



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΟΜΕΣ ΜΑΘΗΜΑΤΙΚΩΝ

Μηχανική Μάθηση και Αναγνώριση Εικόνας
στην Έρευνα της Σωματιδιακής Φυσικής

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

ΧΑΡΑΛΑΜΠΟΥ ΤΖΟΒΑΝΑΚΗ

Επιβλέπων: Πέτρος Στεφανέας
Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, Απρίλιος 2017



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών
Τομές Μαθηματικών

Μηχανική Μάθηση και Αναγνώριση Εικόνας στην Έρευνα της Σωματιδιακής Φυσικής

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

ΧΑΡΑΛΑΜΠΟΥ ΤΖΟΒΑΝΑΚΗ

Επιβλέπων: Πέτρος Στεφανέας
Επίκουρος Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 19η Ιουνίου 2017.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....
Πέτρος Στεφανέας
Επίκουρος Καθηγητής
Ε.Μ.Π.

.....
Παναγιώτης Ψαρράκος
Καθηγητής
Ε.Μ.Π.

.....
Ιωάννης Κολέτσος
Επίκουρος Καθηγητής
Ε.Μ.Π.

Αθήνα, Ιούνιος 2017

(Υπογραφή)

.....

ΧΑΡΑΛΑΜΠΟΣ ΤΖΟΒΑΝΑΚΗΣ

Διπλωματούχος Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών Ε.Μ.Π.

© 2017 – All rights reserved



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών
Τομές Μαθηματικών

Copyright ©–All rights reserved Χαράλαμπος Τζοβανάκης, 2017.

Με επιφύλαξη παντός δικαιώματος.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Ευχαριστίες

Ολόθερμες ευχαριστίες θα ήθελα να εκφράσω στον επιβλέποντα καθηγητή μου κ. Πέτρο Στεφανέα, για την ανάθεση του θέματος, την ενθάρρυνση και τη στήριξή του να εργαστώ στο CERN, την αμέριστη εμπιστοσύνη του και όλη τη βοήθεια για την ολοκλήρωση της διπλωματικής μου εργασίας.

Η διπλωματική αυτή εργασία πραγματοποιήθηκε σχεδόν εξ' ολοκλήρου στο ερευνητικό κέντρο του CERN, στη Γενεύη της Ελβετίας, κατά το έτος 2017 υπό την επίβλεψη των προϊσταμένων μου. Τους ευχαριστώ θερμά για τη βοήθειά τους, την εμπιστοσύνη, την καθοδήγηση και υποστήριξή τους καθ' όλο το διάστημα της εργασίας και παραμονής μου, σε ένα κλίμα φιλικό και ευχάριστο που παράλληλα σε εμπνέει και σου δίνει πρωτοβουλία.

Ευχαριστώ απίστευτα το Ζοζελάκι και τον Γιώργο για τη στήριξη και ανοχή τους όλα αυτά τα χρόνια.

Περίληψη

Το CERN, ο ευρωπαϊκός οργανισμός πυρηνικών ερευνών, δραστηριοποιείται εδώ και χρόνια στη διάδοση των ερευνητικών του αποτελεσμάτων. Για το σκοπό αυτό, αναπτύσσει και χρησιμοποιεί το λογισμικό CDS, ένα πλήρες σύστημα ψηφιακής βιβλιοθήκης για την αποθήκευση, συντήρηση, διαχείριση και διάθεση του έργου του. Το λογισμικό βελτιώνεται και εμπλουτίζεται συνεχώς, ώστε να ανταποκρίνεται στις βιβλιογραφικές ανάγκες των ερευνητών και των επιστημόνων. Ένας από τους στόχους του είναι η δυνατότητα αναγνώριση αντικειμένων μέσω ψηφιακών εικόνων.

Στην παρούσα εργασία εξετάσαμε καταρχήν εκτενώς τις αρχές λειτουργίας και τα χαρακτηριστικά του λογισμικού CDS.

Με βάση αυτά, προτείναμε και υλοποιήσαμε μια σχεδίαση η οποία επιτρέπει την αναγνώριση αντικειμένων. Περιγράψαμε αναλυτικά όλες τις απαραίτητες αλλαγές και προσθήκες για την επίτευξη του σκοπού μας.

Στον πυρήνα της σχεδίασής μας βρίσκεται η βιβλιοθήκη μηχανικής μάθησης Tensorflow που με την βοήθεια του Invenio Framework μπορούμε και πετυχαίνουμε, σε σχεδόν πραγματικό χρόνο, την αναγνώριση αντικειμένων και αποθήκευση αυτών ως μεταδεδομένων στο σύστημα. Με αυτό το τρόπο μέσω της αναζήτησης του CDS μπορούμε να βρούμε φωτογραφίες μέσω του περιεχόμενου τους.

Λέξεις Κλειδιά

ψηφιακή βιβλιοθήκη, καταναεμημένη αναζήτηση, συνεργαζόμενες ψηφιακές βιβλιοθήκες, μηχανική μάθηση, tensorflow, CERN, CDS Invenio

Abstract

CERN, the European Organization for Nuclear Research, has been operating for years in the dissemination of research results. For this purpose, it develops and uses the CDS, a complete library system for the storage, maintenance, management and disposal of all its projects. The software is continuously improved and enriched to correspond and reply to the bibliographic needs of researchers and scientists. One of its objectives is the ability to recognize objects through digital images.

In this thesis, we first extensively examined the principles and the characteristics of the CDS software. Based on this, we proposed and implemented a design that allows object recognition. Then, we described in detail all the necessary changes and additions needed in order to be able to achieve our goal.

Core of our design is the machine learning library Tensorflow that, with the help of the Invenio Framework, achieves, in near real time, to identify objects and to save them as metadata in the system. As a result, through the search of CDS, we can search and find photos depending on their content.

Keywords

digital library, distributed searching, personalization services, machine learning, tensorflow, CERN, CDS Invenio

Περιεχόμενα

Ευχαριστίες	1
Περίληψη	3
Abstract	5
Περιεχόμενα	11
Κατάλογος Πινάκων	13
Κατάλογος Σχημάτων	16
1 Εισαγωγή	19
1.1 Αντικείμενο της εργασίας	19
1.2 Οργάνωση του τόμου	20
2 Οι ψηφιακές βιβλιοθήκες στην έρευνα της σωματιδιακής φυσικής	21
2.1 Σωματιδιακή φυσική ή Φυσική των στοιχειωδών σωματιδίων	21
2.2 Ευρωπαϊκός οργανισμός πυρηνικών ερευνών	22
2.3 Ψηφιακές βιβλιοθήκες	25
2.4 Συνεργαζόμενες ψηφιακές βιβλιοθήκες	25
2.5 Παραδείγματα ψηφιακών βιβλιοθηκών	26

2.5.1	<i>EPrints</i>	26
2.5.2	<i>DSpace</i>	27
2.5.3	<i>Fedora</i>	27
2.5.4	<i>Greenstone</i>	28
3	Ο εξυπηρετητής αρχείων του CERN: CDS	29
3.1	Ιστορικό	29
3.2	Παρούσα κατάσταση	30
3.3	Γενική επισκόπηση	30
3.4	Η τεχνολογία πίσω από το <i>CDS</i>	32
3.4.1	Η τεχνολογία πίσω από το <i>CDS</i>	32
3.4.2	Απόκτηση των μεταδεδομένων	33
3.4.3	Ευρετηρίαση και κατάταξη	34
3.4.4	Διεπαφή και εξατομίκευση	34
3.4.5	Περιγραφή των αυτοτελών λογισμικών μονάδων	35
3.5	Τεχνολογίες σχετικές με το CDS	40
3.5.1	GNU/Linux	40
3.5.2	Python	40
3.5.3	MySQL	41
3.5.4	XML / MARCXML	41
3.5.5	OAI / OAI-PMH	41
4	Τι είναι μηχανική μάθηση	43
4.1	Εισαγωγή	43
4.2	Ορισμός	44
4.3	Θεωρία	44

4.4	Μηχανική μάθηση και στατιστική	45
4.5	Ιστορική αναδρομή	46
4.6	Διαφορά μεταξύ εξόρυξης δεδομένων (data mining), μηχανικής μάθησης (machine learning) και βαθιάς μάθησης deep learning	48
4.6.1	Εξόρυξη δεδομένων (Data Mining)	49
4.6.2	Μηχανική μάθηση (Machine Learning)	49
4.6.3	Βαθιά μάθηση (Deep Learning)	49
4.7	Η μηχανική μάθηση στο σήμερα	50
4.8	Χρησιμότητα μηχανικής μάθησης	50
4.9	Χρήστες μηχανικής μάθησης	51
4.9.1	Χρηματοπιστωτικές υπηρεσίες (Financial services)	51
4.9.2	Κυβέρνηση	51
4.9.3	Ιατρική περίθαλψη	52
4.9.4	Μάρκετινγκ και πωλήσεις	52
4.9.5	Πετρέλαιο και φυσικό αέριο	52
4.9.6	Μεταφορά	52
4.10	Η μηχανική μάθηση στον επιχειρηματικό κλάδο	52
4.11	Άνθρωποι και μηχανές	54
4.12	Τύποι προβλημάτων και εργασιών	54
4.12.1	Επιτηρούμενη μάθηση (supervised learning)	55
4.12.2	Μη επιτηρούμενη μάθηση (unsupervised learning)	55
4.12.3	Ημί-επιτηρούμενη μάθηση	55
4.12.4	Ενισχυτική μάθηση (reinforcement learning)	55
4.13	Μια διαφορετική κατηγοριοποίηση	56
5	Νευρωνικά δίκτυα και εφαρμογές κεφάλαιο	57

5.1	Εισαγωγή στα νευρωνικά δίκτυα	57
5.1.1	Τι είναι ένα νευρωνικό δίκτυο	57
5.1.2	Ιστορική αναδρομή	58
5.1.3	Γιατί να χρησιμοποιήσουμε νευρωνικά δίκτυα	58
5.1.4	Νευρωνικά δίκτυα σε σχέση με τους συμβατικούς υπολογιστές	59
5.1.5	Ανθρώπινοι και τεχνητοί νευρώνες – ποιες είναι οι ομοιότητες	60
5.2	Συνελεκτικά νευρωνικά δίκτυα	60
5.2.1	Βιολογικά-εμπνευσμένο μοντέλο	61
5.2.2	CNN σε 5 βήματα	62
5.2.3	Εκτέλεση	63
5.2.4	Ερμηνεία	64
5.2.5	Συμπέρασμα	64
5.2.6	Εφαρμογές	65
5.2.7	Επικρίσεις και κριτικές	68
6	Βιβλιοθήκες μηχανικής μάθησης στην γλώσσα Python	71
6.1	Tensorflow	71
6.2	Scikit-learn	72
6.3	Theano	74
6.4	Caffe	75
6.5	Lasagne	75
6.6	Blocks	76
6.7	Keras	76
6.8	MXNet	77
6.9	PyTorch	77

7	Αναγνώριση αντικειμένων στο CDS	79
7.1	Το πρόβλημα	79
7.2	Σχεδίαση	79
7.2.1	Μονάδα αναζήτησης	80
7.2.2	Μονάδα καταθέσεων	80
7.2.3	Μονάδα αρχείων	81
7.3	Εκπαίδευση του νευρωνικού δικτύου	83
7.4	Υλοποίηση	84
7.4.1	Διεπαφή ανεβάσματος φωτογραφίας	86
7.4.2	Διεπαφή προγραμματισμού εφαρμογών	87
7.4.3	Διεπαφή της αναζήτησης	91
8	Επίλογος	95
8.1	Συμπεράσματα	95
8.2	Μελλοντικές επεκτάσεις	96
	Βιβλιογραφία	97
	Γλωσσάριο	99

Κατάλογος Πινάκων

7.1 Συλλογές δεδομένων	83
----------------------------------	----

Κατάλογος Σχημάτων

2.1	Ο μεγάλος επιταχυντής αδρονίων. (Φωτογραφία: CERN)	23
2.2	Eprints λογότυπο	26
2.3	DSpace λογότυπο	27
2.4	Fedora λογότυπο	28
2.5	Greenstone λογότυπο	28
5.1	Παράδειγμα τεχνητού νευρωνικού δικτύου	57
6.1	Tensorflow λογότυπο	71
6.2	Scikit-learn λογότυπο	72
6.3	Theano λογότυπο	74
6.4	Keras λογότυπο	76
6.5	MXNet λογότυπο	77
6.6	Pytorch λογότυπο	77
7.1	Ροή εργασίας μονάδας	80
7.2	Όλες οι μονάδες του Invenio Software v3	81
7.3	Καλάθι και αντικείμενα στη μονάδα αρχείων	82
7.4	Παράδειγμα αντικειμένων και κλάσεων	84
7.5	Διεπαφή για το ανέβασμα των φωτογραφιών	86

7.6	Διεπαφή για προβολή ετικετων	87
7.7	Πάντα στο πάρκο	89
7.8	Διεπαφή για προβολή ετικετών	91
7.9	Διεπαφή για προβολή ετικετων	92

Αλγόριθμοι

7.1	Εκπαίδευση	84
7.2	Δημιουργία μονάδας	85
7.3	Σημεία Εισόδου	85
7.4	Διεπαφή Προγραμματισμού Εφαρμογών	87
7.5	Απάντηση ερωτήματος	89
7.6	Κλάση μετάφρασης	90
7.7	Απάντηση ερωτήματος αναζήτησης	92

Κεφάλαιο 1

Εισαγωγή

1.1 Αντικείμενο της εργασίας

Βρισκόμαστε σε μια εποχή έντονης αναζήτησης και έρευνας σε διάφορα επιστημονικά πεδία από πολλούς ακαδημαϊκούς και επιστημονικούς οργανισμούς. Έντονη είναι και η ανάγκη αποθήκευσης των αποτελεσμάτων αυτών των ερευνών και του έργου των διαφόρων οργανισμών καθώς και διατήρηση και διάθεση τους. Στην εποχή μας, οι περιορισμοί της φυσικής αποθήκευσης και διαχείρισης όλου αυτού του υλικού έχουν ξεπεραστεί και έχουν δώσει τη σκυτάλη σε σύγχρονες ψηφιακές λύσεις. Έτσι έχουν γεννηθεί και αναπτύσσονται οι ψηφιακές βιβλιοθήκες, συστήματα λογισμικού δηλαδή, τα οποία ως σκοπό έχουν να προσφέρουν όλες τις υπηρεσίες που προσέφερε μια κλασική βιβλιοθήκη, καθώς και επιπλέον υπηρεσίες που μια κλασική βιβλιοθήκη αδυνατούσε να προσφέρει.

Ένα παράδειγμα επιστημονικού οργανισμού, που παράγει πλήθος σημαντικών επιστημονικών εγγράφων ετησίως, αποτελεί το *CERN*, το οποίο εστιάζει τις έρευνές του στην φυσική υψηλών ενεργειών. Στα πλαίσια των αναγκών αποθήκευσης, συντήρησης, διαχείρισης και διάθεσης των εγγράφων αυτών το *CERN* έχει αναλάβει την ανάπτυξη ενός τέτοιου συστήματος ψηφιακής βιβλιοθήκης, του *CDS Invenio*. Παράλληλα υπάρχουν διάφοροι ακόμη αντίστοιχοι οργανισμοί και ιδρύματα ανά τον κόσμο, που πραγματοποιούν παρόμοια έρευνα και έχουν παρόμοιες απαιτήσεις ψηφιακής οργάνωσης του έργου τους. Στα πλαίσια της συνεργασίας μεταξύ αυτών των οργανισμών γεννιέται η ιδέα μιας από κοινού κεντρικής πρόσβασης στο έργο τους. Στόχος μια τέτοιας συνεργασίας είναι η ευκολία πρόσβασης στην πληροφορία ενώ συγχρόνως διατηρείται η ανεξαρτησία μεταξύ των διαφόρων μελών της.

Για τη συνεχή επέκταση και βελτίωση του παραπάνω λογισμικού, *CDS Invenio*, απασχολούνται στο *CERN* διάφοροι έμπειροι και νέοι επιστήμονες και μηχανικοί. Στις προσπάθειες αυτές συμπεριλαμβάνεται και η παρούσα εργασία, στην οποία αναλύεται το λογισμικό *CDS Invenio* και περιγράφεται μια προτεινόμενη σχεδίαση καθώς και η υλοποίησή της, για την

αναγνώριση αντικειμένων σε φωτογραφίες. Στην επόμενη ενότητα περιγράφεται η οργάνωση του συνολικού κειμένου.

1.2 Οργάνωση του τόμου

Η παρούσα εργασία αποτελείται από επτά κεφάλαια. Σε αυτά μελετώνται και αναλύονται καταρχήν οι υπάρχουσες τεχνολογίες με βάση τις οποίες πραγματοποιήθηκε η εργασία, στη συνέχεια περιγράφεται η σχεδίαση της λύσης και τέλος, γίνεται μια αναφορά στην τρέχουσα αξιοποίησή της, καθώς και στις μελλοντικές πιθανές επεκτάσεις. Ειδικότερα:

Στο πρώτο κεφάλαιο γίνεται μια εισαγωγή του αναγνώστη στο αντικείμενο της εργασίας.

Στο δεύτερο κεφάλαιο παρουσιάζονται κάποιες εισαγωγικές έννοιες σχετικές με τις ψηφιακές βιβλιοθήκες καθώς και κάποια παραδείγματα συστημάτων λογισμικού ψηφιακών βιβλιοθηκών.

Στο τρίτο κεφάλαιο περιγράφεται αναλυτικά το λογισμικό του εξυπηρετητή αρχείων του *CERN*, τα χαρακτηριστικά του και οι σχετικές με αυτόν τεχνολογίες.

Στο τέταρτο και πέμπτο κεφάλαιο παρουσιάζεται η βασική θεωρία της εργασίας ώστε να μπορέσουν να επιλυθούν τα προβλήματα αναγνώρισης αντικειμένων μέσω της ψηφιακής βιβλιοθήκης CDS.

Στο έκτο και έβδομο κεφάλαιο δίνονται οι διάφορες βιβλιοθήκες που βοηθούν στην επίλυση προβλημάτων μηχανικής μάθησης σε Python καθώς και η τελική επίλυση του προβλήματος.

Τέλος στο όγδοο κεφάλαιο παρουσιάζονται τα συμπεράσματα της εργασίας ενώ παράλληλα γίνεται αναφορά στην αξιοποίησή τους στην πράξη και προτείνονται κάποιες μελλοντικές επεκτάσεις.

Κεφάλαιο 2

Οι ψηφιακές βιβλιοθήκες στην έρευνα της σωματιδιακής φυσικής

2.1 Σωματιδιακή φυσική ή Φυσική των στοιχειωδών σωματιδίων

Η Σωματιδιακή φυσική ή Φυσική των στοιχειωδών σωματιδίων είναι ένας κλάδος της Φυσικής που μελετά τα στοιχειώδη σωματίδια της ύλης και την ακτινοβολία τους, καθώς και τις αλληλεπιδράσεις τους. Είναι αλλιώς γνωστή και ως Φυσική Υψηλών Ενεργειών (High Energy Physics HEP) διότι πολλά από τα στοιχειώδη αυτά σωματίδια δεν συναντώνται υπό κανονικές συνθήκες στη φύση αλλά μπορούν να δημιουργηθούν και να παρατηρηθούν μόνο κατά τη διάρκεια ενεργειακών συγκρούσεων με άλλα σωματίδια στο περιβάλλον ενός σωματιδιακού επιταχυντή.

Η ιδέα ότι η ύλη αποτελείται από στοιχειώδη σωματίδια πρωτοεμφανίστηκε τουλάχιστον τον 6ο αιώνα π.Χ. Απασχόλησε αρχαίους Έλληνες φιλοσόφους, Ινδούς φιλοσόφους, επιστήμονες κατά τη διάρκεια του μεσσαίωνα και νεότερους ευρωπαίους επιστήμονες όπως τους *Robert Boyle* και *Isaac Newton*. Τον 20ο αιώνα οι πρόοδοι της πυρηνικής καθώς και χβαντικής φυσικής άνοιξαν νέους ορίζοντες στη σωματιδιακή φυσική και οδήγησαν στην εδραίωση του Καθιερωμένου Μοντέλου τη δεκαετία του 1970.

Η έρευνα πάνω στη σωματιδιακή φυσική σήμερα εστιάζεται στα υποατομικά σωματίδια, όπως ηλεκτρόνια, πρωτόνια, νετρόνια, και, πέραν αυτών, σε πολλά άλλα σωματίδια, παράγωγα των υποατομικών σωματιδίων, σχηματιζόμενα μέσω συγκρούσεων μεγάλων ταχυτήτων (υψηλών ενεργειών) μεταξύ τους. Επεξηγηματικά, στις αρχές της δεκαετίας του '30 οι θεωρητικοί φυσικοί νόμιζαν ότι είχαν ανακαλύψει όλο τον ατομικό κόσμο, τα πρωτόνια, νετρόνια, ηλεκτρόνια καθώς και μερικά ακόμη σωματίδια που είχαν βρεί. Αλλά 30 χρόνια αργότερα, στις

αρχές της δεκαετίας του '60, ανακάλυψαν ένα τεράστιο πλήθος σωματιδίων, περίπου διακόσια σωματίδια, είτε στην κοσμική ακτινοβολία είτε στα πειράματα που έκαναν με τους επιταχυντές της εποχής εκείνης. Προσπάθησαν να τα κατατάξουν σε ομάδες για την καλύτερη εξήγησή τους, συγχρόνως όμως προσπαθούσαν να βρουν μοντέλα που θα εξηγούσαν πως αλληλεπιδρούν μεταξύ τους τα θεμελιώδη δομικά υλικά του σύμπαντος.

