγωγή χαρακτηριστικών ακολουθούμενη από ταξινόμηση. Αυτή η διοχέτευση υποθέτει ότι τα χαρακτηριστικά αποτυπώνουν όλες τις σχετικές πληροφορίες για μια συγκεκριμένη εργασία. Τα χαρακτηριστικά που εξάγονται είναι συχνά "χειροποίητα", μια διαδικασία που απαιτεί σημαντική γνώση του τομέα και μηχανική εφευρετικότητα. Πρόσφατα, η αργή πρόοδος στη βελτίωση της ακρίβειας των καθηκόντων του MIR αποδόθηκε στη χρήση χειροποίητων χαρακτηριστικών και έχουν γίνει αρκετές μελέτες που εξερευνούν την εκμάθηση δυνατοτήτων για μουσικά ακουστικά δεδομένα. Η στοχαστική κλίση (SGD) χρησιμοποιείται για την εκπαίδευση νευρωνικών δικτύων τα τελευταία 25 χρόνια. Αν και έχει πολλά πλεονεκτήματα όπως η ευκολία εφαρμογής και η καλή σύγκλιση, υπάρχουν επίσης αρκετά μειονεκτήματα. Ένα από τα μειονεκτήματα είναι ότι ο αλγόριθμος δεν κλιμακώνεται πολύ καλά σε μεγάλα σύνολα δεδομένων. Ο αλγόριθμος μπορεί να περάσει από εκατοντάδες χιλιάδες επαναλήψεις για να συγκλίνει σε αποδεκτές λύσεις σε μεγάλα σύνολα δεδομένων. Το SGD μπορεί επίσης να προσαρμόσει τα δεδομένα εάν το δίκτυο είναι πολύ βαθύ, ένα αξεπέραστο πρόβλημα στην εκπαίδευση σε βαθιών νευρωνικών δικτύων, μέχρι την πρόσφατη χρήση μη επιτηρούμενης εκπαίδευσης επιπέδου. Πρόσφατα, διαπιστώθηκε ότι η απόδοση των SGD μπορεί να βελτιωθεί με τη χρήση διορθωμένων γραμμικών μονάδων (ReLU). Αυτά τα δίκτυα συγκλίνουν σε παρόμοιες και μερικές φορές καλύτερες λύσεις από τα παραδοσιακά σιγμοειδή δίκτυα.

Τα δίκτυα διόρθωσης όταν συνδυάζονται με το Dropout είναι γνωστό παράγουν καλά αποτελέσματα. Θα ήταν ενδιαφέρον να μπορούμε να συγκρίνουμε τα χαρακτηριστικά που μαθαίνουν τα διορθωτικά δίκτυα με τα χαρακτηριστικά που μαθαίνουν τα σιγμοειδή δίκτυα. Ωστόσο, τα δίκτυα σημασιολογίας είναι δυσκολότερο να βελτιστοποιηθούν με το SGD λόγω της μη γραμμικότητας του στόχου σε σχέση με τις παραμέτρους. Το SGD χρειάζεται επίσης πολύ

μεγάλο αριθμό επαναλήψεων για την εκπαίδευση σιγμοειδών δικτύων, καθιστώντας τον χρόνο εκπαίδευσης απαγορευτικά μεγάλο. Πρόσφατα, ο Μάρτενς πρότεινε μια ισχυρή νέα μέθοδο για την εκπαίδευση νευρωνικών δικτύων, που ονομάζεται βελτιστοποίηση Free Hessian (HF). Το HF είναι μια τεχνική βελτιστοποίησης που έχει αποδειχθεί πολύ αποτελεσματική για την εκπαίδευση βαθιών και επαναλαμβανόμενων νευρωνικών δικτύων. Το HF προσφέρει επίσης άλλα πρακτικά πλεονεκτήματα, όπως σημαντική μείωση του αριθμού των υπερ.-παραμέτρων και σημαντική μείωση του χρόνου εκπαίδευσης.