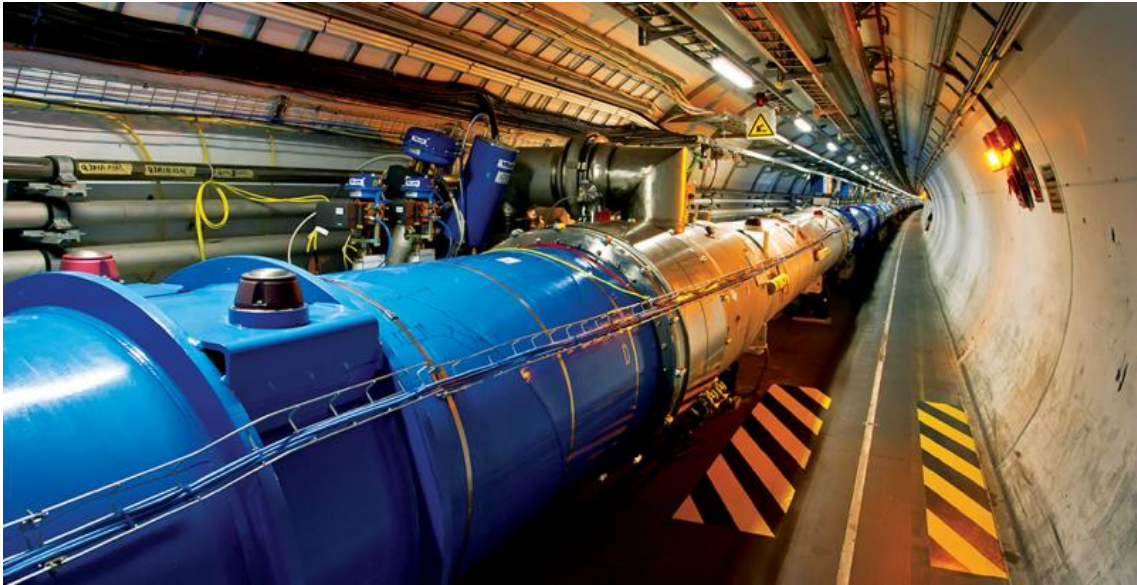
Έτσι, τα είχαν χωρίσει: Πρώτον, στα αδρόνια που είναι όλα τα σωματίδια που αναπτύσσουν τις ισχυρές αλληλεπιδράσεις. Τα αδρόνια χωρίζονται στα μεσόνια (που όλα είναι μποζόνια) και στα βαρυόνια (που όλα είναι φερμιόνια). Δεύτερον, στα λεπτόνια που είναι όλα φερμιόνια και που συμμετέχουν σε ασθενείς και ηλεκτρομαγνητικές αλληλεπιδράσεις κι όχι σε ισχυρές. Τρίτον, στα μποζόνια βαθμίδας, που είναι οι φορείς της ηλεκτρασθενούς δύναμης ($+W$, $-W$, Z) και τέταρτον, στα γκλουόνια (από την ελληνική λέξη γλοιόνια-κολλώδη), που είναι οι φορείς της ισχυρής πυρηνικής δύναμης. Σήμερα, ένας ικανός αριθμός εργαστηρίων ανά τον κόσμο ασχολείται με την εν λόγω έρευνα παράγοντας καθημερινά νέες εργασίες και άρθρα. Μερικά από αυτά είναι το *CERN* (Ευρωπαϊκή Οργανιστική Φορ Νυκλεαρ Ρεσεαρχ) κοντά στη Γενεύη της Ελβετίας, το *DESY* (*Deutsches Elektronen Synchrotron*) κοντά στο Αμβούργο της Γερμανίας, το *Fermilab* (*Fermi National Accelerator Laboratory*) κοντά στο Σικάγο των Η.Π.Α. και το *SLAC* (*Stanford Linear Accelerator Center*) κοντά στο Στάνφορντ των Η.Π.Α.

Οι ψηφιακές βιβλιοθήκες, που αποτελούν αντικείμενο διερεύνησης στην παρούσα εργασία, δημιουργήθηκαν ακριβώς για να εξυπηρετήσουν την καλύτερη επικοινωνία, την ανταλλαγή εμπειριών και τη συνεργασία των διαφόρων επιστημόνων που ασχολούνται με τα θέματα αυτά σε παγκόσμιο επίπεδο. Την αναγκαιότητα αυτή των ψηφιακών βιβλιοθηκών κατανόησε και το *CERN*, το οποίο άρχισε να ασχολείται συστηματικά με το θέμα αυτό ήδη από τις αρχές της δεκαετίας του 1990, όπως θα δούμε παρακάτω.

2.2 Ευρωπαϊκός οργανισμός πυρηνικών ερευνών

Ο Ευρωπαϊκός Οργανισμός Πυρηνικών Ερευνών (*European Organization for Nuclear Research CERN*) είναι το μεγαλύτερο σε έκταση (πειραματικό) κέντρο πυρηνικών ερευνών και σωματιδιακής φυσικής στον κόσμο. Βρίσκεται δυτικά της Γενεύης, στα σύνορα Ελβετίας και Γαλλίας. Ιδρύθηκε το 1954 από δώδεκα ευρωπαϊκές χώρες, μεταξύ των οποίων και η Ελλάδα, και σήμερα αριθμεί 20 κράτη-μέλη. Η κύρια λειτουργία του αφορά στην παροχή επιταχυντών σωματιδίων και άλλων υλικοτεχνικών υποδομών που χρειάζονται για την πειραματική έρευνα στο πεδίο της φυσικής υψηλών ενεργειών. Στο *CERN* βρίσκεται ο Μεγάλος Επιταχυντής Αδρονίων (*Large Hadron Collider LHC*), ένας επιταχυντής στοιχειωδών σωματιδίων, ο οποίος φιλοξενεί διάφορα από τα τρέχοντα πειράματα του οργανισμού μεταξύ των οποίων τα *ALICE*, *ATLAS*, *CMS*, *LHCb*.

Σκοπός των πειραμάτων αυτών είναι, μεταξύ άλλων, ο σχηματισμός του – προς το παρόν μη



Σχήμα 2.1: Ο μεγάλος επιταχυντής αδρονίων. (Φωτογραφία: CERN)

παρατηρήσιμου πλην όμως αναζητούμενου – σωματιδίου Χιγκς (*Higgs Boson*). Ελπίζεται ότι από τις αναμενόμενες συγκρούσεις των πρωτονίων, στη δεδομένη ενέργεια, θα παρατηρηθεί το εν λόγω σωματίδιο, το οποίο αποτελεί και έναν από τους χαμένους κρίκους του Καθιερωμένου Μοντέλου της σωματιδιακής φυσικής. Η πειραματική παρατήρηση του μποζονίου Χιγκς θα ερμηνεύσει πώς τα υπόλοιπα στοιχειώδη σωματίδια αποκτούν μάζα κατά το μοντέλο Χιγκς.

Τα πειράματα που λαμβάνουν χώρα στο *CERN*, καθώς και τους υπόλοιπους παρόμοιους οργανισμούς ανά τον κόσμο, παράγουν καθημερινά νέα συμπεράσματα καθώς και ενδιαμέσως συμπεράσματα τα οποία τεκμηριώνονται εγγράφως από τους σχετικούς επιστήμονες. Τα έγγραφα αυτά, όπως και όλα τα υπόλοιπα έγγραφα που παράγονται, άμεσα και έμμεσα συσχετιζόμενα με τις ερευνητικές διαδικασίες και τη λειτουργία του εκάστοτε οργανισμού, πρέπει να αποθηκεύονται ψηφιακά και να είναι διαθέσιμα για αναζήτηση ανά πάσα στιγμή από τους ενδιαφερόμετους χρήστες.

Αρκετά σημαντικά επιτεύγματα στο πεδίο της φυσικής των σωματιδίων έχουν γίνει μέσα από πειράματα στο *CERN*. Σ'αυτά συμπεριλαμβάνονται:

- 1973: Η ανακάλυψη των ουδέτερων ρευμάτων στο θάλαμο φυσαλίδων Gargamelle
- 1983: Η ανακάλυψη των W και Z μποζονίων στα πειράματα *UA1* και *UA2*
- 1989: Ο προσδιορισμός του αριθμού φωτεινών οικογενειών νετρίνου που λειτουργούν στην κορυφή Z μποζονίου στον επιταχυντή *LEP*
- 1995: Η δημιουργία για πρώτη φορά ατόμων αντιυδρογόνου στο πείραμα *PS210*
- 1999: Η ανακάλυψη της άμεσης παραβίασης CP στο πείραμα *NA48*

- 2010: Η απομόνωση 38 ατόμων του αντιυδρογόνου
- 2011: Η διατήρηση αντιυδρογόνου για πάνω από 15 λεπτά
- 2012: Ένα μποζόνιο με μάζα περίπου 125 GeV το οποίο είναι το μποζόνιο *Higgs*

Το 1984 το βραβείο Νόμπελ Φυσικής απονεμήθηκε στους Carlo Rubbia και Simon van der Meer για το έργο τους που είχε ως αποτέλεσμα τα ευρήματα των μποζονίων *W* και *Z*. Το βραβείο Νόμπελ 1992 για τη Φυσική απονεμήθηκε επίσης και στον ερευνητή Georges Charpak για την εφεύρεση και την ανάπτυξη ανιχνευτών σωματιδίων.

Εκτός των καινοτομιών στον τομέα της φυσικής, είναι σημαντικό να τονιστεί ότι το World Wide Web ξεκίνησε ως ένα έργο του CERN που ονομαζόταν ENQUIRE και ξεκίνησε από τον Tim Berners-Lee το 1989 και τον Robert Cailliau το 1990. Οι Berners-Lee και Cailliau τιμήθηκαν από κοινού από την Association for Computing Machinery το 1995 για τη συμβολή τους στην ανάπτυξη του World Wide Web.

Με βάση την έννοια του υπερκειμένου, το έργο είχε ως στόχο να διευκολύνει την ανταλλαγή πληροφοριών μεταξύ των ερευνητών. Η πρώτη ιστοσελίδα δημιουργήθηκε το 1991. Στις 30 Απριλίου του 1993, το CERN ανακοίνωσε ότι ο Παγκόσμιος Ιστός θα είναι ελεύθερος σε όλους. Αντίγραφο της αρχικής πρώτης ιστοσελίδας, που δημιουργήθηκε από τον Berners-Lee, έχει δημοσιευθεί στην ιστοσελίδα του World Wide Web Consortium ως ένα ιστορικό ντοκουμέντο. Πριν από την ανάπτυξη του Ιστού, το CERN είχε πρωτοστατήσει και στην εισαγωγή της τεχνολογίας του Διαδικτύου, ξεκινώντας στις αρχές του 1980.

Πιο πρόσφατα, το CERN έχει γίνει κέντρο για την ανάπτυξη του υπολογιστικού πλέγματος, φιλοξενώντας έργα, συμπεριλαμβανομένων των Enabling Grids for E-Science (EGEE) και LHC Computing Grid. Επίσης φιλοξενεί το CERN Internet Point Exchange (CIXP), ένα από τα δύο κύρια σημεία ανταλλαγής Διαδικτύου στην Ελβετία.

Το CERN απασχολεί περίπου 2500 άτομα και οι εγκαταστάσεις του είναι διαθέσιμες για χρήση από ερευνητικά ιδρύματα από όλον τον κόσμο. Συνολικά, σχεδόν 11000 επιστήμονες από 100 διαφορετικές εθνικότητες χρησιμοποιούν τις εγκαταστάσεις του CERN. Τα εργαστήρια φιλοξενούν, επίσης, αρκετές εκατοντάδες φοιτητές, υπότροφους, μαθητευόμενους αλλά και επιστήμονες που αποσπώνται από άλλα ιδρύματα.

Το CERN παρέχει εργασία, είτε άμεσα είτε έμμεσα, σε περισσότερους ανθρώπους από οποιαδήποτε άλλη διεθνή οργανισμό στη Γενεύη. Η κατασκευή και η συντήρηση των εγκαταστάσεων του έχει ανατεθεί σε εμπορικές επιχειρήσεις από τα κράτη μέλη, πολλά από τα οποία βασίζονται στην τοπική περιοχή. Επίσης, οι τεχνολογίες που αναπτύχθηκαν για να καλύψουν τις ανάγκες έρευνας του CERN έχουν ωφελήσει την κοινωνία στο σύνολό της. Εκτός από την εφεύρεση του Web, η δέσμευση του CERN για τη μεταφορά της τεχνολογικής γνώσης έχει οδηγήσει σε πρόοδο στην ιατρική απεικόνιση και τη θεραπεία του καρκίνου, σε βιομηχανικές διεργασίες καθώς και σε πολλούς άλλους τομείς.

2.3 Ψηφιακές βιβλιοθήκες

Μια ψηφιακή βιβλιοθήκη είναι μια βιβλιοθήκη η οποία παρέχει το σύνολο του υλικού της σε ψηφιακή μορφή, το οποίο είναι προσβάσιμο μέσω ηλεκτρονικών υπολογιστών. Οι σύγχρονες ψηφιακές βιβλιοθήκες μπορούν να υφίστανται μόνο ηλεκτρονικά, χωρίς δηλαδή κάποια φυσική υποδομή, και η ιστορία τους ξεκινά την τελευταία δεκαετία του εικοστού αιώνα. Τα περιεχόμενα μιας ψηφιακής βιβλιοθήκης μπορεί να είναι φυσικά έγγραφα τα οποία έχουν ψηφιοποιηθεί ή και έγγραφα που είναι εκ της δημιουργίας τους ψηφιακά. Μια ψηφιακή βιβλιοθήκη τυπικά επιτρέπει την αναζήτηση στα περιεχόμενά της, επιτρέποντας έτσι την ευκολότερη εύρεση συγκεκριμένων πόρων της. Συχνές είναι και οι συνεργασίες μεταξύ ψηφιακών βιβλιοθηκών για συγκεντρωτικές αναζητήσεις στα περιεχόμενά τους. Σήμερα οι περισσότεροι ακαδημαϊκοί και επιστημονικοί οργανισμοί φιλοξενούν μια ψηφιακή βιβλιοθήκη του έργου τους, όπως βιβλία, άρθρα, δημοσιεύσεις, περιοδικά και άλλα. Πολλές από αυτές τις βιβλιοθήκες είναι ελεύθερα προσβάσιμες, προσφέροντας έτσι άμεσα ελεύθερη πρόσβαση στην επιστήμη.

Τα πλεονεκτήματα των ψηφιακών βιβλιοθηκών είναι πολλά. Η φυσική υπόσταση της βιβλιοθήκης δεν είναι πλέον απαραίτητη και καθένας μπορεί να την επισκεφτεί ψηφιακά, ανεξαρτήτως ώρας και μέρους. Οι πόροι της βιβλιοθήκης δεν είναι περιορισμένοι, είναι δηλαδή διαθέσιμοι για όλους ανά πάσα στιγμή. Πολύπλοκες αναζητήσεις είναι δυνατές στα περιεχόμενα της βιβλιοθήκης για την εύρεση συγκεκριμένων πόρων. Δεν υπάρχει το πρόβλημα της φθοράς από τη χρήση, και η διατήρηση είναι ευκολότερη. Οι περιορισμοί στο χώρο είναι μηδαμινοί. Τέλος, η ποιότητα αποθήκευσης πληροφοριών είναι πολύ καλύτερη.

Δε λείπουν όμως και τα μειονεκτήματα από τις ψηφιακές βιβλιοθήκες. Οι διαφορετικές τεχνολογίες που χρησιμοποιούνται από διάφορες ψηφιακές βιβλιοθήκες, καθώς και η συνεχής ανάπτυξη και βελτίωσή τους δημιουργούν το πρόβλημα της συνεπούς διατήρησης των πόρων μιας ψηφιακής βιβλιοθήκης στην πάροδο του χρόνου. Προβλήματα επίσης μπορεί να δημιουργήσει η κυριότητα και τα δικαιώματα κάποιου έργου που φιλοξενείται από μια ψηφιακή βιβλιοθήκη λόγω της φύσης της ψηφιακής πληροφορίας και της διάδοσής της. Τέλος, η πολυπλοκότητα της καταλογράφησης και της αποθήκευσης των μεταδεδομένων για κάποια έργα μπορεί να αποτελέσει πρόβλημα στην καταγραφή και αναζήτησή τους.

2.4 Συνεργαζόμενες ψηφιακές βιβλιοθήκες

Η συνεργασία μεταξύ ψηφιακών βιβλιοθηκών βασίζεται κυρίως στη δυνατότητα πρόσβασης στα περιεχόμενά τους, που παρέχουν σε άλλες μηχανές αναζήτησης. Συχνά γίνεται από ψηφιακές βιβλιοθήκες η χρήση του καθιερωμένου πρωτοκόλλου συγκομιδής μεταδεδομένων του *Open Archives Initiative (OAI-PMH)*, για την παροχή και εξαγωγή των μεταδεδομένων τους.

Υπάρχουν δύο κυρίως τεχνικές αναζήτησης σε μία συνεργασία ψηφιακών βιβλιοθηκών: η κατανομημένη αναζήτηση και η αναζήτηση σε μεταδεδομένα που έχουν ήδη συλλεχθεί. Στην περίπτωση της κατανομημένης αναζήτησης ένας εξυπηρετητής εκτελεί παράλληλες αναζητήσεις σε διάφορους εξυπηρετητές-μέλη της συνεργασίας. Τα αποτελέσματα, αφού περάσουν μια επεξεργασία, επιστρέφονται στον τελικό χρήστη. Το πλεονέκτημα αυτής της προσέγγισης είναι ότι αποφεύγεται η τοπική ευρετηρίαση και αποθήκευση του περιεχομένου των διάφορων ψηφιακών βιβλιοθηκών, και συνεπώς και το κόστος σε υπολογιστικούς πόρους που επιφέρουν. Συγχρόνως όμως αυτό αποτελεί και μειονέκτημα εφόσον δεν υπάρχει τοπικός έλεγχος για την ευρετηρίαση και κατάταξη του περιεχομένου των διάφορων ψηφιακών βιβλιοθηκών, και κατά συνέπεια και της συλλογής των απολύτως επιθυμητών αποτελεσμάτων. Στην περίπτωση της αναζήτησης σε μεταδεδομένα τα οποία έχουν ήδη συλλεχθεί, ο εξυπηρετητής φροντίζει να επικοινωνεί περιοδικά με τις διάφορες ψηφιακές βιβλιοθήκες της συνεργασίας, συλλέγοντας μεταδεδομένα και δημιουργώντας ένα τοπικό ευρετήριο πληροφοριών. Το πρωτόκολλο (*OAI-PMH*), για το οποίο έγινε λόγος παραπάνω, χρησιμοποιείται συχνά σε αυτές τις περιπτώσεις. Το πλεονέκτημα αυτής της προσέγγισης είναι ότι ο τελικός μηχανισμός αναζήτησης έχει πλήρη έλεγχο στην ευρετηρίαση και κατάταξη των αποτελεσμάτων, παράγοντας πιο συνεπή αποτελέσματα. Στον αντίποδα, οι διαδικασίες συγκομιδής και ευρετηρίασης είναι ιδιαίτερα απαιτητικές σε υπολογιστικούς πόρους.

2.5 Παραδείγματα ψηφιακών βιβλιοθηκών

Παρακάτω παρουσιάζονται μερικές από τις πιο διαδεδομένες πλατφόρμες ψηφιακών βιβλιοθηκών.

2.5.1 *EPrints*



Σχήμα 2.2: Eprints λογότυπο

Το *EPrints* είναι ένα πακέτο ελεύθερου λογισμικού που χρησιμοποιείται για την κατασκευή αποθετηρίων εγγράφων ελεύθερης πρόσβασης. Είναι συμβατό με το πρωτόκολλο συγκομιδής μεταδεδομένων του Open Archives Initiative, και χρησιμεύει κυρίως ως ένα σύστημα διαχείρισης εγγράφων για ιδρύματα όπως ακαδημαϊκοί και επιστημονικοί οργανισμοί. Η ανάπτυξη του ξεκίνησε το 2000 από το School of Electronics and Computer Science του πανεπιστημίου του Southampton, και διατίθεται βάσει της άδειας *GPL*.

Το *EPrints* χρησιμοποιείται ως μια διαδικτυακή εφαρμογή βασισμένη στην αρχιτεκτονική *LAMP* (*Linux Apache MySQL PHP/Python/Perl*), όπου στην προκειμένη περίπτωση το λογισμικό είναι γραμμένο σε *Perl*. Το λογισμικό χρησιμοποιεί *plugins* για την εισαγωγή και εξαγωγή δεδομένων και την μετατροπή αντικειμένων ενώ *widgets* χρησιμοποιούνται για τη διεπαφή του χρήστη. Η ρύθμιση και προσαρμογή του λογισμικού μπορεί να γίνει μέσω μιας σειράς από *XML* αρχεία, ενώ η εμφάνιση του ελέγχεται από προσαρμόσιμα πρότυπα *HTML*.

2.5.2 DSpace



Σχήμα 2.3: DSpace λογότυπο

Το *DSpace* είναι ένα πακέτο ελεύθερου λογισμικού που παρέχει εργαλεία για τη διαχείριση ψηφιακών πόρων. Προορίζεται για χρήση από ιδρύματα όπως ακαδημαϊκοί και επιστημονικοί οργανισμοί ως εργαλείο διαχείρισης ψηφιακών πόρων, αλλά και ως πλατφόρμα ψηφιακής διατήρησης. Υποστηρίζει ένα πλήθος τύπων δεδομένων, όπως βιβλία, εργασίες, διατριβές, ψηφιακές σαρώσεις αντικειμένων τριών διαστάσεων, αρχεία πολυμέσων και άλλα, τα οποία είναι οργανωμένα σε συλλογές αντικειμένων. Το λογισμικό αναπτύσσεται από το 2002 ως συνεργασία δύο φορέων, του πανεπιστημίου *MIT* της Βοστώνης και των εργαστηρίων έρευνας της εταιρίας *Hewlett – Packard*, και διατίθεται βάσει της άδειας *BSD*.

Η κοινότητα του *DSpace* χρησιμοποιεί ένα μοντέλο ανάπτυξης σύμφωνα με το οποίο μέσα από ένα σύνολο χρηστών, διαφορετικές ομάδες φροντίζουν να αναπτύσουν το λογισμικό και να ελέγχουν την ορθότητα και συμφωνία του με τους καθορισμένους κανόνες ανάπτυξης. Πρόκειται κυρίως για μια διαδικτυακή εφαρμογή. Το λογισμικό είναι γραμμένο στη γλώσσα προγραμματισμού *Java*, με χρήση της διεπαφής προγραμματισμού εφαρμογών *Java Servlet*, ενώ χρησιμοποιείται και μια τυπική σχεσιακή βάση δεδομένων, όπως η *PostgreSQL* ή η *Oracle*. Υποστηρίζεται και η συγκομιδή μεταδεδομένων μέσω του αντίστοιχου πρωτοκόλλου *OAI-PMH* του *Open Archives Initiative*.

2.5.3 Fedora

Το *Fedora* είναι μια πλατφόρμα αρχιτεκτονικής διαχείρισης ψηφιακών πόρων, βασισμένη στη λογική των αυτόνομων λογισμικών μονάδων, και μπορεί να αποτελέσει τη βάση για την κατασκευή ψηφιακών βιβλιοθηκών, αποθετηρίων ακαδημαϊκών και επιστημονικών οργανισμών καθώς και συστημάτων ψηφιακών αρχείων και διατήρησης. Το λογισμικό προορίζεται



Σχήμα 2.4: Fedora λογότυπο

ως η αρχιτεκτονική βάση, αλλά δεν αποτελεί ολοκληρωμένη λύση διαχείρισης, ευρετηρίασης, εύρεσης και διάθεσης, για τη δημιουργία ενός ψηφιακού αποθετηρίου.

Η ανάπτυξη του ξεκίνησε το 1997 από το πανεπιστήμιο *Cornell*, ενώ σήμερα αναπτύσσεται από κοινού από το ίδιο πανεπιστήμιο καθώς και από τη βιβλιοθήκη του πανεπιστημίου της Βιρτζίνια, και διατίθεται βάσει της άδειας Apache. Το βασικό χαρακτηριστικό του λογισμικού είναι ένα στρώμα διαχείρισης γενικής χρήσης για ψηφιακά αντικείμενα. Η διαχείριση των αντικειμένων αυτών γίνεται μέσω μοντέλων περιεχομένου τα οποία αναπαριστούν αντικείμενα δεδομένων ή και συλλογές τους. Τα αντικείμενα αυτά μπορούν να περιέχουν συνδέσμους με εγγραφές, μεταδεδομένα, δεδομένα διαχείρισης ή ακόμη και εντολές/πράξεις. Η πρόσβαση στο στρώμα διαχείρισης μπορεί να γίνει είτε μέσω μιας εφαρμογής πελάτη, είτε μέσω μιας διεπαφής προγραμματισμού εφαρμογών για διαδικτυακή πρόσβαση. Η εισαγωγή και εξαγωγή ψηφιακών αντικειμένων μπορεί να πραγματοποιηθεί με αρχεία *XML* διαφόρων τύπων.

2.5.4 *Greenstone*



Σχήμα 2.5: Greenstone λογότυπο

Το *Greenstone* αποτελεί μια σειρά εργαλείων λογισμικού για την κατασκευή και διάθεση ψηφιακών βιβλιοθηκών. Χρησιμοποιείται για τη δημιουργία εκτενών συλλογών ψηφιακών εγγράφων με δυνατότητα αναζήτησης και προσφέρει μια διεπαφή διαχείρισης τους για βιβλιοθηκονόμους. Χειρίζεται ένα πλήθος μορφών αρχείων, όπως κείμενο και πολυμέσα, ενώ ειδικά για τα αρχεία κειμένου χρησιμοποιεί μια εσωτερική δομή βασισμένη σε *XML*. Είναι ελεύθερο λογισμικό, αναπτύσσεται από το πρόγραμμα ψηφιακών βιβλιοθηκών της Νέας Ζηλανδίας στο πανεπιστήμιο του *Waikato*, και διατίθεται βάσει της άδειας *GPL*.

Κεφάλαιο 3

Ο εξυπηρετητής αρχείων του CERN: CDS

3.1 Ιστορικό

Το 1993 πρωτοεμφανίστηκε στο διαδίκτυο με τη μορφή του *CERN Preprint Server* το *CDS (CERN Document Server)*, με σκοπό τη συλλογή για διάδοση όλων των ερευνητικών εγγράφων σχετικών με φυσική υψηλών ενεργειών. Χρησιμοποιήθηκε κυρίως ως θεσμικό αποθετήριο/πηγή πληροφοριών (*institutional repository*) με αρχικά δύο ξεχωριστές συλλογές εγγράφων: άρθρα από το ίδιο το *CERN* και επιστημονικά έγγραφα και άρθρα από όλο τον κόσμο τα οποία αναλάμβανε να σαρώνει (σκαννάρει / scan) η βιβλιοθήκη του *CERN*.

Το 1996 μετονομάστηκε σε *CERN Library server (weblib)* χρησιμοποιώντας το ίδιο λογισμικό για να παρέχει πρόσβαση σε επιπλέον υλικό όπως περιοδικά βιβλία και το περισσότερο υλικό που ήταν διαθέσιμο στη βιβλιοθήκη.

Το 2000 η νέα έκδοση του λογισμικού ήταν διαθέσιμη, με όνομα αυτή τη φορά: *CERN Document Server Software: CDSware*. Η νέα αυτή έκδοση επέτρεψε την εισαγωγή αρχείων πολυμέσων στην ψηφιακή βιβλιοθήκη, όπως φωτογραφίες, βίντεο, φυλλάδια, αφίσες και άλλα. Συγχρόνως το λογισμικό έγινε σύμμορφο προς τους κανόνες του *OAI (Open Archives Initiative)* και άρχισε να διανέμεται και να χρησιμοποιείται σε περισσότερα μέρη όπως το *San Diego Supercomputer Center* στο Σαν Ντιέγκο των Η.Π.Α., το *HBZ NRW* στην Κολωνία της Γερμανίας και το Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης. Υπό αυτή τη μορφή και από το 2004 και μετά το *CDSware* επιλέχτηκε και ως το επίσημο σύστημα διαχείρισης αρχείων της διεύθυνσης του *CERN* για όλα τα εισερχόμενα και εξερχόμενα έγγραφα. Την ίδια χρονιά ξεκίνησε η συνεργασία μεταξύ του *CERN* και του *EPFL* για την συμμετοχική ανάπτυξη του λογισμικού.

Το 2006 το *CDSware* μετονομάστηκε σε *CDS*, το οποίο είναι και το τρέχον όνομα του λογισμικού, και έγινε πιο ξεκάθαρος ο ρόλος του *CDS Software Consortium* ως ομάδας που αναπτύσσει και συντηρεί λογισμικό διαχείρισης εγγράφων και συνεδρίων με κυρίως προϊόντα το *CDS* και το *CDS Indico*.

Το 2007 ξεκίνησε η συνεργασία μεταξύ του *CDS* και του *SPIRES*¹ με σκοπό τη δημιουργία του *INSPIRE*, μιας συγκεντρωτικής ψηφιακής βιβλιοθήκης για όλα τα έγγραφα σχετικά με φυσική υψηλών ενεργειών. Το σχέδιο, που είναι μια συνεργασία μεταξύ τεσσάρων οργανισμών ανά τον κόσμο (*CERN*, *DESY*, *Fermilab*, *SLAC*) προβλέπεται να είναι σε χρήση μέσα στο 2010. Η εκκίνηση αυτής της συνεργασίας έχει οδηγήσει στην ανάπτυξη νέων εργαλείων για το *CDS* όπως εξελιγμένα εργαλεία συντήρησης και εμπλουτισμού μεταδεδομένων (*metadata*), οντολογικά εργαλεία ανάλυσης παραθέσεων και λέξεων-κλειδιών καθώς και εργαλείων αποσαφήνισης συγγραφέων.

3.2 Παρούσα κατάσταση

Επί του παρόντος το *CDS* βρίσκεται προ των πυλών της έκδοσης 3.0 του λογισμικού του, η οποία αναμένεται προς το τέλος του έτους 2017. Η έκδοση 3.0 του *CDS* θα περιέχει πολλά από τα εργαλεία και χαρακτηριστικά της επερχόμενης διάθεσης του προαναφερθέντος *INSPIRE*. Η κυρίως ανάπτυξη του *CDS* γίνεται από το *CDS Software Consortium*, μια ομάδα μηχανικών λογισμικού βασισμένη στο *CERN* αλλά συγχρόνως συνεισφορές παρέχονται και από τις διάφορες συνεργασίες με άλλες ομάδες, ακόμη από μεμονωμένους προγραμματιστές και μηχανικούς λογισμικού. Το ίδιο το *CDS Software Consortium* συγκροτείται από μερικά μόνιμα μέλη-εργαζομένους στο *CERN* και από διάφορους σπουδαστές, τελειόφοιτους και μεταπτυχιακούς — συμπεριλαμβανομένου και του γράφοντα, που κατά καιρούς αποτελούν μέρος της ομάδας. Η σύσταση αυτή της ομάδας αλλά και οι συνεισφορές από τους διάφορους συνεργάτες επιτρέπουν στο *CDS* να ανανεώνεται συνεχώς με νέες ιδέες, λύσεις και τεχνικές ανάπτυξης καθιστώντας το ιδιαίτερα δυναμικό.

3.3 Γενική επισκόπηση

Το *CDS* είναι ένα σύνολο από εφαρμογές που αποτελούν το πλαίσιο και τα εργαλεία για τη δημιουργία και τη διαχείριση ενός αυτόνομου εξυπηρετητή ψηφιακής βιβλιοθήκης. Το λογισμικό του αποτελεί ελεύθερο λογισμικό και λογισμικό ανοιχτού κώδικα και διατίθεται υπό την άδεια GNU General Public Licence. Η τεχνολογία που προσφέρεται από το λογισμικό αυτό καλύπτει όλες τις πλευρές της διαχείρισης μια ψηφιακής βιβλιοθήκης. Η υψηλή απόδοση και η προσαρμοστικότητα του το καθιστούν μια ολοκληρωμένη λύση για την πλήρη διαχείριση ενός αποθετηρίου εγγράφων μεσαίου και μεγάλου μεγέθους. Στο *CERN* η υπάρχουσα υλοποίηση του *CDS* φιλοξενεί πάνω από 1.000.000 βιβλιογραφικά αρχεία διάφορων κατηγο-

ριών, όπως άρθρα, βιβλία, περιοδικά, φωτογραφίες, βίντεο και άλλα. Εκτός από την τρέχουσα τοπική υλοποίηση, το *CDS* χρησιμοποιείται σήμερα από τουλάχιστον ακόμη 15 επιστημονικούς οργανισμούς παγκοσμίως. Το λογισμικό έχει υποστεί μια συνεχή σταδιακή ανάπτυξη ώστε να καταλήξει από την αρχική βασική του έκδοση ως ψηφιακού εξυπηρετητή στο τρέχον σύστημα αποθήκευσης αρχείων υψηλής απόδοσης. Παρά την πολύπλοκότητα που έχει επιφέρει αυτή η πολυετής ανάπτυξη, το *CDS* έχει καταφέρει να διατηρήσει την υψηλή του απόδοση, τη φιλικότητα προς το χρήστη και τη δυνατότητα παραμετροποίησής του σε μεγάλο βαθμό με το να παραμένει σύμμορφο σε μια καθιερωμένη αρχιτεκτονική αυτόνομων λογισμικών μονάδων (*modular architecture*). Από καθαρά τεχνικής άποψης, το *CDS* τρέχει πάνω σε συστήματα *GNU/Unix*, χρησιμοποιώντας τον εξυπηρετητή βάσεων δεδομένων *MySQL* και τον εξυπηρετητή εφαρμογών διαδικτύου *Apache/Python*. Η διαμόρφωση κατά τη διάρκεια της μετάφρασης (*compile – time configuration*) επιτυγχάνεται μέσω του εργαλείου *GNU Autoconf* ενώ η διαμόρφωση κατά τη διάρκεια της εκτέλεσης (*run-time configuration*) μέσω διάφορων πινάκων δεδομένων διαμόρφωσης της βάσης δεδομένων *MySQL* (*MySQL configuration tables*). Ο κώδικας που αποτελεί το λογισμικό είναι σχεδόν εξ' ολοκλήρου γραμμένος στη γλώσσα προγραμματισμού *Python*.

Το χαρακτηριστικό-κλειδί της αρχιτεκτονικής του *CDS* βρίσκεται στη λογική των αυτόνομων λογισμικών μονάδων. Κάθε μία από αυτές τις μονάδες περιλαμβάνει μια συγκεκριμένη και ορισμένη λειτουργία του συνολικού συστήματος της ψηφιακής βιβλιοθήκης. Οι μονάδες αλληλεπιδρούν μεταξύ τους καθώς και με τη βάση δεδομένων και το στρώμα διεπαφής. Η λογική κάθε μονάδας, η λειτουργία και διαλειτουργικότητά της είναι επεκτάσιμες και προσαρμόσιμες / επιδεκτικές προσαρμογής.

Το διάγραμμα αναπαριστά σχηματικά τη ροή εργασίας από πάνω προς τα κάτω ενός εγγράφου στο *CDS*. Στην κορυφή του διαγράμματος, η απόκτηση των δεδομένων πραγματοποιείται από 3 διαφορετικές πηγές: άμεση υποβολή από τον ίδιο το συγγραφέα (χρησιμοποιώντας ηλεκτρονικό ταχυδρομείο ή τη διαδικτυακή διεπαφή) και συγκομιδή από πηγές δεδομένων σύμμορφες με το *OAI* (*Open Archives Initiative*) ή και όχι. Τα μεταδεδομένα (*metadata*) που συγκεντρώνονται μετατρέπονται κατευθείαν σε μια πρότυπη εσωτερική δομή απεικόνισης μεταδεδομένων (*MARCXM*) ενώ τα πλήρη κείμενα υποβάλλονται και αποθηκεύονται κατευθείαν στον εξυπηρετητή αρχείων αφού μετατραπούν πρώτα σε *PDF* (*Portable Document Format*). Μόλις τα μεταδεδομένα φορτωθούν στον εξυπηρετητή μπορεί να υποβληθούν σε διαδικασίες αξιολόγησης της ποιότητάς τους από τους καταλογογράφους της βιβλιοθήκης. Επίσης, τα μεταδεδομένα μπορούν να εμπλουτιστούν με παραπομπές που εξάγονται από αντίστοιχα πλήρη κείμενα. Στη συνέχεια ο εξυπηρετητής μπορεί, εάν του ζητηθεί, να υπολογίσει και να παράγει περιεχόμενα, τάξεις, ομάδες και βιβλιογραφικές διατάξεις, κατάλληλα για ταχεία ανάκτηση. Η πληροφορία εν τέλει διανέμεται, όπως φαίνεται στο κάτω μέρος του διαγράμματος, στον τελικό χρήστη και σε παρόχους υπηρεσιών *OAI* μέσω αιτημάτων *OAI-PMH* (*The Open Archives Initiative Protocol for Metadata Harvesting*), ειδοποιήσεων ηλεκτρονικού ταχυδρομείου και της διαδικτυακής μηχανής αναζήτησης. Οι πε-

ρισσότερες από τις αλληλεπιδράσεις μεταξύ των διαφόρων αυτόνομων λογισμικών μονάδων-κλειδιών και της βάσης δεδομένων συγχρονίζονται μέσω ενός προγραμματιστή εργασιών (*task scheduler Celery*), ο οποίος δύναται να χρησιμοποιηθεί και για την εκτέλεση περιοδικών εργασιών.

3.4 Η τεχνολογία πίσω από το CDS

Σε αυτήν την ενότητα δίνεται έμφαση σε κάποια από τα εσωτερικά τεχνολογικά χαρακτηριστικά του CDS, κυρίως αυτά που το διαφοροποιούν από άλλα ελεύθερα λογισμικά και πλατφόρμες ψηφιακών βιβλιοθηκών και αποθετηρίων εγγράφων όπως το *DSPace* ή το *EPrints*. Επίσης περιγράφονται συνοπτικά οι διάφορες αυτόνομες λογισμικές μονάδες που συγκροτούν το λογισμικό.

3.4.1 Η τεχνολογία πίσω από το CDS

Ολα τα βιβλιογραφικά δεδομένα στο CDS αναπαριστώνται εσωτερικά με τη δομή *MARC 21*. Αυτή η δομή εσωτερικής αναπαράστασης μεταδεδωμένων επιλέχτηκε για διάφορους λόγους:

- Είναι ένα εδραιωμένο πρότυπο για τον κόσμο των βιβλιοθηκών, χρησιμοποιούμενο από το 1960.
- Συνδυάζεται καλά με σύγχρονες περιγραφικές τεχνολογίες όπως η γλώσσα *XML*. (το CDS χρησιμοποιεί την προσφάτως τυποποιημένη διάταξη *MARCXML* που παρέχεται από τη Βιβλιοθήκη του Κογκρέσου των Ηνωμένων Πολιτειών)
- Είναι ευπροσάρμοστο και συνεπώς μακροπρόθεσμα αξιόπιστο.
- Είναι απόλυτα επεκτάσιμο ώστε να προσαρμόζεται σε οποιαδήποτε δομή μεταδεδωμένων (η τρέχουσα διάταξη *MARCXML* στο CDS περιλαμβάνει περισσότερα από 150 πεδία μεταδεδωμένων)

Τα αποθετήρια που φιλοξενούν ομοιογενή έγγραφα συνήθως δε χρειάζεται να χρησιμοποιήσουν ένα πλήρες σύστημα καταλογογράφησης *MARC*. Σε αυτή την περίπτωση μπορούν να υιοθετήσουν την προεπιλεγμένη περιγραφική δομή που προσφέρει το CDS και η οποία καλύπτει τα πιο κοινά πεδία μεταδεδωμένων. Σε διαφορετική περίπτωση, ο χρήστης μπορεί να υλοποιήσει, επιπλέον της προεπιλεγμένης δομής, επιπρόσθετα περιγραφικά πεδία για ειδικούς και τοπικούς τύπους μεταδεδωμένων. Σε μια ακόμη πιο προχωρημένη περίπτωση ο χρήστης μπορεί να θέλει να ορίσει ιδιαίτερα συγκεκριμένα και ξεχωριστά αντικείμενα δεδομένων, υλοποιώντας έτσι μια καινούργια περιγραφική δομή εξ αρχής. Αυτή η ακραία περιγραφική επεκτασιμότητα μαζί με την παραμετροποίηση του CDS επιτρέπουν πρακτικά στο σύστημα να

χειρίζεται οποιοδήποτε τύπο μεταδεδομένων (από εκθέματα ενός μουσείου ως παρουσιάσεις πολυμέσων).

3.4.2 Απόκτηση των μεταδεδομένων

Η απόκτηση των μεταδεδομένων πραγματοποιείται από αυτόματες και ημιαυτόματες διαδικασίες συγκομιδής εγγράφων (μονάδα BibHarvest) – εφαρμόζοντας είτε τυποποιημένες προσεγγίσεις όπως η χρήση της συγκομιδής δεδομένων σύμμορφης με το *OAI – PMH* είτε ειδικές διαδικασίες όπως τη ρηχή διαδικτυακή συγκομιδή. Η υποβολή εγγράφων γίνεται κατευθείαν από τους συγγραφείς τους μέσω του διαδικτύου ή ηλεκτρονικού ταχυδρομείου μέσω των μονάδων *WebSubmit* και *ElmSubmit*. Και στις δύο περιπτώσεις τα μεταδεδομένα συγκεντρώνονται πρώτα ακατέργαστα, στη συνέχεια μετατρέπονται στην εσωτερική δομή αναπαράστασης του *CDS* και τέλος καταχωρούνται στον εξυπηρετητή βιβλιογραφικών μεταδεδομένων.