Διορθωτικές Γραμμικές Μονάδες

Πρόσφατα, υπήρξαν πολλά αποδεικτικά στοιχεία που καταδεικνύουν τα πλεονεκτήματα της χρήσης των ReLUs έναντι των παραδοσιακών σιγμοειδών μονάδων. Τα πλεονεκτήματα αυτά περιλαμβάνουν την καλή σύγκλιση χωρίς την ανάγκη για προ εκπαίδευση και φυσικά αραιά χαρακτηριστικά στα κρυφά επίπεδα. Η συνάρτηση ενεργοποίησης ReLU ορίζεται ως $f(x) = \max(0, x)$. Σε αντίθεση με τις μονάδες sigmoid, οι ReLUs δεν τείνουν στο 1 και η μερική παράγωγος της συνάρτησης ενεργοποίησης σε σχέση με τις παραμέτρους του μοντέλου δεν είναι ποτέ 0, με την προϋπόθεση ότι ο νευρώνας είναι ενεργός. Τα νευρωνικά δίκτυα με ReLUs ως κρυφές μονάδες μπορούν να φτάσουν το ίδιο επίπεδο σφάλματος στο σύνολο δεδομένων εκπαίδευσης πολύ πιο γρήγορα από τα σιγμοειδή δίκτυα. Αυτό είναι επωφελές, καθώς επιτρέπει τον πειραματισμό με πολύ μεγαλύτερα δίκτυα, το οποία δεν θα ήταν δυνατόν να χρησιμοποιήσουν σιγμοειδή δίκτυα που έχουν εκπαιδευτεί με SGD.

Dropout

Αν και τα ReLUs έχουν διάφορες ιδιότητες που τα καθιστούν επιθυμητά ως κρυφές μονάδες για βαθιά νευρωνικά δίκτυα, τείνουν να υπερεξοπλίζουν τα εκπαιδευτικά δεδομένα. Ο υπερεξοπλισμός αποτελεί κοινό πρόβλημα με τα δίκτυα μεγάλης χωρητικότητας και γενικά επιλύεται με τη χρήση διαφόρων τεχνικών διευθέτησης. Το Dropout είναι μια τεχνική διευθέτησης. Το Dropout όταν χρησιμοποιείται με ReLUs μειώνει το πρόβλημα της υπερπροσαρμογής στα εκπαιδευτικά δεδομένα. Το Dropout προσθέτει θόρυβο στο δίκτυο κατά τη διάρκεια της εκπαίδευσης, αποθέτοντας ή αφαιρώντας ένα προκαθορισμένο ποσοστό ενεργοποιήσεων σε κάθε επίπεδο. Οι μονάδες που αποσύρονται επιλέγονται τυχαία. Επομένως, η εκπαίδευση ενός δικτύου με Dropout είναι η ίδια με την εκπαίδευση ενός συνόλου "αραιωμένων" ή υπό-δειγματοληπτικών νευρωνικών δικτύων που μοιράζονται παραμέτρους.

Υπάρχουν δύο τρόποι ερμηνείας του τρόπου με τον οποίο το dropout τροποποιεί την εκπαίδευση. Το πρώτο είναι ότι το Dropout αποφεύγει πολύπλοκες συν προσαρμογές των χαρακτηριστικών στα δεδομένα της εκπαίδευσης. Η δεύτερη ερμηνεία είναι ότι το dropout είναι παρόμοιο με τον μέσο όρο του μοντέλου.