Η μετατροπή λαμβάνει χώρα στη μονάδα *BibConvert*, η οποία επιτρέπει μετατροπές μεταξύ διαφόρων διατάξεων ακολουθίας (π.χ. *ISO2709*) και ημι-δομημένων διατάξεων (π.χ. *XML*) αλλά και μεταξύ διαφόρων διατάξεων μεταδεδομένων (π.χ. *MARCXML*, *DublinCore*, *RFC1807*). Η εν λόγω μονάδα παρέχει επίσης τη δυνατότητα λεπτομερούς μορφοποίησης κειμένου, συμπεριλαμβανομένων και των κανονικών εκφράσεων που προσδιορίζονται στη γλώσσα διαμόρφωσης του *BibConvert*. Η γλώσσα διαμόρφωσης αυτή παρέχει έναν ορισμό της σύνταξης και σημασιολογίας της περιγραφής της μετατροπής των μεταδεδομένων, ο οποίος κωδικοποιείται σε ένα σύνολο προτύπων μετατροπής. Τα πρότυπα αυτά αναλαμβάνουν την εξόρυξη της περιγραφής του κάθε εγγράφου, παρέχουν την περιγραφή κάθε πεδίου μεταδεδομένων για κάθε έγγραφο καθώς και τη διαρρύθμισή τους.

Επιπλέον, το *BibConvert* παρέχει ένα μηχανισμό ελέγχου αντιστοιχίας των νέων εγγράφων με τα υπάρχοντα βιβλιογραφικά μεταδεδομένα στον εξυπηρετητή με σκοπό την αποφυγή διπλοεγγράφων (π.χ. σε περίπτωση πολλαπλής καταχώρησης). Τα έγγραφα τα οποία σημειώνονται ως διφορούμενα σε αυτό το στάδιο απορρίπτονται προσωρινά και προωθούνται για έλεγχο από το διαχειριστή απόκτησης των μεταδεδομένων. Εν συνεχεία τα επιβεβαιωμένα μεταδεδομένα καταχωρούνται στον εξυπηρετητή υπό τη δομή *MARC21*.

Η μετατροπή των μεταδεδομένων με το *BibConvert* επιτρέπει σε μεγάλο βαθμό την αυτοματοποίηση της όλης διαδικασίας: έγγραφα από διαφορετικές πηγές μπορούν εύκολα να καταχωρηθούν με τη διάταξη *MARCXML* και άμεσα να εισαχθούν στο σύστημα με χρήση απλά μερικών τυποποιημένων προτύπων διαμόρφωσης. Προκειμένου να διαπιστωθεί η γνησιότητα των μεταδεδομένων που εισάγονται στο σύστημα, οι καταλογογράφοι της βιβλιοθήκης μπορούν να πραγματοποιήσουν ελέγχους αξιολόγησης της ποιότητάς τους μέσω της μονάδας *BibCheck*.

3.4.3 Ευρετηρίαση και κατάταξη

Οι μονάδες ευρετηρίασης και κατάταξης αποτελούν τον πυρήνα του CDS. Καθιερώθηκαν ειδικά σχεδιασμένοι δείκτες ευρετηρίου ώστε να παρέχουν εξαιρετική ταχύτητα ανταπόκρισης σε αποθετήρια μεγέθους ενός εκατομμυρίου εγγράφων. Εκτός από την αναζήτηση στα μεταδεδομένα, το CDS παρέχει επίσης τη δυνατότητα ευρετηρίου και αναζήτησης σε πλήρη κείμενα και τις αναφορές τους έτσι ώστε η συνδυασμένη αναζήτηση μεταδεδομένων / πλήρους κειμένου / αναφορών να είναι δυνατή. Για παράδειγμα μπορεί κάποιος να κάνει την παρακάτω αναζήτηση: find all documents written by Ellis in 2002 that mention the term Higgs boson in the fulltext and that refer to Physical Review D 1997 paper.

Τα αποτελέσματα που ανακτώνται από τη μηχανή αναζήτησης μπορούν να καταταχθούν επιπλέον σύμφωνα με διάφορα κριτήρια. Η προεπιλεγμένη εγκατάσταση του CDS περιλαμβάνει το κλασικό μοντέλο διαστήματος δεικτών συχνότητας λέξεων που επιτρέπει την ανάκτηση παρόμοιων εγγράφων. Επιπροσθέτως, περιλαμβάνει ένα μηχανισμό μεθόδων κατάταξης που βασίζεται σε συγκεκριμένες τιμές μεταδεδομένων. Για παράδειγμα, η μέθοδος κατάταξης σύμφωνα με τον συντελεστή απήχησης (*impact factor*) των περιοδικών η οποία κατατάσσει τα έγγραφα χρησιμοποιώντας μια παραμετροποιήσιμη γνωσιακή βάση περιοδικών και των αντίστοιχων συντελεστών απήχησης. Τέλος, πρόσφατες μέθοδοι παρέχουν τη δυνατότητα κατάταξης σύμφωνα με τον αριθμό των αναφορών και τον αριθμό των μεταφορτώσεων.

Τα αποτελέσματα της αναζήτησης ομαδοποιούνται σε συλλογές (*collections*). Ο διαχειριστής του συστήματος έχει τη δυνατότητα να ορίσει δέντρα κανονικών αλλά και εικονικών συλλογών διευκολύνοντας έτσι την πλοήγηση του χρήστη στο σώμα των εγγράφων που φιλοξενεί η ψηφιακή βιβλιοθήκη. Παράλληλα, μελετώνται διάφορες τεχνικές αυτοματοποιημένης ομαδοποίησης που θα παρέχουν έναν έξυπνο και αυτόματο τρόπο ομαδοποίησης των αποτελεσμάτων και πλοήγησης σε αυτά.

3.4.4 Διεπαφή και εξατομίκευση

Το CDS μπορεί να χειριστεί σχεδόν οποιοδήποτε είδος ηλεκτρονικού υλικού, χάρη στην ευέλικτη δομή *MARC*, όπως περιγράφεται ανωτέρω. Για να απεικονίζει σωστά όμως όλες αυτές τις διάφορες μορφές εγγράφων σύμφωνα με τα χαρακτηριστικά που ορίζει η κάθε μια, χρησιμοποιεί έναν ευέλικτο μηχανισμό εξόδου που δίνει τη δυνατότητα της αυτόματης δημιουργίας συνδέσμων προς εξωτερικούς πόρους με βάση το περιεχόμενο του εγγράφου.

Το γραφικό περιβάλλον του λογισμικού παρέχει στο χρήστη μια σειρά από επιλογές εξατομίκευσης και συλλογικότητας. Ο τελικός χρήστης μπορεί να ορίσει προσωπικές συλλογές εγγράφων (καλάθια *baskets*) και περιοδικές ειδοποιήσεις (*alerts*) σχετικές με την προσθήκη νέων εγγράφων στον εξυπηρετητή με βάση τα πεδία του ενδιαφέροντός του. Ως μέρος των συλλογικών χαρακτηριστικών που παρέχει το λογισμικό ο χρήστης μπορεί να ορίσει το περιε-

χόμενο ενός καλαθιού ως μοιραζόμενο στα μέλη μιας ομάδας, καθώς επίσης και να σχολιάσει ή να αξιολογήσει έγγραφα.

3.4.5 Περιγραφή των αυτοτελών λογισμικών μονάδων

Το χαρακτηριστικό-κλειδί της αρχιτεκτονικής του CDS βρίσκεται στη λογική των αυτόνομων λογισμικών μονάδων. Κάθε μία από αυτές τις μονάδες αναλαμβάνει μια συγκεκριμένη λειτουργία του όλου συστήματος. Κατά αυτό τον τρόπο το σύστημα παραμένει ευέλικτο και παραμετροποιήσιμο και η συντήρησή του διευκολύνεται. Ο σκοπός της κάθε μονάδας είναι ορισμένος και διαφορετικός από τις υπόλοιπες μονάδες, η λειτουργία τους όμως δεν είναι πάντα ανεξάρτητη. Αν χρειάζεται, μια μονάδα μπορεί να εισάγει και να χρησιμοποιεί λειτουργίες κάποιας άλλης.

Αυτοτελείς λογισμικές μονάδες σχετικές με την απόκτηση δεδομένων

Η εισαγωγή δεδομένων στο CDS γίνεται ή μπορεί να γίνει με δύο τρόπους : είτε με μια αυτοματοποιημένη διαδικασία την οποία ορίζει ο διαχειριστής, είτε με άμεση εισαγωγή από το χρήστη. Στην πρώτη περίπτωση η μονάδα *BibHarvest* αναλαμβάνει τη συγκομιδή των δεδομένων από αποθετήρια *OAI*, η μονάδα *BibConvert* τη μετατροπή των δεδομένων αυτών στην προεπιλεγμένη δομή *XML MARC* και τέλος η μονάδα *BibUpload* την καταχώρηση των τελικών δεδομένων στον εξυπηρετητή. Στη δεύτερη περίπτωση ο χρήστης αρχικά καλείται να εισάγει και να υποβάλει έγγραφα χρησιμοποιώντας τη μονάδα *WebSubmit* και στη συνέχεια τη μονάδα *BibEdit* για να επεξεργαστεί τα αντίστοιχα μεταδεδομένα αν το επιθυμεί. Οι σχετικές λογισμικές μονάδες περιγράφονται παρακάτω.

BibHarvest Η μονάδα αυτή ουσιαστικά αναλαμβάνει τη συγκομιδή και διάθεση μεταδεδομένων, σύμμορφη με τους κανόνες και προδιαγραφές του *Open Archives Initiative Protocol for Metadata Harvesting (OAI – PMH)*. Σύμφωνα με το εν λόγω πρωτόκολλο, το αποθετήριο των μεταδεδομένων αυτών είναι χτισμένο απευθείας πάνω από τη βάση δεδομένων και διαθέτει ένα διαχειριστή αποθετηρίου *OAI* που επιτρέπει την εκτέλεση των διοικητικών καθηκόντων στον αποθηκευτικό χώρο πέραν των κυρίως διαχειριστικών λειτουργιών που προσφέρονται. Η βάση δεδομένων μπορεί να είναι μερικώς ή εξ ολοκλήρου ανοιχτή για συγκομιδή στο πεδίο εφαρμογής του *OAI-PMH* πρωτοκόλλου. Στην περίπτωση αυτή τα μεταδεδομένα παρέχονται σε ακατέργαστη μορφή και η σημασιολογία των διαφόρων πεδίων τους ορίζεται στο πλαίσιο της δομής *MARC 21*.

BibConvert Η μονάδα αυτή επιτρέπει την μετατροπή μεταδεδομένων από οποιαδήποτε δομημένη ή ημιδομημένη μορφή σε οποιαδήποτε άλλη. Στο CDS η μετατροπή γίνεται κατά βάση στη δομή *XML MARC*, η οποία είναι και η προεπιλεγμένη. Παρόλα αυτά, οι μορφές εισόδου και εξόδου για τη μετατροπή είναι απόλυτα παραμετροποιήσιμες. Η δύναμη της εν

λόγω μονάδας έγκειται στο γεγονός ότι δεν θεωρεί δεδομένη κάποιου είδους διάρθρωση για τα δεδομένα εισόδου, αλλά δίνει τη δυνατότητα στο διαχειριστή να προσδιορίσει μια γλώσσα διαμόρφωσης ορίζοντας τη σύνταξη και τη σημασιολογία της κάθε μετατροπής. Αναπόφευκτα, η γλώσσα διαμόρφωσης είναι ιδιαίτερα πολύπλοκη. Το *CDS* παρέχει τις διαμορφώσεις για τις πιο συχνές μετατροπές μορφών μεταξύ τους. Γενικά μια διαμόρφωση θα περιέχει την περιγραφή της δομής των δεδομένων εισαγωγής καθώς και αυτών της εξαγωγής. Ο επεξεργαστής της μονάδας αναλαμβάνει να αναλύσει τα δεδομένα εισόδου δημιουργώντας εν τέλει τη δομή εξόδου με τα δεδομένα που προκύπτουν. Στις περισσότερες περιπτώσεις η χρήση του *BibConvert* στοχεύει σε δεδομένα που στερούνται XML αναπαράστασης και η επεξεργασία τους συνεπάγεται διάφορα βήματα περιλαμβανομένου του διαχωρισμού εγγράφων, της εξόρυξης πεδίων, της μετατροπής των τιμών τους και τη μορφοποίησής τους. Προϋπόθεση είναι τα δεδομένα εισόδου να βρίσκονται σε οποιαδήποτε δομημένη ή ημιδομημένη μορφή (για παράδειγμα να μην είναι εκφρασμένα σε κάποια φυσική γλώσσα που αποτελεί αντικείμενο εξαγωγής δεδομένων). *BibMatch* Η μονάδα αυτή ελέγχει τα νεοεισρχθέντα δεδομένα ταιριάζοντας τα με τα ήδη υπάρχοντα στη βάση δεδομένων προς αποφυγή διπλών εγγράφων.

BibUpload Η μονάδα αυτή επιτρέπει την καταχώρηση νέων δεδομένων και εγγράφων στη βάση δεδομένων. Είθισται τα μεταδεδομένα που καταχωρούνται να βρίσκονται σε ένα καλώς δομημένο *XML* αρχείο, σύμφωνα με την τρέχουσα προεπιλεγμένη δομή, και συνήθως παρέχονται απευθείας από τη μονάδα *BibConvert*.

WebSubmit Η μονάδα αυτή είναι μια διεπαφή χρήστη και αποτελεί ουσιαστικά ένα ολοκληρωμένο σύστημα υποβολής που επιτρέπει σε εξουσιοδοτημένα άτομα (συντάκτες, γραμματείς, προσωπικό συντήρησης του αποθετηρίου) την υποβολή μεμονωμένων εγγράφων στον εξυπηρετητή. Το σύστημα αυτό διαθέτει ένα μηχανισμό ροής ελέγχου που εξασφαλίζει την έγκριση των δεδομένων από εξουσιοδοτημένες μονάδες. Συνολικά, υπάρχουν διάφορα διαθέσιμα σχήματα αξιοποίησης της υποβολής, περιλαμβανομένου ενός αυτοματοποιημένου μετατροπέα από διάφορες μορφές κειμένου και εικόνων σε πλήρη έγγραφα. Η μονάδα παρέχει επίσης τη δυνατότητα εξόρυξης πληροφοριών με έμφαση σε βιβλιογραφικές οντότητες, όπως αναφορές, συγγραφείς, λέξεις-κλειδιά και άλλα.

ElmSubmit Η μονάδα αυτή δίνει τη δυνατότητα αυτοματοποιημένης υποβολής εγγράφων μέσω ηλεκτρονικού ταχυδρομείου από αξιόπιστες πηγές.

BibEdit Η μονάδα αυτή είναι μια διεπαφή χρήστη που αποτελεί ένα πλήρες γραφικό περιβάλλον για την επεξεργασία μεταδεδομένων. Λειτουργίες όπως η προβολή του ιστορικού των αλλαγών, η προσθήκη και αφαίρεση πεδίων, το κλείδωμα της επεξεργασίας του εγγράφου από τρίτους κατά τη διάρκεια της τρέχουσας επεργασίας του, είναι επίσης διαθέσιμες μεταξύ άλλων.

Αυτοτελείς λογισμικές μονάδες σχετικές με την παροχή δεδομένων

Η παρουσίαση και παροχή των εγγράφων και των μετεδεδωμένων στο χρήστη πραγματοποιείται από διάφορες μονάδες: Η BibIndex ευρετηριάζει τα δεδομένα, η BibRank τα κατατάσσει, η BibFormat τα μορφοποιεί για την τελική προβολή στο χρήστη ενώ η WebSearch παρέχει τη δυνατότητα της εκτενούς αναζήτησης τους μέσω μιας διεπαφής χρήστη. Στη συνέχεια αναφερόμαστε ξεχωριστά σε κάθε μία από τις μονάδες αυτές.

BibIndex Η μονάδα αυτή αναλαμβάνει την ευρετηρίαση των μεταδεδωμένων, των αναφορών και των εγγράφων. Υπάρχουν δύο είδη ευρετηρίασης: λέξεων και φράσεων. Ο χρήστης μπορεί να ορίσει διάφορους λογικούς δείκτες ευρετηρίασης (για παράδειγμα συγγραφέων, τίτλων) και την αντιστοιχία τους στα πεδία μεταδεδωμένων του MARC21. Ένας δείκτης ευρετηρίασης αποτελείται από δύο μέρη. Πρώτον, ο δείκτης που δίνει τη λέξη (ή φράση) που βρέθηκε σε ένα συγκεκριμένο πεδίο μεταδεδωμένων μαζί με μια λίστα αναγνωριστικών των εγγράφων στα οποία βρέθηκε η συγκεκριμένη λέξη (ή φράση). Δεύτερον, ένας αντίστροφος δείκτης αναγνωριστικών εγγράφων με το σύνολο των λέξεων (ή φράσεων) προς τον πρώτο δείκτη. Η εν λόγω τεχνική επιτρέπει την αποδοτική ανανέωση μόνο των λέξεων που έχουν αλλάξει στα μεταδεδωμένα εισόδου ενός εγγράφου. Οι δείκτες ευρετηρίασης σχεδιάστηκαν με σκοπό να προσφέρουν ιδιαίτερα γρήγορη απόκριση στις αναζητήσεις των χρηστών, ταχύτερη από αυτή των δεικτών της βάσης δεδομένων MySQL.

BibRank Η μονάδα αυτή επιτρέπει την δημιουργία διαφόρων κριτηρίων κατάταξης που χρησιμοποιούνται στο στάδιο της αναζήτησης από την αντίστοιχη μηχανή. Για παράδειγμα, κατάταξη σύμφωνα με τη συχνότητα εμφάνισης μια λέξης, ή σύμφωνα με την τιμή ενός πεδίου μεταδεδωμένων όπως το βαθμό απήχησης ενός περιοδικού, ή ακόμη και σύμφωνα με τον αριθμό μεταφορτώσεων ενός εγγράφου.

BibClassify Η μονάδα αυτή δίνει τη δυνατότητα αυτόματης εξόρυξης λέξεων κλειδιών από έγγραφα, με βάση τη συχνότητα εμφάνισης ειδικών όρων που λαμβάνονται από ένα ελεγχόμενο λεξιλόγιο. Ένα ελεγχόμενο λεξιλόγιο μπορεί να εκφραστεί ως ένα απλό λεξικό αποθησαύρισης ή ως μια δομημένη ταξινόμηση σύμφωνα με το Resource Description Framework (RDF) ώστε να καταστεί δυνατή η σημασιολογική ταξινόμηση.

BibFormat Η μονάδα αυτή είναι υπεύθυνη για τη μορφοποίηση των βιβλιογραφικών μεταδεδωμένων με τους διάφορους διαθέσιμους τρόπους. Με άλλα λόγια επιτρέπει τον πλήρη διαχωρισμό της διαχείρισης των δεδομένων αυτών καθ' αυτών και της διαρρύθμισης της μορφοποίησής τους. Η λειτουργία της BibFormat μπορεί να εκτελεσθεί όταν ζητηθεί για την μορφοποίηση ενός εγγράφου ή ακόμη και να προεκτελεστεί και να αποθηκεύσει ένα συχνά ζητούμενο είδος μορφοποίησης για ένα ή και περισσότερα έγγραφα (όπως για παράδειγμα την προεπιλεγμένη διάταξη μορφοποίησης που χρησιμοποιείται για την παρουσίαση αποτελεσμάτων αναζήτησης).

WebSearch Η μονάδα αυτή χειρίζεται πρακτικά τα αιτήματα των χρηστών για την αναζήτηση συγκεκριμένων λέξεων ή φράσεων στη βάση δεδομένων. Υπάρχουν δύο είδη αναζήτησης: λέξεις και φράσεις. Επιτρέπεται η αναζήτηση πολύπλοκων boolean ερωτημάτων, αναζήτηση κανονικών εκφράσεων, ακόμη και η συνδυασμένη αναζήτηση μεταδεδομένων, αναφορών και εγγράφων. Αν δεν είναι δυνατή η εύρεση αποτελεσμάτων για το ακριβές αίτημα του χρήστη, το σύστημα προτείνει εναλλακτικά αποτελέσματα ως οδηγό αναζήτησης για το χρήστη. Οι δείκτες ευρετηρίασης της αναζήτησης σχεδιάστηκαν έτσι ώστε να παρέχουν ιδιαίτερα γρήγορες αποκρίσεις για συλλογές δεδομένων μεσαίου μεγέθους (έως 106 εγγραφών). Το σύνολο των μεταδεδομένων είναι οργανωμένο σε συλλογές μεταδεδομένων οι οποίες είναι άμεσα προσβάσιμες μέσω της λειτουργίας περιήγησης, ακολουθώντας τη λογική των δημοφιλών καταλόγων του διαδικτύου. Μια διαφορετική οργάνωση των συλλογών των εγγράφων μπορεί να είναι διαθέσιμη στο χρήστη μέσω των εικονικών συλλογών. Για παράδειγμα ένα έγγραφο μπορεί να ταξινομηθεί σύμφωνα με το είδος τους (βιβλίο, άρθρο) αλλά και σύμφωνα με την ημερομηνία έκδοσης του. Αυτή η ευέλικτη οργάνωση των εγγράφων σε συλλογές επιτρέπει την εύκολη πλοήγηση του χρήστη και τη δυνατότητα αναζήτησης περιεχομένου στο σύστημα.

WebStat Η μονάδα αυτή δίνει τη δυνατότητα συγκέντρωσης στατιστικών στοιχείων για την κατάσταση του εξυπηρετητή, τη χρήση του συστήματος, καθώς και για άλλα χαρακτηριστικά του συστήματος.

Αυτοτελείς λογισμικές μονάδες σχετικές με την εξατομίκευση

Το CDS Invenio μπορεί να εξατομικευθεί ώστε να ανταποκρίνεται στις ανάγκες κάθε τελικού χρήστη. Η μονάδα WebSession αναγνωρίζει τους χρήστες και τις προσωπικές τους ρυθμίσεις, η WebBasket τους δίνει τη δυνατότητα να ορίσουν και να μοιραστούν προσωπικές συλλογές εγγράφων (καλάθια), ενώ η WebAlert επιτρέπει τη δημιουργία και ρύθμιση προσωπικών ειδοποιήσεων ηλεκτρονικού ταχυδρομείου.

WebSession Η μονάδα αυτή αναλαμβάνει τη διαχείριση των χρηστών και των προσωπικών τους ρυθμίσεων κατά την περίοδο που βρίσκονται συνδεδεμένοι στο σύστημα. Φροντίζει τη διαφοροποίηση των χρηστών μεταξύ τους και χρησιμεύει στην εξατομίκευση της διεπαφής του κάθε χρήστη και στη χρήση των υπολοίπων υπηρεσιών εξατομίκευσης.

WebBasket Η μονάδα αυτή δίνει τη δυνατότητα στο χρήστη να δημιουργεί και να διαχειρίζεται προσωπικές συλλογές εγγράφων που ονομάζονται καλάθια. Ο κάθε χρήστης μπορεί να δημιουργήσει και να κατέχει πολλά καλάθια στα οποία μπορεί να αποθηκεύσει έγγραφα του ενδιαφέροντος του. Το κάθε καλάθι μπορεί να είναι προσωπικό και περιορισμένο, ή να μοιράζεται με άλλους χρήστες μέλη κάποιας ομάδας, ή και με όλους τους χρήστες του συστήματος.

WebAlert Η μονάδα αυτή επιτρέπει στο χρήστη να λαμβάνει ειδοποιήσεις ηλεκτρονικού ταχυδρομείου κάθε φορά που καταχωρείται στο σύστημα ένα έγγραφο που ταιριάζει με τα

κριτήρια τα οποία έχει ορίσει. Τα κριτήρια αυτά ανταποκρίνονται σε ένα τυπικό αίτημα αναζήτησης του χρήστη. Για παράδειγμα ο χρήστης μπορεί να ειδοποιείται κάθε φορά που ένα έγγραφο που περιέχει κάποιες συγκεκριμένες λέξεις εισάγεται στο σύστημα. Κάθε χρήστης μπορεί να ορίσει διάφορες ειδοποιήσεις, των οποίων η συχνότητα μπορεί να είναι είτε ημερήσια, είτε εβδομαδιαία, είτε μηνιαία. Ο χρήστης ενημερώνεται μέσω ηλεκτρονικού ταχυδρομείου για τα αποτελέσματα της αναζήτησης που πραγματοποίησε μια ειδοποίηση, μπορεί όμως επίσης να ορίσει κατ' επιλογήν να αποθηκεύονται αυτόματα αυτά τα δεδομένα σε ένα από τα καλάθια που του ανήκουν.

WebComment Η μονάδα αυτή αποτελεί ένα εργαλείο της κοινότητας των χρηστών του συστήματος για την ταξινόμηση των εγγράφων και την ανταλλαγή σχολίων πάνω σε αυτά.

WebMessage Η μονάδα αυτή δίνει τη δυνατότητα στους χρήστες να επικοινωνούν μεταξύ τους ανταλλάσσοντας μηνύματα καθώς και προσκλήσεις για τη συμμετοχή σε ομάδες, κ.α.

Αυτοτελείς λογισμικές μονάδες σχετικές με την ενοποίηση του συστήματος

Κάποιες μονάδες εξυπηρετούν την ενοποίηση όλων των προαναφερθεισών μονάδων με σκοπό την ολοκληρωμένη λειτουργία του συστήματος. Η μονάδα BibSched επιτρέπει τον προγραμματισμό και τη διαχείριση βιβλιογραφικών εργασιών, η WebAccess τον προσδιορισμό ρόλων, επίσης τη λειτουργία ενός συστήματος πρόσβασης στις υπηρεσίες που προσφέρει το CDS Invenio. Η WebStyle επιτρέπει τον ορισμό μια κοινής αισθητικής για το σύστημα.

BibSched Η μονάδα αυτή αποτελεί τον προγραμματιστή και διαχειριστή βιβλιογραφικών εργασιών του συστήματος. Είναι κεντρική μονάδα και επιτρέπει σε όλες τις υπόλοιπες μονάδες να έχουν πρόσβαση στα βιβλιογραφικά δεδομένα του εξυπηρετητή με προγραμματισμένο τρόπο, εμποδίζοντας την από κοινού πρόσβαση σε δεδομένα όταν αυτό δεν επιτρέπεται και φροντίζοντας τη συνεπή εκτέλεση των εργασιών ενημέρωσης της βάσης δεδομένων. Η διεπαφή διαχείρισης του προγραμματιστή εργασιών δίνει τη δυνατότητα ελέγχου και παρακολούθησης της ουράς/ακολουθίας εργασιών και παρέχει διάφορες επιλογές ανθρώπινης παρέμβασης, όπως για παράδειγμα την παύση της ουράς, την αναδιάταξή της, τον επαναπρογραμματισμό κάποιων εργασιών κ.α.

WebAccess Η μονάδα αυτή είναι υπεύθυνη για τη παροχή πρόσβασης στους χρήστες στις διάφορες λειτουργίες του συστήματος. Χρησιμοποιείται μια τεχνική ελέγχου πρόσβασης ρόλων (Role-Based Access Control RBAC) στην οποία κάθε χρήστης ανήκει σε διάφορες ομάδες ανάλογα με το ρόλο του στο σύστημα. Σε κάθε ομάδα μπορεί να χορηγηθεί το δικαίωμα να εκτελέσει μια λειτουργία ανάλογα με την εκάστοτε χρήση της λειτουργίας αυτής. Επι του παρόντος η μονάδα WebAccess χρησιμοποιείται κυρίως για τη διεπαφή διαχείρισης του συστήματος ελέγχοντας τη διαμόρφωση των διαχειριστικών μονάδων και την εκτέλεση των

σχετικών εργασιών.

WebStyle Η μονάδα αυτή ελέγχει το σύνολο των ρυθμίσεων που καθορίζουν την αισθητική του συστήματος.

3.5 Τεχνολογίες σχετικές με το CDS

Παρακάτω περιγράφονται οι κύριες τεχνολογίες στις οποίες βασίζεται και τις οποίες χρησιμοποιεί το CDS Invenio

3.5.1 GNU/Linux

Μια εγκατάσταση του CDS Invenio βασίζεται στο λειτουργικό σύστημα GNU/Linux. Πρόκειται για ένα λειτουργικό σύστημα σύμφωνο με τις αρχές του συστήματος Unix, βασισμένο στο πυρήνα Λινυξ, που χρησιμοποιεί λογισμικό GNU. Κατά συνέπεια, είναι ένα λειτουργικό σύστημα εξ ολοκλήρου βασισμένο στο ελεύθερο λογισμικό και λογισμικό ανοιχτού κώδικα. Ιστορικά χρησιμοποιείται κυρίως σε εξυπηρετητές λόγω της σταθερότητας και ασφάλειας που προσφέρει, ενώ η φιλοσοφία του, που επιτρέπει σε κάθε χρήστη να συμμετέχει στην ανάπτυξή του, έχει συμβάλλει ιδιαίτερα στην κατεύθυνση αυτή. Στην περίπτωση του CDS Invenio το GNU/Linux προσφέρει το περιβάλλον και όλα εκείνα τα απαραίτητα εργαλεία για την ανάπτυξη και προγραμματισμό του λογισμικού καθώς και την λειτουργία του ως εξυπηρετητή μιας ψηφιακής βιβλιοθήκης.

3.5.2 Python

Η Python είναι μια γλώσσα προγραμματισμού υψηλού επιπέδου γενικής χρήσης. Η βασική της φιλοσοφία είναι η αναγνωσιμότητα σε συνδυασμό με τη δυναμικότητα, και η βασική της βιβλιοθήκη είναι εκτενής και συνδυάζει αυτά τα χαρακτηριστικά. Η Python υποστηρίζει διάφορα προγραμματιστικά παραδείγματα, όπως το αντικειμενοστραφές, το προστακτικό και το συναρτησιακό. Χαρακτηριστικό της είναι η αυτόματη διαχείριση μνήμης και ως δυναμική γλώσσα προγραμματισμού μπορεί να χρησιμοποιηθεί και ως γλώσσα δέσμης ενεργειών. Λόγω της φύσης της αυτής, είναι κατάλληλη μεταξύ άλλων και για διαδικτυακές εφαρμογές. Το σύνολο αυτών των χαρακτηριστικών κάνουν τη γλώσσα προγραμματισμού Python να αποτελεί μια εξαιρετική επιλογή για την ανάπτυξη και προγραμματισμό του CDS.

3.5.3 MySQL

Το MySQL είναι ένα σύστημα διαχείρισης σχεσιακών βάσεων δεδομένων. Μέσα από αυτό ο χρήστης μπορεί να ορίσει και να διαχειριστεί ένα πλήθος βάσεων δεδομένων, και να προσφέρει πρόσβαση σε αυτές. Πρόκειται για ελεύθερο λογισμικό, το οποίο διατίθεται βάσει της άδειας GPL. Είναι ένα ιδιαίτερα δημοφιλές σύστημα, και έχει μεγάλη απήχηση στις διαδικτυακές εφαρμογές, συνήθως ως κομμάτι του μοντέλου ανάπτυξης LAMP (Linux Apache MySQL PHP/Python/Perl). Προσφέρει διεπαφές προγραμματισμού εφαρμογών για διάφορες γλώσσες προγραμματισμού και μερικά από τα χαρακτηριστικά του είναι τα παρακάτω : εκτενής υποστήριξη της γλώσσας SQL για την υποβολή ερωτημάτων στη βάση δεδομένων, υποστήριξη διάφορων υπολογιστικών πλατφορμών, αποθηκευμένες διαδικασίες, εναύσματα, δρομείς, ενημερώσιμες όψεις, μερική υποστήριξη του διεθνούς προτύπου κωδικοποίησης Ύνιςοδε, και άλλα. Το CDS Invenio χρησιμοποιεί το σύστημα MySQL ήδη εδώ και δέκα χρόνια με επιτυχία, καθώς αποτελεί ένα παράγοντα σταθερότητας και ταχύτητας στην αποθήκευση και διάθεση των δεδομένων.

3.5.4 XML / MARCXML

Η XML είναι ουσιαστικά μια γλώσσα / σύνολο κανόνων κωδικοποίησης κειμένων ψηφιακά. Τα κύρια χαρακτηριστικά της είναι η απλότητα, η γενικότητα και η ευκολία χρήσης στο διαδίκτυο. Υποστηρίζει πλήρως το διεθνές πρότυπο κωδικοποίησης Ύνιςοδε, και προορίζεται κυρίως για κείμενα, αλλά χρησιμοποιείται ευρέως και για την απεικόνιση άλλων δομών δεδομένων. Στο διαδίκτυο συχνά χρησιμοποιείται για την ανταλλαγή δεδομένων. Υπάρχουν πολλές προγραμματιστικές διεπαφές για την επεξεργασία και ανάλυση της ως το μέσο επεξεργασίας κειμένων. Μια χρήση της είναι στην επιστήμη της πληροφορίας και των βιβλιοθηκών. Η δομή MARCXML, βασισμένη στο πρότυπο MARC21, υπογορεύει τη σύνταξη βιβλιογραφικών μεταδεδομένων στη γλώσσα ΞΜΑ. Το CDS Invenio χρησιμοποιεί τη δομή MARCXML για τη διαχείριση, ανταλλαγή και μετατροπή μεταδεδομένων, εκμεταλλευόμενο τις δυνατότητες που προσφέρει η XML.

3.5.5 OAI / OAI-PMH

Το Open Archives Initiative (OAI) αποτελεί μια προσπάθεια προώθησης προτύπων διαλειτουργικότητας που αποσκοπούν στη διευκόλυνση της αποτελεσματικής διάδοσης του περιεχομένου. Έχει τις ρίζες του στο κίνημα της ελεύθερης πρόσβασης και των ακαδημαϊκών και επιστημονικών αποθετηρίων. Η συνέχιση της υποστήριξης αυτής της προσπάθειας εξακολουθεί να αποτελεί τον ακρογωνιαίο λίθο του Open Archives Initiative. Σήμερα, ωστόσο, το έργο του OAI έχει επεκταθεί για να προωθήσει την ευρεία πρόσβαση σε ψηφιακούς πόρους. Ένα από τα κύρια έργα της πρωτοβουλίας είναι το πρωτόκολλο για τη συγκομιδή μεταδεδομένων,

Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), που αποτελεί ένα μηχανισμό διαλειτουργικότητας των αποθετηρίων. Οι πάροχοι μεταδεδομένων μπορούν αρχικά να εκθέσουν και να διαθέσουν μεταδεδομένα μέσω των κανόνων που ορίζει το πρωτόκολλο, ενώ οι πάροχοι υπηρεσιών στη συνέχεια μπορούν να πραγματοποιήσουν αιτήσεις συγκομιδής τους, ώστε να διαθέσουν τελικά ένα σύνολο μεταδεδομένων προερχόμενα από διάφορους παρόχους. Οι αιτήσεις αυτές γίνονται στη γλώσσα XML με βάση το πρωτόκολλο επικοινωνίας HTTP. Το CDS Invenio δύναται να χρησιμοποιήσει το πρωτόκολλο ώστε να εκθέσει το σύνολο των μεταδεδομένων των εγγραφών μιας εγκατάστασής του.

Κεφάλαιο 4

Τι είναι μηχανική μάθηση

4.1 Εισαγωγή

Η Μηχανική μάθηση (machine learning) αποτελεί υποκατηγορία της επιστήμης των υπολογιστών και αναπτύχθηκε από τη μελέτη της αναγνώρισης προτύπων και της υπολογιστικής θεωρίας στην τεχνητή νοημοσύνη. Σύμφωνα με τον Arthur Samuel (1959), η μηχανική μάθηση ορίζεται ως το «πεδίο μελέτης που δίνει στους υπολογιστές την ικανότητα να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί». Η μηχανική μάθηση διερευνά τη μελέτη και την κατασκευή αλγορίθμων που μπορούν να μαθαίνουν από τα δεδομένα και να κάνουν προβλέψεις σχετικά με αυτά. Η λειτουργία τέτοιων αλγορίθμων βασίζεται στην κατασκευή μοντέλων από πειραματικά δεδομένα, με στόχο τις προβλέψεις που βασίζονται στα δεδομένα ή την εξαγωγή αποφάσεων ως το αποτέλεσμα αυτή της διαδικασίας (αντί να ακολουθούν στατικά τις οδηγίες του προγράμματος).

Η μηχανική μάθηση είναι στενά συνδεδεμένη και συχνά συγχέεται με την υπολογιστική στατιστική, έναν κλάδο που επίσης επικεντρώνεται στην πρόβλεψη μέσω της χρήσης των υπολογιστών. Έχει ισχυρούς δεσμούς με την μαθηματική βελτιστοποίηση, η οποία παρέχει μεθόδους, θεωρίες και τομείς εφαρμογής στο συγκεκριμένο πεδίο. Ο βασικός τομέας εφαρμογής της μηχανικής μάθησης είναι μια σειρά από υπολογιστικές εργασίες, όπου τόσο ο σχεδιασμός όσο και ο ρητός προγραμματισμός των αλγορίθμων είναι ανέφικτος. Χαρακτηριστικά παραδείγματα εφαρμογών αποτελούν οι μηχανές αναζήτησης, η υπολογιστική όραση, τα φίλτρα spam (spam filtering) και η οπτική αναγνώριση χαρακτήρων (OCR). Η μηχανική μάθηση μερικές φορές συγχέεται με την εξόρυξη δεδομένων, η οποία επικεντρώνεται περισσότερο στην εξερευνητική ανάλυση των δεδομένων, γνωστή και ως μη επιτηρούμενη μάθηση.

Στο πεδίο της ανάλυσης δεδομένων, η μηχανική μάθηση είναι μια μέθοδος που χρησιμοποιείται για την επινόηση πολύπλοκων μοντέλων και αλγορίθμων που οδηγούν στην πρόβλεψη. Τα αναλυτικά μοντέλα επιτρέπουν στους ερευνητές, στους επιστήμονες δεδομένων, στους μη-

χανικούς και στους αναλυτές να παράγουν αξιόπιστες αποφάσεις και αποτελέσματα και να αναδεικνύουν συσχετίσεις μέσω της μάθησης από ιστορικές σχέσεις και τάσεις στα δεδομένα.

Η μηχανική μάθηση χωρίζεται σε δύο κατηγορίες, το μοντέλο επαγωγικής μάθησης (inductive learning) και το μοντέλο μεταγωγικής μάθησης (transductive learning).

Στο επαγωγικό μοντέλο στόχος είναι ο υπολογισμός μιας γενικευμένης συνάρτησης, η οποία μπορεί να υπολογίζει την τιμή όλων των ορισμάτων από το πεδίο των δεδομένων. Η μεταγωγική μάθηση δέχεται σαν είσοδο δυο μεταβλητές δεδομένων, από το σύνολο της εκπαίδευσης και το σύνολο της δοκιμής, με σκοπό να υπολογίσει το αποτέλεσμα του συνόλου δοκιμής.

Η μεταγωγική μάθηση έχει ως βάση την αρχή του Vapnik, κατά την οποία, στην διαδικασία επίλυσης ενός προβλήματος, πρέπει να αποφεύγεται η λύση ενός πιο γενικού προβλήματος ως μέρος της τελικής λύσης του.

4.2 Ορισμός

Σύμφωνα με τον Tom M. Mitchell, «Ένα πρόγραμμα υπολογιστή λέγεται ότι μαθαίνει από εμπειρία E ως προς μια κλάση εργασιών T και ένα μέτρο επίδοσης P , αν η επίδοσή του σε εργασίες της κλάσης, όπως αποτιμάται από το μέτρο, βελτιώνεται με την εμπειρία E ». Αυτός ο ορισμός είναι σημαντικός για τον καθορισμό της μηχανικής μάθησης σε βασικό λειτουργικό πλαίσιο, ωστόσο ο Alan Turing στην εργασία του «Υπολογιστικές μηχανές και Νοημοσύνη», αντικαθιστά το ερώτημα που κυριαρχούσε για το αν οι μηχανές μπορούν να σκεφτούν και θέτει ως βασικό ερώτημα το αν μπορούν οι μηχανές να κάνουν αυτό που εμείς (ως σκεπτόμενες οντότητες) μπορούμε να κάνουμε τελικά. Στόχος της μηχανικής μάθησης είναι η δημιουργία συστημάτων που να μπορούν συνεχώς να βελτιώνουν την απόδοσή τους σε ένα συγκεκριμένο έργο που επιτελούν, χρησιμοποιώντας την εμπειρία που αποκομίζουν κατά την εκτέλεση της εργασίας.

4.3 Θεωρία

Ο βασικός στόχος ενός μαθητευόμενου είναι να γενικεύει από την εμπειρία του. Σε αυτό το πλαίσιο, ως γενίκευση ορίζεται η ικανότητα μιας μηχανής μάθησης να αποδίδει με ακρίβεια σε καινούργιες, πρωτόγνωρες εργασίες, αφού πρώτα έχει εκπαιδευτεί σε ένα σύνολο δεδομένων εκπαίδευσης. Τα προς εκπαίδευση παραδείγματα προέρχονται κυρίως από κάποια άγνωστη κατανομή πιθανότητας, η οποία θεωρείται αντιπροσωπευτική του χώρου των καταστάσεων, και ο μαθητής πρέπει να κατασκευάσει ένα γενικό μοντέλο για αυτό το χώρο που του επιτρέπει να παράγει αρκετά ακριβείς προβλέψεις σε νέες περιπτώσεις.

Η υπολογιστική ανάλυση των αλγορίθμων των μηχανών μάθησης και η απόδοσή τους είναι ένας κλάδος της θεωρητικής πληροφορικής γνωστός ως Υπολογιστική θεωρία μάθησης. Λόγω του πεπερασμένου είδους των εκπαιδευτικών συνόλων και της αβεβαιότητας του μέλλοντος, η θεωρία μάθησης δεν εγγυάται πάντα την απόδοση των αλγορίθμων. Αντί αυτού είναι συχνή η χρήση των πιθανό-θεωρητικών ορίων της απόδοσης.