Βελτιστοποίηση Hessian Free

Τα σιγμοειδή δίκτυα είναι δύσκολο να εκπαιδευτούν με SGD, επειδή η επιφάνεια σφάλματος που ορίζεται από τη συνάρτηση απώλειας είναι πολύ περίπλοκη λόγω των μη-γραμμικών σιγμοειδών σε κάθε επίπεδο. Το πρόβλημα αυτό επιδεινώνεται καθώς αυξάνεται το βάθος του δικτύου. Το μεγαλύτερο βάθος οδηγεί επίσης σε ζητήματα όπως η εξαφάνιση της κλίσης λόγω του κορεσμού της σιγμοειδούς συνάρτησης. Έχουν γίνει προσπάθειες να διορθωθεί αυτό χρησιμοποιώντας διαφορετικούς ρυθμούς μάθησης για διαφορετικά επίπεδα, προσαρμοζόμενα χρονοδιαγράμματα για τον ρυθμό μάθησης. Ωστόσο, για αρκετά βαθιά δίκτυα, το SGD αποδεικνύεται ανεπαρκές. Η βελτιστοποίηση Hessian Free (HF) είναι μια τεχνική βελτιστοποίησης που στοχεύει στην ελαχιστοποίηση μιας αντικειμενικής συνάρτησης $f(\theta)$ σε σχέση με τις παραμέτρους θ . Ο αλγόριθμος ενημερώνει διαδοχικά τις παραμέτρους σε κάθε βήμα βελτιστοποιώντας μια τοπική προσέγγιση προς το στόχο. Πιο συγκεκριμένα, η τοπική προσέγγιση προς την αντικειμενική συνάρτηση σε κάποια επανάληψη n ορίζεται ως:

$$M_{\theta_n} = f(\theta_n) + f'(\theta_n)^T \delta n + \delta_n^T B \delta_n / 2$$

Στην εξίσωση 1, B είναι ο πίνακας Gauss-Newton που χρησιμοποιείται αντί για τον Hessian και δ_n είναι οι κατευθύνσεις αναζήτησης για την ενημέρωση των παραμέτρων θ_n . Αν και οι μέθοδοι Hessian Free υπάρχουν στη λογοτεχνία και έχουν μελετηθεί για πολύ καιρό, ήταν κατάφωρα ανεφάρμοστες για εφαρμογές σε προβλήματα εκμάθησης μηχανών μέχρι πρόσφατα. Το HF χρησιμοποιεί τη μέθοδο των Διαβάθμιση Κλίσης (CG) για να βρει μια λύση στην εξίσωση 1 σε κάθε βήμα. Η CG είναι πολύ ευαίσθητη στη μορφή του πίνακα καμπυλότητας και μικρά σύνολα δεδομένων είναι αναποτελεσματικά για εκπαίδευση. Γενικά, χρησιμοποιούνται πολύ μεγάλα σύνολα δεδομένων για κάθε επανάληψη του HF. Εκτός από την καλύτερη απόδοση των CG, η χρήση μεγάλων συνόλων δεδομένων μειώνει σημαντικά τον αριθμό των επαναλήψεων που απαιτούνται για την εκπαίδευση ενός δικτύου. Ένα άλλο πλεονέκτημα της χρήσης HF είναι ότι μειώνει σημαντικά τον αριθμό των υπερ-παραμέτρων για τις οποίες πρέπει να γίνει ρύθμιση. (Siddharth Sigtia)

4.3 Αλγόριθμοι Εκμάθησης Χαρακτηριστικών

Σε αυτό το σημείο, δείχνουμε με ποιους τρόπους η υποκείμενη δομή των δεδομένων να μπορεί να ανακαλυφθεί μέσω αλγορίθμων εκμάθησης χαρακτηριστικών.

Συγκρίνουμε τρεις αλγόριθμους εκμάθησης χαρακτηριστικών για την κωδικοποίηση των προεξεργασμένων δεδομένων σε δισδιάστατα διανύσματα:

- τον K-means clustering
- τον Sparse Coding
- τον Sparse Restricted Boltzmann Machine

4.3.1 K-means clustering

Ο K-means clustering μαθαίνει K κεντροειδή από τα δεδομένα εισόδου και εκχωρεί μία δεδομένη είσοδο σε ένα από τα κεντροειδή K. Αυτό μπορεί να θεωρηθεί ως γραμμική προσέγγιση προς τα διανύσματα εισόδου, $x \approx Ds$, όπου D είναι ένα λεξικό (κεντροειδή) και το s είναι ένα εξαιρετικά αραιό διάνυσμα που όλα τα στοιχεία του είναι μηδενικά εκτός από ένα "1" που αντιστοιχεί στο εκχωρημένο κεντροειδές. Χρησιμοποιούμε το κωδικοποιημένο διάνυσμα, s σαν μαθημένα χαρακτηριστικά. (Juhan Nam, 2012)

4.3.2 Sparse Coding (SC)

Η Sparse coding είναι η μελέτη των αλγορίθμων που στοχεύουν στην εκμάθηση μιας χρήσιμης αραιής απεικόνισης οποιωνδήποτε δεδομένων. Στη συνέχεια κάθε δεδομένο κωδικοποιείται ως αραιός κωδικός:

- Ο αλγόριθμος χρειάζεται μόνο δεδομένα εισόδου για να μάθει τη αραιή αναπαράσταση. Αυτό είναι πολύ χρήσιμο, καθώς μπορείτε να το εφαρμόσετε απευθείας σε οποιοδήποτε είδος δεδομένων, ονομάζεται μη επιτηρούμενη μάθηση
- Θα εντοπίζει αυτόματα την αναπαράσταση χωρίς να χάνει καμία πληροφορία.

Για να γίνει αυτό, οι αραιοί αλγόριθμοι κωδικοποίησης προσπαθούν να ικανοποιήσουν δύο περιορισμούς ταυτόχρονα:

- Για ένα δεδομένο στοιχείο αναπαριστώμενο ως διάνυσμα x , θα προσπαθήσει να μάθει πώς να δημιουργήσει "χρήσιμη" αραιή αναπαράσταση ως διάνυσμα h χρησιμοποιώντας έναν πίνακα κωδικοποιητή E .
- Για κάθε αναπαράσταση ως διάνυσμα h , θα προσπαθήσει να ανακατασκευάσει το στοιχείο x χρησιμοποιώντας έναν πίνακα αποκωδικοποίησης D .
- Τα D και E είναι παράμετροι που μαθαίνονται.

Η μαθηματική αναπαράσταση της γενικής αντικειμενικής συνάρτησης του προβλήματος αυτού μπορεί να βοηθήσει:

$$\min_D \frac{1}{N} \sum_{k=1}^N \min_{h_k} \frac{1}{2} \|x_k - Dh_k\|_2^2 + \lambda \|h_k\|_1$$

$$\text{subject to } \|D_i\|_2^2 \leq C \quad \forall i = 1, \dots, N$$

Όπου:

- N είναι ο αριθμός του στοιχείου στα δεδομένα
- x_k είναι το k διάνυσμα των δεδομένων
- h_k είναι η αραιή αναπαράσταση του x_k
- D είναι η μήτρα αποκωδικοποίησης
- λ είναι ο συντελεστής αραιότητας
- C είναι μια δεδομένη σταθερά (Morgan, 2019)

4.3.3 Sparse Restricted Boltzmann Machine (sparse RBM):

Η Restricted Boltzmann Machine είναι ένα διμερές μη κατευθυνόμενο γραφικό μοντέλο που αποτελείται από ορατούς κόμβους x και κρυφούς κόμβους h . Οι ορατοί κόμβοι αντιπροσωπεύουν διανύσματα εισόδου και οι κρυφοί κόμβοι αντιπροσωπεύουν τα χαρακτηριστικά που μαθαίνονται από την εκπαίδευση του RBM. Η πιθανότητα ένωσης για τους κρυφούς και ορατούς κόμβους ορίζεται στην εξίσωση 2 όταν οι ορατές τιμές είναι μονάδες

Gauss με πραγματικές τιμές και οι κρυφές τιμές είναι δυαδικές μονάδες. Το RBM έχει συμμετρικές συνδέσεις μεταξύ των δύο επιπέδων που υποδεικνύονται από έναν πίνακα βάρους W , αλλά δεν υπάρχουν συνδέσεις εντός κρυφών κόμβων ή ορατών κόμβων. Αυτή η συγκεκριμένη διαμόρφωση διευκολύνει τον υπολογισμό των κατανομών πιθανότητας υπό όρους, όταν παρατηρούνται κόμβοι σε οποιοδήποτε επίπεδο (Εξισώσεις 3 και 4).