Το πόσο καλά ένα μοντέλο που έχει εκπαιδευτεί σε υπαρκτά παραδείγματα μπορεί να προβλέψει άγνωστες καταστάσεις, ονομάζεται γενίκευση. Για την καλύτερη δυνατή γενίκευση, η πολυπλοκότητα της υπόθεσης θα πρέπει να είναι αντίστοιχη της πολυπλοκότητας της συνάρτησης των δεδομένων. Εάν η πολυπλοκότητα του μοντέλου αυξάνεται σε απόκριση, τότε το σφάλμα εκπαίδευσης μειώνεται. Ωστόσο, αν η υπόθεση είναι πολύ περίπλοκη, τότε το μοντέλο υπόκειται σε υπέρ-προσαρμογή και η γενίκευση θα είναι φτωχότερη. Σε αυτή την περίπτωση, θα πρέπει στη συνέχεια, να βρεθεί η υπόθεση που έχει το ελάχιστο σφάλμα εκπαίδευσης.

Πέρα όμως από την απόδοση, οι θεωρητικοί της υπολογιστικής μάθησης μελετούν την χρονική πολυπλοκότητα καθώς και τη σκοπιμότητα της μάθησης. Στην υπολογιστική θεωρία μάθησης ένας υπολογισμός θεωρείται εφικτός αν μπορεί να υλοποιηθεί σε πολυωνυμικό χρόνο. Υπάρχουν δύο είδη αποτελεσμάτων αναφορικά με την χρονική πολυπλοκότητα. Τα θετικά αποτελέσματα, τα οποία υποδηλώνουν ότι μια συγκεκριμένη κλάση αντιστοιχίσεων μπορεί να επιτευχθεί σε πολυωνυμικό χρόνο και τα αρνητικά αποτελέσματα που δείχνουν ότι ορισμένες τάξεις δεν μπορούν να μαθευτούν σε πολυωνυμικό χρόνο.

4.4 Μηχανική μάθηση και στατιστική

Η μηχανική μάθηση και η στατιστική συμπερασματολογία είναι δύο στενά συνδεδεμένοι επιστημονικοί τομείς. Σύμφωνα με τον Michael Jordan, οι ιδέες της μηχανικής μάθησης, από τις μεθοδολογικές αρχές μέχρι τα θεωρητικά εργαλεία, προϋπάρχουν στην στατιστική. Ο ίδιος ήταν αυτός που πρότεινε και τον όρο «Επιστήμη των Δεδομένων» για το συνολικό πεδίο.

Ο Leo Breiman διέκρινε δύο υποδείγματα στατιστικής μοντελοποίησης: το μοντέλο δεδομένων και το αλγοριθμικό μοντέλο. Το αλγοριθμικό μοντέλο ταυτίζεται σχεδόν με τους αλγορίθμους μηχανικής μάθησης όπως τα Τυχαία Δάση (Random Forest). Τέλος, ορισμένοι στατιστικολόγοι υιοθετούν μεθόδους μηχανικής μάθησης με αποτέλεσμα την δημιουργία ενός συνδυασμένου τομέα που ονομάζεται στατιστική μάθηση. Οι δύο βασικοί στόχοι της μηχανικής μάθησης έγκεινται στην:

A. Κατανόηση της 'Φύσης των Πληροφοριών / Μάθησης'

- Πώς λειτουργεί ο εγκέφαλος
- Βασικά δομικά στοιχεία τα οποία μας κάνουν έξυπνους

- Μπορούν οι μηχανές να δουν, να ακούσουν, και να ενεργούν σαν τους ανθρώπους;
- Μπορούν οι μηχανές να καταλάβουν, να σκεφτούν, να δημιουργήσουν όπως οι άνθρωποι;

B. Λήψη αποφάσεων από τα δεδομένα

- Κατανοώ – Τη φύση των δεδομένων
- Μαθαίνω από τα ιστορικά δεδομένα από - Αιτία και αποτέλεσμα
- Προβλέπω το μέλλον - Παρατήρηση για πρόβλεψη
- Μετράω επιπτώσεις - Στατιστικά και επιχειρηματικά εργαλεία μέτρησης
- Προσαρμόζω – Όπως αλλαγές των δεδομένων, διαφορετικά πεδία, κλπ.

4.5 Ιστορική αναδρομή

Από την απαρχή της επιστήμης, της τεχνολογίας και της τεχνητής νοημοσύνης, οι επιστήμονες που ακολούθησαν τον Blaise Pascal και τον Von Leibniz προσπαθούν να δημιουργήσουν ένα μηχάνημα που είναι διανοητικά ικανό όσο ένας άνθρωπος. Διάσημοι συγγραφείς όπως ο Jules Verne, Marry Shelly (Frankenstein), George Lucas (Star Wars) ονειρεύτηκαν τεχνητά όντα με συμπεριφορές παρόμοιες με του ανθρώπου ή ακόμη περισσότερο, με υιοθέτηση ανθρωποποιημένων δεξιοτήτων σε διαφορετικά περιβάλλοντα.

Το πρώτο βήμα προς την επικρατούσα μηχανική μάθηση προτάθηκε από τον Hebb το 1949, με βάση ένα σκεύασμα νευρο-ψυχολογικής μάθησης το οποίο καλείται Hebbian Learning theory. Με μια απλή εξήγηση, επιδιώκει συσχετίσεις μεταξύ των κόμβων ενός Περιοδικού Νευρωνικού Δικτύου και απομνημονεύει τυχόν κοινά σημεία στο δίκτυο με αποτέλεσμα να λειτουργεί έπειτα σαν μνήμη.

Το 1952, ο Arthur Samuel στην IBM, ανέπτυξε ένα πρόγραμμα που παίζει Ντάμα. Το πρόγραμμα ήταν σε θέση να παρατηρεί τις θέσεις και να μαθαίνει ένα μοντέλο που δίνει καλύτερες κινήσεις για τις τελευταίες περιπτώσεις. Ο Samuel έπαιξε τόσα πολλά παιχνίδια με το πρόγραμμα και παρατήρησε ότι το πρόγραμμα ήταν σε θέση να παίζει καλύτερα όσο περνούσε ο χρόνος. Με αυτό το πρόγραμμα ο Samuel αναίρεσε την γενική ομολογία που υπαγόρευε ότι οι μηχανές δεν μπορούν να υπερβούν τον ήδη γραμμένο κώδικα και να μάθουν τα πρότυπα όπως τα ανθρώπινα όντα.

Το 1957, το Perceptron του Rosenblatt ήταν το δεύτερο μοντέλο που προτάθηκε και πάλι με νευρο-επιστημονικό υπόβαθρο και είναι πιο κοντά στα σημερινά μοντέλα μηχανικής μάθησης. Ήταν μια πολύ συναρπαστική ανακάλυψη εκείνη την εποχή και ήταν πιο εφαρμόσιμη

στην πράξη από την ιδέα του Hebb. Ο Rosenblatt εισήγαγε το Perceptron με τις ακόλουθες γραμμές:

‘Το Perceptron έχει σχεδιαστεί για να απεικονίζει μερικές από τις θεμελιώδεις ιδιότητες των ευφυών συστημάτων σε γενικές γραμμές, χωρίς να εισέρχεται σε εξειδικευμένες και άγνωστες συνθήκες που ισχύουν για συγκεκριμένους βιολογικούς οργανισμούς’.

Τρία χρόνια αργότερα, ο Widrow εισήγαγε τον κανόνα μάθησης Δέλτα (Delta Learning) - γνωστό και ως πρόβλημα Ελαχίστων Τετραγώνων - ο οποίος εν συνεχεία χρησιμοποιήθηκε ως πρακτική διαδικασία για την εκπαίδευση του Perceptron. Ο συνδυασμός αυτών των δύο ιδεών δημιουργεί έναν καλό γραμμικό ταξινομητή. Το 1969, ο Minsky πρότεινε το περίφημο πρόβλημα XOR και επεσήμανε την ανικανότητα του Perceptron σε τέτοιου είδους γραμμικά αδιαχώριστες κατανομές δεδομένων.

Στη συνέχεια ακολούθησε η εισαγωγή του Multi-Layer Perceptron (MLP) από τον Werbos το 1981 με συγκεκριμένο αλγόριθμο ανάστροφης διάδοσης BP (δεδομένης της υπάρχουσας ιδέας BP η οποία είχε προταθεί το 1970 από Linnainmaa με το όνομα «αντίστροφη λειτουργία αυτόματης διαφοροποίησης»). Ακόμα και σήμερα η BP είναι το βασικό συστατικό των σημερινών αρχιτεκτονικών. Με αυτές τις νέες ιδέες, οι έρευνες επιταχύνθηκαν και πάλι. Από το 1985 έως και το 1986, οι NN ερευνητές παρουσίασαν διαδοχικά την ιδέα του MLP σε συνδυασμό με την πρακτική BP εκπαίδευση (Rumelhart, Hinton, Williams - HetCH, Nielsen).

Σε ένα άλλο φάσμα, ένας πολύ γνωστός αλγόριθμος μηχανικής μάθησης, τα Δέντρα Απόφασης και κυρίως ο ID3, προτάθηκε από τον Θ. R. Quinlan το 1986. Επιπλέον, ο ID3 θεωρήθηκε και ως λογισμικό ικανό να βρει περισσότερες εφαρμογές και χρήσεις στην πραγματική ζωή με τους απλοϊκούς του κανόνες και τη σαφή συναγωγή του (σε αντίθεση με το μαύρο κουτί των μοντέλων NN).

Μετά τον ID3, πολλές διαφορετικές εναλλακτικές λύσεις ή βελτιώσεις έχουν διερευνηθεί από την κοινότητα (π.χ. ID4, Δέντρα παλινδρόμησης, CART) ενώ εξακολουθεί να αποτελεί ένα από τα ενεργά θέματα στην μηχανική μάθηση. Η μηχανική μάθηση έχει αρχίσει να αναδιοργανώνεται πλέον ως ένα ξεχωριστό πεδίο.

Μια από τις πιο σημαντικές επαναστάσεις στη μηχανική μάθηση ήταν τα SVM (Support Vector Machines) Δίκτυα, που προτείνει ο Vapnik και ο Cortes το 1995 με πολύ ισχυρό θεωρητικό υπόβαθρο και εμπειρικά αποτελέσματα. Αυτό το γεγονός χώρισε τη κοινότητα της μηχανικής μάθησης σε δύο ομάδες, των και των SVM υποστηρικτών. Ωστόσο, ο ανταγωνισμός μεταξύ των δύο κοινοτήτων δεν ήταν πάντα τόσο εύκολος για την πλευρά των NN, ειδικά μετά την έκδοση Kernelized των SVM κοντά στη δεκαετία του 2000. Τα SVM ήταν σε θέση να εκμεταλλευτούν όλη την βαθιά γνώση της κυρτής βελτιστοποίησης, τη θεωρία γενίκευσης περιθωρίου και τους πυρήνες (kernels) σε αντίθεση με τα μοντέλα NN. Ως εκ τούτου, υπήρξε μια μεγάλη ώθηση από διαφορετικούς επιστημονικούς κλάδους προκαλώντας

μια πολύ γρήγορη θεωρητική και πρακτική βελτίωση.

Τα NN δίκτυα επηρεάστηκαν αρνητικά από το έργο και τη διατριβή του Hochreiter το 1991. Έπειτα, ένα άλλο μοντέλο μηχανικής μάθησης, το adaboost, προτάθηκε από τους Freund και Schapire το 1997 για το οποίο κέρδισαν και το βραβείο Godel. Το Adaboost εκπαιδεύει αδύναμα σύνολα ταξινομητών που είναι εύκολο να εκπαιδευτούν, δίνοντας μεγαλύτερη σημασία σε δύσκολες περιπτώσεις. Αυτό το μοντέλο εξακολουθεί να είναι η βάση για πολλές διαφορετικές εργασίες, όπως η αναγνώριση και ανίχνευση προσώπου, αποτελώντας παράλληλα και μια υλοποίηση της PAC (Probably Approximately Correct) θεωρίας μάθησης. Γενικά, οι αδύναμοι ταξινομητές επιλέγονται ως ενιαίοι κόμβοι στα δέντρα απόφασης.

Ακόμη ένα μοντέλο, το Random Forests (RF), εξερευνήθηκε από τον Breiman το 2001 και συλλέγει πολλαπλά δέντρα απόφασης, όπου το καθένα από αυτά απαρτίζεται από ένα τυχαίο υποσύνολο των περιπτώσεων και κάθε κόμβος επιλέγεται από ένα τυχαίο υποσύνολο των χαρακτηριστικών. Σε περιπτώσεις όπου το AdaBoost δείχνει αδυναμία στην υπέρ-τοποθέτηση και σε ακραία περιστατικά στα δεδομένα, το μοντέλο RF είναι μια πιο ισχυρή λύση ενάντια σε αυτές τις δυσκολίες.

Το Random Forests αποτελεί έναν συνδυασμό προγνωστικών δέντρων όπου κάθε δέντρο εξαρτάται από τις τιμές ενός τυχαίου ανεξάρτητου διανύσματος και όλα τα δέντρα στο δάσος έχουν την ίδια κατανομή. Το σφάλμα γενίκευσης για τα δάση συγκλίνει σε ένα όριο όσο ο αριθμός των δέντρων στο δάσος μεγαλώνει.

Όσο πλησιάζουμε το σήμερα, μια νέα εποχή των εισάγεται στο προσκήνιο, γνωστή πλέον και ως Deep Learning. Αναφέρεται σε μοντέλα νευρωνικών δικτύων (NN) με πολλά και ευρεία διαδοχικά στρώματα. Η τρίτη άνοδος των NN ξεκίνησε περίπου το 2005 με το συνδυασμό πολλών διαφορετικών ευρημάτων από το παρελθόν και το παρόν από τον Hinton, τον LeCun, τον Bengio, τον Andrew Ng και άλλους ερευνητές.

4.6 Διαφορά μεταξύ εξόρυξης δεδομένων (data mining), μηχανικής μάθησης (machine learning) και βαθιάς μάθησης deep learning

Παρά το γεγονός ότι όλες οι προαναφερθείσες μέθοδοι έχουν τον ίδιο στόχο - να εξάγουν πληροφορίες, πρότυπα και σχέσεις για την υποστήριξη της διαδικασίας λήψης αποφάσεων - έχουν διαφορετικές προσεγγίσεις και ικανότητες.

4.6.1 Εξόρυξη δεδομένων (Data Mining)

Η εξόρυξη δεδομένων μπορεί να θεωρηθεί ως ένα υπερσύνολο πολλών διαφορετικών μεθόδων για την εξαγωγή ιδεών από τα δεδομένα. Μπορεί να περιλαμβάνει παραδοσιακές στατιστικές μεθόδους και μηχανική μάθηση. Πιο συγκεκριμένα, εφαρμόζει μεθόδους από πολλές διαφορετικές περιοχές για τον εντοπισμό άγνωστων μοτίβων τα οποία προκύπτουν από τα δεδομένα. Αυτό μπορεί να περιλαμβάνει στατιστικούς αλγόριθμους, μηχανική μάθηση, ανάλυση κειμένου, ανάλυση χρονοσειρών και άλλες περιοχές της αναλυτικής. Η εξόρυξη δεδομένων περιλαμβάνει ακόμη τη μελέτη της αποθήκευσης και διαχείρισης δεδομένων.

4.6.2 Μηχανική μάθηση (Machine Learning)

Η κύρια διαφορά της μηχανικής μάθησης είναι ότι ακριβώς όπως τα στατιστικά μοντέλα, στοχεύει στη κατανόηση της δομής των δεδομένων - ταιριάζει θεωρητικές κατανομές στα δεδομένα οι οποίες είναι απολύτως κατανοητές. Έτσι, στα στατιστικά μοντέλα υπάρχει μια θεωρία πίσω από το μοντέλο που είναι μαθηματικά αποδεδειγμένο, αλλά αυτό προϋποθέτει επιπλέον ότι τα δεδομένα πληρούν ορισμένες ισχυρές παραδοχές. Η μηχανική μάθηση έχει αναπτυχθεί βασιζόμενη στην ικανότητα της χρήσης υπολογιστών για να εξέταση δεδομένων για τη δημιουργία δομών, ακόμα και στις περιπτώσεις που δεν έχουμε μια θεωρία για τη μορφή της δομής. Η δοκιμή για ένα μοντέλο μηχανικής μάθησης είναι ένα σφάλμα επικύρωσης σε νέα δεδομένα, και όχι μια θεωρητική εξέταση που αποδεικνύει την μηδενική υπόθεση. Καθώς η μηχανική μάθηση συχνά χρησιμοποιεί μια επαναληπτική προσέγγιση για να «μάθει» από τα δεδομένα, η διαδικασία μπορεί εύκολα να αυτοματοποιηθεί. Στην ουσία, πραγματοποιούνται περάσματα μέσα από τα δεδομένα μέχρι να βρεθεί ένα ισχυρό πρότυπο.

4.6.3 Βαθιά μάθηση (Deep Learning)

Η Βαθιά μάθηση συνδυάζει προόδους στην υπολογιστική ισχύ και ειδικούς τύπους νευρωνικών δικτύων για να μάθει περίπλοκα σχέδια σε περιπτώσεις μεγάλου όγκου δεδομένων. Οι τεχνικές του Deep learning συνίστανται επί του παρόντος σε τεχνικές αναγνώρισης αντικειμένων σε εικόνες, λέξεις και ήχους. Οι ερευνητές ψάχνουν τώρα να εφαρμόσουν αυτές τις επιτυχημένες τεχνικές στην αναγνώριση προτύπων σε πολυπλοκότερες εργασίες, όπως η αυτόματη μετάφραση, οι ιατρικές διαγνώσεις και σε πολλά σημαντικά κοινωνικά και επιχειρηματικά προβλήματα.

4.7 Η μηχανική μάθηση στο σήμερα

Λόγω των νέων τεχνολογιών πληροφορικής, η μηχανική μάθηση σήμερα έχει εξελιχθεί σε σχέση με το παρελθόν. Γεννήθηκε από την αναγνώριση προτύπων και τη θεωρία ότι οι υπολογιστές μπορούν να μάθουν, χωρίς να έχουν προγραμματιστεί να εκτελούν συγκεκριμένες λειτουργίες και καθήκοντα· οι ερευνητές που δείχνουν ενδιαφέρον για την τεχνητή νοημοσύνη θέλουν να δουν αν οι υπολογιστές μπορούν να μάθουν από τα δεδομένα. Η επαναληπτική πτυχή της μηχανικής μάθησης είναι ιδιαίτερα σημαντική επειδή προσφέρει τη δυνατότητα στα μοντέλα να προσαρμόζονται ανεξάρτητα ενώ εκτίθενται σε νέα δεδομένα. Μαθαίνουν από προηγούμενους υπολογισμούς να παράγουν αξιόπιστες, επαναλαμβανόμενες αποφάσεις και αποτελέσματα. Αποτελεί πλέον μια επιστήμη που κερδίζει μέρα με τη μέρα νέα δυναμική.

Ενώ πολλοί αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται κατά τη διάρκεια των τελευταίων ετών, η ικανότητα να εφαρμόζονται αυτόματα πολύπλοκοι μαθηματικοί υπολογισμοί σε μεγάλο όγκο δεδομένων - ξανά και ξανά, όλο και πιο γρήγορα - είναι μια πιο πρόσφατη εξέλιξη. Στη συνέχεια παρουσιάζονται μερικά ευρέως διαδεδομένα παραδείγματα εφαρμογών μηχανικής μάθησης:

- Το αυτοκίνητο της Google που οδηγείται μόνο του (self-driving)· Η ουσία της μηχανικής μάθησης.
- Η ηλεκτρονική σύσταση προσφορών, όπως αυτές από το Amazon και το Netflix· εφαρμογές μηχανικής μάθησης στην καθημερινή ζωή.
- Η γνώση για το τι λέγεται από τους πελάτες κάθε εταιρίας στα μέσα κοινωνικής δικτύωσης (Twitter). Μηχανική μάθηση σε συνδυασμό με τη δημιουργία του γλωσσικού κανόνα.
- Ανίχνευση απάτης, μία από τις πιο προφανείς και σημαντικές χρήσεις στη σημερινή κοινωνία.

4.8 Χρησιμότητα μηχανικής μάθησης

Ο αυξανόμενος όγκος, η ποικιλία των διαθέσιμων δεδομένων, η υπολογιστική επεξεργασία η οποία είναι οικονομικότερη και πιο ισχυρή καθώς και η πιο προσιτή οικονομικά αποθήκευση δεδομένων αποτελούν βασικούς παράγοντες που έχουν αναζωπυρώσει το ενδιαφέρον για τη μηχανική μάθηση και την έχουν τοποθετήσει μαζί με την εξόρυξη δεδομένων και τη Μπεϋζιανή (Bayesian) ανάλυση στο επίκεντρο των σημερινών κοινωνιών. Γεγονός που σημαίνει, ότι καθίσταται δυνατή η γρήγορη και αυτόματη παραγωγή μοντέλων που έχουν τη δυνατότητα να αναλύουν μεγαλύτερα, πολυπλοκότερα δεδομένα και να αποδίδουν γρηγορότερα και ακριβέστερα αποτελέσματα - ακόμα και σε πολύ μεγάλη κλίμακα. Οικοδομώντας με αυτό τον

τρόπο ακριβή μοντέλα, ένας οργανισμός έχει καλύτερη πιθανότητα εντοπισμού κερδοφόρων ευκαιριών ή αποφυγή άγνωστων κινδύνων αντίστοιχα.