$$-\log P(\mathbf{x}, \mathbf{h}) \propto E(\mathbf{x}, \mathbf{h}) = \frac{1}{2\sigma^2} \mathbf{x}^T \mathbf{x} - \frac{1}{\sigma^2} (\mathbf{c}^T \mathbf{x} + \mathbf{b}^T \mathbf{h} + \mathbf{h}^T W \mathbf{x}) \quad (2)$$

$$p(h_j | \mathbf{x}) = \text{sigmoid}(\frac{1}{\sigma^2} (b_j + \mathbf{w}_j^T \mathbf{x})) \quad (3)$$

$$p(x_i | \mathbf{h}) = \mathcal{N}((c_i + \mathbf{w}_i^T \mathbf{h}), \sigma^2), \quad (4)$$

όπου το σ^2 είναι συντελεστής κλιμάκωσης, τα b και c είναι όροι πόλωσης και το W είναι πίνακας βάρους. Οι παράμετροι υπολογίζονται μεγιστοποιώντας την πιθανότητα καταγραφής των ορατών κόμβων. Αυτό εκτελείται με δειγματοληψία τύπου μπλοκ Γκιμπς μεταξύ δύο επιπέδων, ιδιαίτερα, χρησιμοποιώντας έναν κανόνα μάθησης αντίφασης-απόκλισης που περιλαμβάνει μόνο ένα βήμα της δειγματοληψίας Γκιμπς. Κανονικοποιούμε περαιτέρω αυτό το μοντέλο με αραιότητα ενθαρρύνοντας κάθε κρυφή μονάδα να έχει μια προκαθορισμένη αναμενόμενη ενεργοποίηση χρησιμοποιώντας μια ποινή κανονικοποίησης:

$$\lambda \sum_j (\rho - \frac{1}{m} (\sum_{k=1}^m \mathbf{E}[h_j | \mathbf{x}^k]))^2, \quad (5)$$

όπου $\{\mathbf{x}^1, \dots, \mathbf{x}^m\}$ είναι το σύνολο εκπαίδευσης και το ρ καθορίζει την αραιότητα του στόχου των κρυφών ενεργοποιήσεων μονάδας. Παρόμοια με τον K-means clustering και του SC, η εξίσωση 4 μπορεί να ερμηνευτεί ως διανύσματα εισόδου κατά προσέγγιση, $\mathbf{x}$, με γραμμικό συνδυασμό στοιχείων από το λεξικό W . Δηλαδή, $\mathbf{x} \approx W\mathbf{h}$ (παράβλεψη του όρου πόλωσης γ). Το πλεονέκτημα της RBM έναντι των δύο αλγορίθμων είναι ότι η RBM έχει ένα ρητό σχήμα

κωδικοποίησης, $h = \text{sigmoid}(\frac{1}{\sigma^2}(b + W^T x))$. Αυτό επιτρέπει πολύ γρηγορότερο υπολογισμό των μαθημένων χαρακτηριστικών σε σχέση με το SC. (Juhan Nam, 2012)

Βιβλιογραφία

An introduction to Machine Learning. (χ.χ.). Ανάκτηση από GeeksForGeeks:

<https://www.geeksforgeeks.org/introduction-machine-learning/>

Basic Concept of Classification (Data Mining). (χ.χ.). Ανάκτηση από Geeks For Geeks:

<https://www.geeksforgeeks.org/basic-concept-classification-data-mining/>

Corey Kereliuk, M. I. (2015). *Deep Learning and Music Adversaries*.

Data Mining Tutorial: Process, Techniques, Tools, EXAMPLES. (χ.χ.). Ανάκτηση από Guru99:

<https://www.guru99.com/data-mining-tutorial.html>

Eck, P. H. (2010). *LEARNING FEATURES FROM MUSIC AUDIO WITH DEEP BELIEF*.

geeksforgeeks. (χ.χ.). Ανάκτηση από geeksforgeeks: <https://www.geeksforgeeks.org/machine-learning/>

Gupta, P. (2017, May 17). *Decision Trees in Machine Learning*. Ανάκτηση από towards data science: <https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052>

Hall, C. &. (2012). *Music Data Mining*. CRC Press.

Harrison, O. (2018, September 10). *Machine Learning Basics with the K-Nearest Neighbors Algorithm*. Ανάκτηση από towards data science:

<https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761>

Hasan, M. (χ.χ.). *Top 20 Best AI Examples and Machine Learning Applications*. Ανάκτηση από Ubuntu Pit: <https://www.ubuntupit.com/top-20-best-machine-learning-applications-in-real-world/>

Juhan Nam, J. H. (2012). *LEARNING SPARSE FEATURE REPRESENTATIONS FOR MUSIC*.

Keunwoo Choi, G. F. (2016). *CONVOLUTIONAL RECURRENT NEURAL NETWORKS FOR MUSIC CLASSIFICATION*.

Kumar, N. (χ.χ.). *Naive Bayes Classifiers*. Ανάκτηση από Geeks for Geeks:
<https://www.geeksforgeeks.org/naive-bayes-classifiers/>

Li, T. L. (2010). *Music Data Mining: An Introduction*.

Maurya, S. (2019, August 29). *Decision Trees — Introduction (ID3)*. Ανάκτηση από towards data science: <https://towardsdatascience.com/decision-trees-introduction-id3-8447fd5213e9>

Mijwil, M. M. (2018, January 27). *Artificial Neural Networks Advantages and Disadvantages*. Ανάκτηση από Linked In: <https://www.linkedin.com/pulse/artificial-neural-networks-advantages-disadvantages-maad-m-mijwel>

ML | *Introduction to Data in Machine Learning*. (χ.χ.). Ανάκτηση από GeeksForGeeks:
<https://www.geeksforgeeks.org/ml-introduction-data-machine-learning/>

ML | *What is Machine Learning ?* (χ.χ.). Ανάκτηση από GeeksForGeeks:
<https://www.geeksforgeeks.org/ml-machine-learning/>

Morgan. (2019, September 29). *Sparse coding: A simple exploration*. Ανάκτηση από MetaFlow:
<https://blog.metaflow.fr/sparse-coding-a-simple-exploration-152a3c900a7c>

Nicholson, C. (χ.χ.). *A Beginner's Guide to Neural Networks and Deep Learning*. Ανάκτηση από skyminid.

R. A. S. C. Jayasanka, M. D. (2013, November). *Sentiment Analysis for Social Media* .
Ανάκτηση από Reasearch Gate:
https://www.researchgate.net/publication/268817500_Sentiment_Analysis_for_Social_Media

Siddharth Sigtia, S. D. (χ.χ.). *IMPROVED MUSIC FEATURE LEARNING WITH DEEP NEURAL NETWORKS*.

- Soni, D. (2018, June 8). *Introduction to Bayesian Networks*. Ανάκτηση από towards data science: <https://towardsdatascience.com/introduction-to-bayesian-networks-81031eed94e>
- Tao Li, L. L. (2010, October 17). *Music Data Mining: An Introduction*. Ανάκτηση από semanticscholar: <https://pdfs.semanticscholar.org/6ca7/b53455585b484afef120cfea75ebfc84f7d6.pdf>
- Twin, A. (2019, 08 18). *Data Mining*. Ανάκτηση από Investopedia: <https://www.investopedia.com/terms/d/datamining.asp>
- What is Database? What is SQL?* (χ.χ.). Ανάκτηση από Guru99: <https://www.guru99.com/introduction-to-database-sql.html>
- What Is Deep Learning?* (χ.χ.). Ανάκτηση από Deep Learning: <https://www.mathworks.com/discovery/deep-learning.html>
- What is Naive Bayes Algorithm?* (χ.χ.). Ανάκτηση από educba: <https://www.educba.com/naive-bayes-algorithm/>
- Yiu, T. (2019, June 2). *Understanding Neural Networks*. Ανάκτηση από towards data science: <https://towardsdatascience.com/understanding-neural-networks-19020b758230>
- Yonatan Vaizman, B. M. (2013). *Codebook based Audio Feature Representation for*.