Οι απαιτήσεις για τη δημιουργία αποτελεσματικών συστημάτων μηχανικής μάθησης περιλαμβάνουν:

- Δυνατότητες προετοιμασίας δεδομένων.
- Αλγόριθμοι - βασικοί και προηγμένοι.
- Αυτοματοποίηση και επαναληπτικές διαδικασίες.
- Επεκτασιμότητα (Scalability).
- Συνολική μοντελοποίηση.

4.9 Χρήστες μηχανικής μάθησης

Οι περισσότερες βιομηχανίες που εργάζονται με μεγάλο όγκο δεδομένων έχουν αναγνωρίσει την αξία της τεχνολογίας της μηχανικής μάθησης. Με τη συλλογή στοιχείων που προκύπτουν από την ανάλυση των δεδομένων - συχνά σε πραγματικό χρόνο - οι οργανισμοί είναι σε θέση να εργαστούν πιο αποτελεσματικά και να αποκτήσουν ανταγωνιστικό πλεονέκτημα.

4.9.1 Χρηματοπιστωτικές υπηρεσίες (Financial services)

Οι τράπεζες και άλλες επιχειρήσεις στο χρηματοπιστωτικό κλάδο κάνουν χρήση της μηχανικής μάθησης για δύο βασικούς σκοπούς: τον εντοπισμό σημαντικών πληροφοριών από τα δεδομένα και την ανίχνευση της απάτης. Οι πληροφορίες και η γνώση που αποκτούν συμβάλλουν στην εύρεση επενδυτικών ευκαιριών ή στη παροχή βοήθειας στους επενδυτές για τις εμπορικές τους συναλλαγές. Με την εξόρυξη δεδομένων μπορούν επίσης να αναγνωριστούν οι πελάτες με προφίλ υψηλού κινδύνου, ή να ανιχνευτεί η απάτη μέσω προειδοποιητικών σημαδιών από την εποπτεία του κυβερνοχώρου.

4.9.2 Κυβέρνηση

Οι κρατικές υπηρεσίες, όπως η δημόσια ασφάλεια και οι επιχειρήσεις κοινής ωφέλειας έχουν ιδιαίτερη ανάγκη τη μηχανική μάθηση, δεδομένου ότι έχουν πολλαπλές πηγές δεδομένων οι οποίες μέσω της εξόρυξης μπορούν να παράγουν ιδέες. Αναλύοντας, για παράδειγμα, τα δεδομένα που προκύπτουν από τους αισθητήρες, προσδιορίζονται τρόποι αύξησης της αποδοτικότητας και εξοικονομούνται χρήματα. Μέσω της μηχανικής μάθησης διευκολύνεται ο εντοπισμός και η ανίχνευση απάτης καθώς και ελαχιστοποιείται ο κίνδυνος κλοπής της ταυτότητας.

4.9.3 Ιατρική περίθαλψη

Η μηχανική μάθηση είναι μια ταχέως αναπτυσσόμενη τάση στον κλάδο της υγειονομικής περίθαλψης, χάρη στην έλευση των συσκευών που μπορούν να φορεθούν από τον καθένα μας (wearable) και των αισθητήρων που μπορούν να χρησιμοποιήσουν τα δεδομένα για την αξιολόγηση της υγείας του ασθενούς σε πραγματικό χρόνο. Η τεχνολογία μπορεί επίσης να βοηθήσει ιατρικούς εμπειρογνώμονες στην ανάλυση των δεδομένων για τον προσδιορισμό των τάσεων που μπορούν να συμβάλλουν σε βελτιωμένες διαγνώσεις και θεραπείες.

4.9.4 Μάρκετινγκ και πωλήσεις

Οι ιστοσελίδες που προτείνουν αντικείμενα που μπορεί να αρέσουν στο χρήστη βασισμένες στις προηγούμενες αγορές του χρησιμοποιούν μηχανική μάθηση για να την ανάλυση του ιστορικού των αγορών του και την προώθηση άλλων αντικειμένων που ενδεχομένως τον ενδιαφέρουν. Η προαναφερθείσα ικανότητα να συλλέγονται, να αναλύονται τα δεδομένα και να χρησιμοποιούνται για να διαμορφώνεται μια προσωποποιημένη εμπειρία αγορών (ή αντίστοιχα μια εκστρατεία μάρκετινγκ) αποτελεί το μέλλον του λιανικού εμπορίου.

4.9.5 Πετρέλαιο και φυσικό αέριο

Η εύρεση νέων πηγών ενέργειας, η ανάλυση μετάλλων του εδάφους, η πρόβλεψη αποτυχίας των αισθητήρων στα διυλιστήρια, ο εξ' ορθολογισμού της διανομής πετρελαίου ώστε να καταστεί πιο αποτελεσματική και αποδοτική είναι μόνο μερικές από τις περιπτώσεις που εφαρμόζεται η μηχανική μάθηση στον κλάδο αυτό ενώ ο αριθμός τους εξακολουθεί να επεκτείνεται.

4.9.6 Μεταφορά

Η ανάλυση των δεδομένων για την αναγνώριση προτύπων και τάσεων είναι το κλειδί για την βιομηχανία μεταφορών, το οποίο στοχεύει στο σχεδιασμό αποτελεσματικότερων διαδρομών και τη πρόβλεψη πιθανών προβλημάτων για την αύξηση της κερδοφορίας. Οι πτυχές της ανάλυσης δεδομένων και της μοντελοποίησης της μηχανικής μάθησης είναι σημαντικά εργαλεία για τις εταιρείες διανομής, τα δημόσια μέσα μεταφοράς και άλλους οργανισμούς μεταφορών.

4.10 Η μηχανική μάθηση στον επιχειρηματικό κλάδο

Το 2015, οι διεκδικητές του πρωταθλήματος της Εθνικής Ομοσπονδίας Καλαθοσφαίρισης των ΗΠΑ στηρίχθηκαν στην ανάλυση της Second Spectrum, μιας start-up εταιρίας μηχανικής

μάθησης από την Καλιφόρνια. Με την ψηφιοποίηση των παιχνιδιών των τελευταίων σεζόν, δημιουργούνται μοντέλα πρόβλεψης που επιτρέπουν στον προπονητή να μπορεί να ξεχωρίσει ανάμεσα σε, όπως ανέφερε και ο Διευθύνων Σύμβουλος Rajiv Maheswaran, «ένα κακό σπότερ που ρίχνει καλές βολές και ένα καλό που ρίχνει κακές βολές» έτσι ώστε να ρυθμίσει τις αποφάσεις του ανάλογα.

Ένα άλλο παράδειγμα είναι η General Electric, η οποία βρίσκεται ακόμα στην αγορά μετά από 119 χρόνια. Η GE κερδίζει ήδη εκατοντάδες εκατομμύρια δολάρια «σπάζοντας» τα δεδομένα που συλλέγει από τις πετρελαιοπηγές βαθέων υδάτων ή τους κινητήρες των τζετ για να βελτιστοποιήσει την απόδοση, την πρόβλεψη βλαβών, και τον εξ ορθολογισμό της συντήρησής τους. Ωστόσο, ο Colin Paris, ο οποίος εντάχθηκε στην GE από την IBM το 2015 ως αντιπρόεδρος της έρευνας λογισμικού, πιστεύει ότι η συνεχής πρόοδος στην εξουσία επεξεργασίας δεδομένων, στους αισθητήρες και στους προβλεπτικούς αλγόριθμους θα δώσει σύντομα στην εταιρεία του την ίδια ευκρίνεια για τις ατομικές ιδιαιτερότητες ενός κινητήρα τζετ όση έχει και το Google για την διαδικτυακή συμπεριφορά ενός 24 χρονών νέου από Δυτικό Χόλιγουντ.

Στην Ευρώπη, πολλές τράπεζες έχουν αντικαταστήσει παλαιότερα στατιστικά μοντέλα με τεχνικές μηχανικής μάθησης και, σε ορισμένες περιπτώσεις, παρουσίασαν 10% αύξηση των πωλήσεων των νέων προϊόντων, 20% εξοικονόμηση σε κεφαλαιουχικές δαπάνες, 20% σε συλλογή μετρητών, και 20% μείωση σε πελάτες που αλλάζουν υπηρεσία. Οι τράπεζες έχουν επιτύχει αυτά τα κέρδη επινοώντας νέες συστάσεις σαν κίνητρο για τους πελάτες στο λιανικό εμπόριο, στις μικρές και στις μεσαίες επιχειρήσεις. Έχουν, επίσης, ενσωματώσει στοχευμένα μοντέλα (micro – targeted) που προβλέπουν με μεγαλύτερη ακρίβεια ποιος θα ακυρώσει την υπηρεσία ή θα αμελήσει το δάνειό του και προσπαθεί να εντοπίσει τον καλύτερο τρόπο για να παρέμβει.

Παρόμοιες ιστορίες επιτυχίας έχουν παρατηρηθεί στον τομέα της διαχείρισης ταλέντων για την πρόβλεψη, αποκλειστικά και μόνο με την εξέταση των κατατεθειμένων βιογραφικών, πιθανών υποψηφίων.

Καθώς όλο και μεγαλύτερο τμήμα του αναλογικού κόσμου ψηφιοποιείται, η ικανότητά μας να μαθαίνουμε από τα δεδομένα, με την ανάπτυξη και τη δοκιμή των αλγορίθμων γίνεται όλο και πιο σημαντικό για το τι θεωρείται σήμερα παραδοσιακή επιχείρηση. Ο επικεφαλής οικονομολόγος της Google, Hal Varian, ονόμασε την κατάσταση «computer kaizen». «Για όσο η μαζική παραγωγή άλλαξε τον τρόπο συναρμολόγησης των προϊόντων και η συνεχής βελτίωση άλλαξε τον τρόπο υλοποίησης της παραγωγής», ανέφερε, «τόσο ο συνεχής (και συχνά αυτοματοποιημένος) πειραματισμός θα βελτιώνει τον τρόπο με τον οποίο βελτιστοποιούνται οι επιχειρηματικές διαδικασίες στο οργανισμό μας».

4.11 Άνθρωποι και μηχανές

Η αλλαγή βρίσκεται στο κατώφλι μας και τα δεδομένα δημιουργούνται τόσο γρήγορα με αποτέλεσμα η συμμετοχή του ανθρώπου σε αυτό να καταστεί γρήγορα ανέφικτη. Σε χρονικό ορίζοντα τριών έως πέντε ετών, αναμένεται να παρατηρηθούν πολύ υψηλά επίπεδα τεχνητής νοημοσύνης, καθώς και μεγάλη ανάπτυξη καταναμημένων αυτόνομων επιχειρήσεων. Αυτές οι αυτό-παρακινούμενες με αυτόνομα μέσα επιχειρήσεις, θα είναι σε θέση να πραγματοποιήσουν στόχους που έχουν τεθεί αυτόνομα, χωρίς οποιαδήποτε άμεση ανθρώπινη επίβλεψη. Πολλές πτυχές θα υλοποιούνται μέσω του αυτό-προγραμματισμού.

Ωστόσο, μια πλευρά της κοινής γνώμης υποστηρίζει ότι οι καταναμημένες αυτόνομες επιχειρήσεις μπορούν να γίνουν απειλητικές και επιζήμιες για τον πολιτισμό μας. Μέχρι τη στιγμή της πλήρης άνθησης τους, η μηχανική μάθηση θα έχει γίνει πολιτισμικά αόρατη με τον ίδιο τρόπο που οι τεχνολογικές εφευρέσεις του 20ου αιώνα, εξαφανίστηκαν στο παρασκήνιο. Ο ρόλος των ανθρώπων θα είναι να κατευθυντήριος και καθοδηγητικός προς τους αλγόριθμους που θα προσπαθούν να επιτύχουν τους στόχους που τους δίνονται. Αυτό το μάθημα προέκυψε από το ρόλο που έπαιξαν οι αλγόριθμοι αυτόματων συναλλαγών που προκάλεσαν σημαντική ζημιά κατά τη διάρκεια της οικονομικής κρίσης του 2008.

Τέλος, όποιες και όσες καινούριες γνώσεις ηλεκτρονικών υπολογιστών έρθουν στην επιφάνεια, μόνο οι άνθρωποι διαχειριστές μπορούν να αποφασίσουν τις βασικές ερωτήσεις, όπως το ποια αποτελούν τα πιο κρίσιμα επιχειρηματικά προβλήματα που μια εταιρεία προσπαθεί πραγματικά να λύσει. Ακριβώς όπως οι άνθρωποι χρειάζονται τακτικές αξιολογήσεις και εκτιμήσεις, έτσι και αυτές οι «έξυπνες μηχανές» και το έργο τους, θα πρέπει να αξιολογείται τακτικά, να βελτιώνεται, και, ίσως ακόμη να αλλάζει και να ακολουθεί τελείως διαφορετικές πορείες - από στελέχη με εμπειρία, κρίση και εξειδίκευση στον εκάστοτε τομέα.

Οι νικητές δε θα είναι ούτε οι μηχανές ούτε οι άνθρωποι που θα δουλεύουν ξεχωριστά, αλλά ο συνδυασμός αυτών των δύο παραγόντων που εργάζονται μαζί αποτελεσματικά.

4.12 Τύποι προβλημάτων και εργασιών

Οι εργασίες μηχανικής μάθησης ανάλογα με τη φύση του εκπαιδευτικού σήματος ή την ανατροφοδότηση που είναι διαθέσιμα σε ένα σύστημα εκμάθησης, συνήθως ταξινομούνται σε τέσσερις μεγάλες κατηγορίες:

4.12.1 Επιτηρούμενη μάθηση (supervised learning)

Το υπολογιστικό πρόγραμμα δέχεται τις παραδειγματικές εισόδους καθώς και τα επιθυμητά αποτελέσματα από έναν «δάσκαλο», και ο στόχος του είναι να μάθει έναν γενικό κανόνα προκειμένου να αντιστοιχίσει τις εισόδους με τα αποτελέσματα.

4.12.2 Μη επιτηρούμενη μάθηση (unsupervised learning)

Χωρίς να παρέχεται κάποια εμπειρία στον αλγόριθμο μάθησης, πρέπει να βρει την δομή των δεδομένων εισόδου. Η μη επιτηρούμενη μάθηση μπορεί να είναι αυτοσκοπός (ανακαλύπτοντας κρυμμένα μοτίβα σε δεδομένα) ή μέσο για ένα τέλος (χαρακτηριστικό της μάθησης). Η μη επιτηρούμενη μάθηση αναφέρεται κυρίως στο πρόβλημα της εύρεσης της δομής σε ένα σύνολο δεδομένων, χωρίς να γνωρίζουμε τα επιθυμητά αποτελέσματα – εξόδους. Συχνά θεωρείται ότι το θεμελιώδες πρόβλημα της επιτηρούμενης μάθησης είναι ο υπολογισμός της πυκνότητας που παρήγαγε το ζητούμενο. Εφόσον για το σύνολο δεδομένων που δίνονται στον υπολογιστή σαν παράδειγμα δεν έχουμε τις σωστές εξόδους, δεν υπάρχει κριτήριο λάθους για την πιθανή λύση.

4.12.3 Ημί-επιτηρούμενη μάθηση

Η Ημί - Επιτηρούμενη Μάθηση (Semi-supervised learning – SSL), βρίσκεται σαν φιλοσοφία ανάμεσα στην επιτηρούμενη και την επιτηρούμενη μάθηση. Η Ημί-επιτηρούμενη μάθηση χρησιμοποιείται για την εκπαίδευση και των δύο τύπων δεδομένων, και αυτών που είναι κατηγοριοποιημένα (λαβελεδ δατα) και αυτών που δεν είναι (unlabeled data). Συνήθως στην ημι-επιτηρούμενη μάθηση ο όγκος των πρώτων δεδομένων είναι σημαντικά μικρότερος από τον αντίστοιχο των δεδομένων που δεν έχουν κατηγοριοποιηθεί. Έτσι συνδυάζονται οι διαφορετικές φιλοσοφίες της επιτηρούμενης και της μη επιτηρούμενης μάθησης. Σύμφωνα με τις έρευνες, ο συνδυασμός των δύο τύπων δεδομένων συνήθως έχει καλύτερα αποτελέσματα στη μηχανική μάθηση.

4.12.4 Ενισχυτική μάθηση (reinforcement learning)

Ενισχυτική μάθηση είναι μια περιοχή της μηχανικής μάθησης εμπνευσμένη από τη συμπεριφορά και ασχολείται με το πώς το λογισμικό θα πρέπει να αναλάβει δράσεις σε ένα περιβάλλον έτσι ώστε να μεγιστοποιηθεί κάποια έννοια της αθροιστικής ανταμοιβής. Το πρόβλημα, λόγω της γενικότητας του, μελετάται σε πολλούς άλλους επιστημονικούς κλάδους, όπως η θεωρία παιγνίων, η θεωρία ελέγχου, η επιχειρησιακή έρευνα, η θεωρία πληροφοριών, η βελτιστοποίηση και προσομοίωση με βάση τα συστήματα πολλαπλών πρακτόρων. Στη βιβλιογραφία εργασιών

έρευνας και ελέγχου, το πεδίο όπου μελετώνται σχετικές μέθοδοι μάθησης ονομάζεται προσέγγιση δυναμικού προγραμματισμού. Το πρόβλημα έχει μελετηθεί στη θεωρία του βέλτιστου ελέγχου, αν και οι περισσότερες μελέτες ασχολούνται με την ύπαρξη των βέλτιστων λύσεων και τον χαρακτηρισμό τους, και όχι με τις πτυχές της μάθησης ή την προσέγγιση. Στα οικονομικά και τη θεωρία παιγνίων η ενίσχυση της μάθησης μπορεί να χρησιμοποιηθεί για να εξηγήσει πώς μπορεί να προκύψει ισορροπία από τον οριοθετημένο ορθολογισμό.

Στη μηχανική μάθηση, το περιβάλλον διαμορφώνεται τυπικά ως μια διαδικασία λήψης αποφάσεων Markov (MDP), αφού πολλοί αλγόριθμοι ενισχυτικής μάθησης χρησιμοποιούν τεχνικές δυναμικού προγραμματισμού. Η κύρια διαφορά μεταξύ των κλασικών τεχνικών και των αλγορίθμων ενισχυτικής μάθησης είναι ότι οι τελευταίοι δεν χρειάζονται MDP γνώση και στοχεύουν σε μεγάλες MDPs όπου ακριβείς μέθοδοι δεν είναι δυνατόν να χρησιμοποιηθούν.

4.13 Μια διαφορετική κατηγοριοποίηση

Μια άλλη κατηγοριοποίηση των προβλημάτων μηχανικής μάθησης προκύπτει όταν κάποιος αναλογιστεί το επιθυμητό αποτέλεσμα του συστήματος μηχανικής μάθησης:

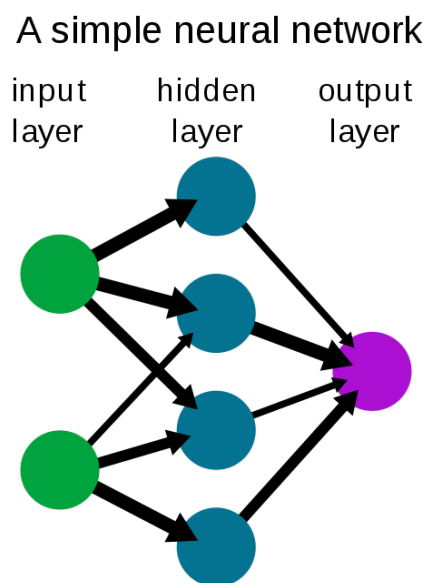
- Στην ταξινόμηση, τα δεδομένα εισόδου χωρίζονται σε δύο ή περισσότερες κλάσεις, και η μηχανή πρέπει να κατασκευάσει ένα μοντέλο, το οποίο θα αντιστοιχίζει τα δεδομένα σε μία ή περισσότερες (multi-label ταξινόμηση) κλάσεις. Αυτό συνήθως εμπίπτει στην επιτηρούμενη μάθηση. Τα φίλτρα Spam είναι ένα παράδειγμα ταξινόμησης, όπου οι εισοδοί είναι τα emails ή άλλα μηνύματα και οι κλάσεις είναι "spam" και "μη spam".
- Στην παλινδρόμηση, η οποία αποτελεί επίσης πρόβλημα επιτηρούμενης μάθησης, τα αποτελέσματα είναι συνεχή και όχι διακριτά.
- Στην συσταδοποίηση, ένα σύνολο εισόδων χωρίζεται σε ομάδες. Σε αντίθεση με την ταξινόμηση, οι ομάδες δεν είναι γνωστές εκ των προτέρων, καθιστώντας αυτόν τον διαχωρισμό τυπική εργασία μη επιτηρούμενης μάθησης.
- Η εκτίμηση πυκνότητας βρίσκει την κατανομή των δεδομένων εισόδου σε κάποιο χώρο.
- Σε προβλήματα μείωσης διαστασιμότητας (dimensionality reduction), τα δεδομένα απλοποιούνται και αντιστοιχίζονται σε ένα χώρο λιγότερων διαστάσεων. Το στατιστικό μοντέλο θεμάτων (Topic modeling) είναι ένα σχετικό πρόβλημα, όπου η μηχανή καλείται να βρει έγγραφα που καλύπτουν παρόμοια θέματα από ένα σύνολο εγγράφων γραμμένων σε φυσική γλώσσα.

Κεφάλαιο 5

Νευρωνικά δίκτυα και εφαρμογές κεφάλαιο

5.1 Εισαγωγή στα νευρωνικά δίκτυα

5.1.1 Τι είναι ένα νευρωνικό δίκτυο



Σχήμα 5.1: Παράδειγμα τεχνητού νευρωνικού δικτύου

Ένα τεχνητό νευρωνικό δίκτυο (ANN) είναι ένα παράδειγμα επεξεργασίας πληροφοριών που είναι εμπνευσμένο από τον τρόπο με τον οποίο βιολογικά νευρικά συστήματα, όπως ο εγκέφαλος, επεξεργάζονται πληροφορίες. Το βασικό στοιχείο αυτού του παραδείγματος είναι η νέα δομή του συστήματος επεξεργασίας πληροφοριών. Αποτελείται από ένα μεγάλο αριθμό

εξαιρετικά διασυνδεδεμένων στοιχείων επεξεργασίας (νευρώνες) που εργάζονται από κοινού για την επίλυση συγκεκριμένων προβλημάτων. Τα ANN, όπως και οι άνθρωποι, μαθαίνουν από τα παραδείγματα. Ένα ANN έχει διαμορφωθεί για μια συγκεκριμένη εφαρμογή, όπως η αναγνώριση προτύπων ή ταξινόμηση δεδομένων, μέσα από μια διαδικασία μάθησης. Η μάθηση σε βιολογικά συστήματα αφορά προσαρμογές στις συναπτικές συνδέσεις που υπάρχουν ανάμεσα στους νευρώνες.

5.1.2 Ιστορική αναδρομή

Οι προσομοιώσεις νευρωνικών δικτύων φαίνονται να είναι μια πρόσφατη εξέλιξη. Ωστόσο, αυτός ο τομέας ξεκίνησε πριν από την έλευση των υπολογιστών, και έχει επιζήσει τουλάχιστον ένα σημαντικό πλήγμα και διαφορετικές εποχές.

Πολλές εξελίξεις έχουν σημειωθεί χάρη στη χρήση των ανέξοδων εξομοιώσεων υπολογιστών. Μετά την αρχική περίοδο του ενθουσιασμού, ο τομέας επέζησε μια περίοδο απογοήτευσης και ανυποληψίας. Κατά τη διάρκεια αυτής της περιόδου όταν η χρηματοδότηση και η επαγγελματική υποστήριξη ήταν ελάχιστη, σημαντική πρόοδος σημειώθηκε από τους λίγους ερευνητές. Οι πρωτοπόροι κατάφεραν να αναπτύξουν πειστική τεχνολογία η οποία ξεπέρασε τους περιορισμούς που αναγνωρίστηκαν από τον Minsky και τον Papert. Οι δύο τελευταίοι εξέδωσαν ένα βιβλίο το 1969 όπου περιέγραφαν ένα γενικό αίσθημα απογοήτευσης ενάντια στα νευρωνικά δίκτυα από τους διάφορους ερευνητές της εποχής. Το βιβλίο έγινε αποδεκτό στους περισσότερους χωρίς περαιτέρω ανάλυση. Επί του παρόντος, ο τομέας των νευρωνικών δικτύων παρουσιάζει μια αναζωπύρωση του ενδιαφέροντος και αύξηση χρηματοδοτήσεων.

Ο πρώτος τεχνητός νευρώνας παρήχθη το 1943 από τον νευροφυσιολόγο Warren McCulloch και τον προγραμματιστή Walter Pitts. Δυστυχώς η τεχνολογία που ήταν διαθέσιμη εκείνη την εποχή δεν τους επέτρεψε να προχωρήσουν ιδιαίτερα.

5.1.3 Γιατί να χρησιμοποιήσουμε νευρωνικά δίκτυα

Τα Νευρωνικά δίκτυα, με την αξιοσημείωτη ικανότητα τους να δίνουν νόημα σε πολύπλοκα ή ασαφή δεδομένα, μπορούν να χρησιμοποιηθούν για την εξαγωγή μοντέλων και τον εντοπισμό τάσεων που είναι πάρα πολύ περίπλοκο να γίνουν αντιληπτά από τον άνθρωπο αλλά και από τους ηλεκτρονικούς υπολογιστές. Ένα εκπαιδευμένο νευρωνικό δίκτυο μπορεί να θεωρηθεί ως 'ειδικό' στην κατηγορία των πληροφοριών που του έχει δοθεί για ανάλυση. Αυτό το δίκτυο μπορεί να χρησιμοποιηθεί τότε στην παροχή προβλέψεων δεδομένου των νέων καταστάσεων που παρουσιάζουν ενδιαφέρον και να απαντήσει σε ερωτήσεις του τύπου «τι θα γινόταν αν».

Άλλα πλεονεκτήματα περιλαμβάνουν:

- Προσαρμοστική μάθηση: δυνατότητα να μάθουν πώς να κάνουν εργασίες με βάση τα στοιχεία που παρέχονται για την αρχική εκπαίδευση ή εμπειρία.
- Αυτό-Οργάνωση: Ένα ANN μπορεί να δημιουργήσει τη δική του οργάνωση και εκπροσώπηση των πληροφοριών που λαμβάνει κατά τη διάρκεια της μάθησης.
- Real Time Λειτουργία: Οι ANN υπολογισμοί μπορούν να διεξαχθούν παράλληλα και οι ειδικές συσκευές υλικού σχεδιάζονται και κατασκευάζονται για να επωφεληθούν από αυτή τη δυνατότητα.
- Ανοχή σε σφάλματα μέσω κωδικοποίησης περιττή πληροφορίας: Μερική καταστροφή ενός δικτύου οδηγεί στην αντίστοιχη υποβάθμιση της απόδοσης. Ωστόσο, ορισμένες δυνατότητες του δικτύου μπορούν να διατηρηθούν, ακόμη και με μεγάλες ζημιές.

5.1.4 Νευρωνικά δίκτυα σε σχέση με τους συμβατικούς υπολογιστές

Τα νευρωνικά δίκτυα παρουσιάζουν μια διαφορετική προσέγγιση στην επίλυση προβλημάτων σε σχέση με τους συμβατικούς υπολογιστών. Οι συμβατικοί υπολογιστές χρησιμοποιούν μια αλγοριθμική προσέγγιση δηλαδή ο υπολογιστής ακολουθεί μια σειρά από οδηγίες για να λύσει ένα πρόβλημα. Εάν ο υπολογιστής δεν γνωρίζει τα βήματα που πρέπει να ακολουθήσει δεν μπορεί να λύσει το πρόβλημα. Αυτό περιορίζει την ικανότητα επίλυσης προβλημάτων από τους συμβατικούς υπολογιστές στα προβλήματα που ήδη καταλαβαίνουμε και ξέρουμε πώς να λύσουμε. Οι υπολογιστές θα ήταν πολύ πιο χρήσιμοι αν μπορούσαν να κάνουν πράγματα που δεν γνωρίζουμε ακριβώς πώς να τα κάνουμε.

Τα νευρωνικά δίκτυα επεξεργάζονται τις πληροφορίες με παρόμοιο τρόπο με τον ανθρώπινο εγκέφαλο. Το δίκτυο αποτελείται από ένα μεγάλο αριθμό εξαιρετικά διασυνδεδεμένων στοιχείων επεξεργασίας (νευρώνες) που εργάζονται παράλληλα με την επίλυση ενός συγκεκριμένου προβλήματος. Τα νευρωνικά δίκτυα μαθαίνουν από τα παραδείγματα. Δεν μπορούν να προγραμματιστούν για να εκτελέσουν μια συγκεκριμένη εργασία. Τα παραδείγματα πρέπει να επιλέγονται προσεκτικά, διαφορετικά σπαταλιέται χρήσιμος χρόνος ή ακόμη χειρότερα το δίκτυο μπορεί να λειτουργεί εσφαλμένα. Το μειονέκτημα είναι ότι επειδή το δίκτυο ανακαλύπτει πώς να λύσει το πρόβλημα από μόνο του, η λειτουργία του μπορεί να είναι απρόβλεπτη.

Από την άλλη πλευρά, οι συμβατικοί υπολογιστές χρησιμοποιούν μια γνωστική προσέγγιση για την επίλυση προβλημάτων. Ο τρόπος που το πρόβλημα έχει λυθεί πρέπει να αναφέρεται σε μικρές σαφείς οδηγίες. Αυτές οι οδηγίες στη συνέχεια μετατρέπονται σε ένα πρόγραμμα σε γλώσσα υψηλού επιπέδου και έπειτα σε κώδικα μηχανής ώστε ο υπολογιστής να μπορεί να καταλάβει. Αυτοί οι υπολογιστές είναι εντελώς προβλέψιμοι, αν κάτι πάει στραβά οφείλεται σε λογισμικό ή υλικό σφάλμα.

Τα νευρωνικά δίκτυα και οι συμβατικοί υπολογιστές δεν ανταγωνίζονται το ένα το άλλο

αλλά αλληλοσυμπληρώνονται. Υπάρχουν εργασίες που είναι πιο κατάλληλες για μια αλγοριθμική προσέγγιση, όπως αριθμητικές πράξεις και άλλες που είναι πιο κατάλληλες για νευρωνικά δίκτυα. Ακόμη περισσότερο, μια μεγάλη σειρά από εργασίες, απαιτούν συστήματα που χρησιμοποιούν ένα συνδυασμό των δύο προσεγγίσεων (συνήθως ένας συμβατικός υπολογιστής χρησιμοποιείται για να επιβλέπει το νευρωνικό δίκτυο), προκειμένου να εκτελέσει με μέγιστη αποδοτικότητα.

5.1.5 Ανθρώπινοι και τεχνητοί νευρώνες – ποιες είναι οι ομοιότητες

Πως μαθαίνει ο ανθρώπινος εγκέφαλος

Πολλά εξακολουθούν να είναι ακόμη άγνωστα για το πώς ο εγκέφαλος εκπαιδεύεται μόνος του ώστε να επεξεργάζεται τις πληροφορίες, έτσι λοιπόν οι θεωρίες αφθονούν. Στον ανθρώπινο εγκέφαλο, ένας τυπικός νευρώνας συλλέγει σήματα από τους άλλους μέσα από μια σειρά από λεπτές δομές που ονομάζονται δενδρίτες. Ο νευρώνας στέλνει αιχμές της ηλεκτρικής δραστηριότητας μέσα από ένα μακρύ, λεπτό σύνδεσμο γνωστό ως άξονα ο οποίος χωρίζει σε χιλιάδες κλαδιά.

Στο τέλος του κάθε κλάδου, μια δομή που ονομάζεται σύναψης μετατρέπει τη δραστηριότητα από τον άξονα σε ηλεκτρική και αναστέλλει ή διεγείρει τη δραστηριότητα στους νευρώνες που συνδέεται. Όταν ένας νευρώνας λαμβάνει επαρκής διέγερση, στέλνει μια ακίδα της ηλεκτρικής δραστηριότητας κάτω στον άξονα. Η μάθηση συμβαίνει με την αλλαγή της αποτελεσματικότητας των συνάψεων, έτσι ώστε η επιρροή ενός νευρώνα σ' έναν άλλον να αλλάζει.

Από τους ανθρώπινους νευρώνες στους τεχνητούς

Διεξάγουμε αυτά τα νευρωνικά δίκτυα προσπαθώντας να συναγάγουμε τα ουσιαστικά χαρακτηριστικά των νευρώνων και των διασυνδέσεών τους. Τότε τυπικά προγραμματίζουμε έναν υπολογιστή να προσομοιώσει αυτά τα χαρακτηριστικά. Ωστόσο, επειδή η γνώση των ανθρώπων περί των νευρώνων είναι ελλιπής και η υπολογιστική ισχύ μας είναι περιορισμένη, τα μοντέλα είναι κατ' ανάγκη γενικές εξιδανικεύσεις των πραγματικών δικτύων των νευρώνων.

5.2 Συνελεκτικά νευρωνικά δίκτυα

Μπορούμε να δούμε ένα παράδειγμα που προσπαθούν να λύσουν τα συνελεκτικά νευρωνικά δίκτυα ώστε να καταλάβουμε πιο εύκολα το σκοπό τους. Οι νευρώνες, μαθαίνουν σ' ένα συνελεκτικό νευρωνικό δίκτυο, πώς να μετατρέψουν τα σήματα εισόδου (π.χ. εικόνα μιας

γάτας) σε αντίστοιχα σήματα εξόδου (π.χ. η ετικέτα 'cat') και αυτό αποτελεί τη βάση της αυτοματοποιημένης αναγνώρισης.

Η νοημοσύνη των νευρωνικών δικτύων είναι παράξενη. Ενώ τεχνητά νευρωνικά δίκτυα έχουν ερευνηθεί ήδη από παλαιότερα ήταν μόνο στα τέλη της δεκαετίας του 2000, όταν η σε βάθος μάθηση με χρήση νευρωνικών δικτύων απογειώθηκε. Ο βασικός καταλύτης ήταν η ανάπτυξη της υπολογιστικής ισχύς και των σύνολα δεδομένων καθώς και η πρωτοποριακή έρευνα σε βάθος μάθηση από την Google. Τον Ιούλιο του 2012, οι ερευνητές στο Google εξέθεσαν ένα προηγμένο νευρωνικό δίκτυο σε μια σειρά από μη επισημασμένες στατικές εικόνες κομμένες από YouTube βίντεο. Προς έκπληξή τους, ανακάλυψαν ότι το νευρωνικό δίκτυο έμαθε από μόνο του να ανιχνεύει μια γάτα, υποστηρίζοντας το λαϊκό ισχυρισμό ότι «το Διαδίκτυο είναι κατασκευασμένο από γάτες».

5.2.1 Βιολογικά-εμπνευσμένο μοντέλο

Η τεχνική που χρησιμοποιήθηκε από τους ερευνητές της Google ονομάζεται Συνελεκτικά Νευρωνικά Δίκτυα (CNN), ένα είδος προηγμένου τεχνητού νευρωνικού δικτύου. Διαφέρει από τα τυπικά νευρωνικά δίκτυα από την άποψη της ροής των σημάτων μεταξύ των νευρώνων. Τα τυπικά νευρωνικά δίκτυα περνούν τα σήματα κατά μήκος του καναλιού εισόδου-εξόδου σε μία μόνο κατεύθυνση, χωρίς να επιτρέπουν σήματα σε βρόχο πίσω στο δίκτυο. Αυτό ονομάζεται προς τα εμπρός τροφοδοσία.

Ενώ τα εμπρός δίκτυα τροφοδοσίας έχουν χρησιμοποιηθεί με επιτυχία για την εικόνα και την αναγνώριση κειμένου, απαιτούν όλους τους νευρώνες να είναι συνδεδεμένοι, με αποτέλεσμα μια υπερβολικά πολύπλοκη δομή του δικτύου. Το κόστος της πολυπλοκότητας μεγαλώνει όταν το δίκτυο πρέπει να έχει εκπαιδευτεί σε μεγάλα σύνολα δεδομένων τα οποία, σε συνδυασμό με τους περιορισμούς των ταχυτήτων επεξεργασίας υπολογιστή, οδηγούν σε μεγάλους χρόνους κατάρτισης. Ως εκ τούτου, τα εμπρός δίκτυα τροφοδοσίας δεν χρησιμοποιούνται πλέον από την επικρατούσα μηχανικής μάθησης στη σημερινή εποχή.

Το 1986, οι ερευνητές Hubel και Wiesel εξέταζαν τον οπτικό φλοιό της γάτας όταν ανακάλυψαν ότι το δεκτικό πεδίο αποτελείται από επιμέρους περιοχές που στρώθηκαν η μία πάνω από την άλλη για να καλύψουν το σύνολο του οπτικού πεδίου. Αυτά τα στρώματα ενεργούν ως φίλτρα που επεξεργάζονται την είσοδο εικόνων, οι οποίες στη συνέχεια διαβιβάζονται στις επόμενες στρώσεις. Αυτό αποδείχθηκε ότι είναι μια απλούστερη και πιο αποτελεσματική τεχνική για να μεταφέρουν τα σήματα. Το 1998, ο Yann LeCun και ο Yoshua Bengio προσπάθησαν να συλλάβουν την οργάνωση των νευρώνων στον οπτικό φλοιό της γάτας ως μια μορφή τεχνητών καθαρών νευρώνων, για την ίδρυση της βάσης του πρώτου CNN.

5.2.2 CNN σε 5 βήματα

Υπάρχουν τέσσερα κύρια στάδια CNN: συνέλιξη, υπό-δειγματοληψία, ενεργοποίηση και πλήρη συνεκτικότητα.

Βήμα 1: Συνέλιξη

Τα πρώτα στρώματα που λαμβάνουν ένα σήμα εισόδου ονομάζονται φίλτρα συνέλιξης. Η συνέλιξη είναι μια διαδικασία όπου το δίκτυο προσπαθεί να ονομάσει το σήμα εισόδου με βάση αυτό που έχει μάθει στο παρελθόν. Αν το σήμα εισόδου μοιάζει με τις προηγούμενες εικόνες γατών, το σήμα αναφοράς ‘γάτα’ θα πρέπει να αναμιγνύεται ή να ανελίσσεται με το σήμα εισόδου. Το σήμα εξόδου που προκύπτει στη συνέχεια διαβιβάζεται στο επόμενο στρώμα.

Η ανέλιξη έχει την ωραία ιδιότητα να είναι μεταφραστικά αμετάβλητη. Διαισθητικά, αυτό σημαίνει ότι κάθε φίλτρο ανελίζεως αντιπροσωπεύει ένα χαρακτηριστικό που το ενδιαφέρει (π.χ. μουστάκια, γούνα) και ο αλγόριθμος CNN μαθαίνει ποια χαρακτηριστικά περιλαμβάνουν το σήμα αναφοράς (δηλαδή τη γάτα). Η ισχύς του σήματος εξόδου δεν εξαρτάται από που τα χαρακτηριστικά βρίσκονται, αλλά απλώς από το αν τα χαρακτηριστικά είναι παρόντα. Ως εκ τούτου, μια γάτα θα μπορούσε να κάθεται σε διαφορετικές θέσεις και ο αλγόριθμος CNN θα εξακολουθούσε να είναι σε θέση να το αναγνωρίσει.

Βήμα 2: Υπό-δειγματοληψία

Οι είσοδοι από το προηγούμενο στάδιο μπορούν να ‘εξομαλυνθούν’ για να μειωθεί η ευαισθησία των φίλτρων σε θορύβους και παραλλαγές. Αυτή η διαδικασία εξομάλυνσης ονομάζεται υπό-δειγματοληψία και μπορεί να επιτευχθεί με τη λήψη μέσου ή τη λήψη της μέγιστης τιμής πάνω από ένα δείγμα του σήματος. Παραδείγματα μεθόδων υπό-δειγματοληψίας (για σήματα εικόνας) περιλαμβάνουν τη μείωση του μεγέθους της εικόνας, ή της αντίθεση χρωμάτων σε κόκκινο, πράσινο, μπλε (RGB) κανάλι.

Βήμα 3: Ενεργοποίηση

Η στιβάδα ενεργοποίησης ελέγχει τον τρόπο με τον οποίο το σήμα ρέει από το ένα στρώμα στο επόμενο, μιμείται τον τρόπο που οι νευρώνες τροφοδοτούνται στον εγκέφαλο μας. Σήματα εξόδου που συνδέονται στενά με το παρελθόν θα μπορούσαν να ενεργοποιήσουν περισσότερους νευρώνες, επιτρέποντας σε σήματα να πολλαπλασιάζονται πιο αποτελεσματικά για την ταυτοποίηση.

Το CNN είναι συμβατό με μια μεγάλη ποικιλία πολύπλοκων λειτουργιών ενεργοποίησης

για να μοντελοποιηθεί η διάδοση του σήματος. Η πιο κοινή λειτουργία είναι η Διορθωμένη Γραμμική Μονάδα (Relu), η οποία ευνοείται λόγω της μεγαλύτερης ταχύτητας κατάρτισης.

Βήμα 4: Πλήρη συνεκτικότητα

Οι τελευταίες στρώσεις στο δίκτυο είναι πλήρως συνδεδεμένες, πράγμα που σημαίνει ότι οι νευρώνες των προηγούμενων στρωμάτων συνδέονται με κάθε νευρώνα στις επόμενες στρώσεις. Αυτό μιμείται υψηλές συλλογιστικές μεθόδους όπου όλες οι πιθανές οδοί από την είσοδο στην έξοδο λαμβάνονται υπόψη.

Βήμα 5: Απώλεια

Κατά την κατάρτιση του νευρικού δικτύου, υπάρχει ένα επιπλέον στρώμα που ονομάζεται στρώμα απώλειας. Αυτό το στρώμα παρέχει ανάδραση στο νευρωνικό δίκτυο με το αν εντοπίζονται σωστά οι είσοδοι, και αν όχι, πόσο μακριά έπεσαν οι εικασίες. Αυτό βοηθά το νευρωνικό δίκτυο να γνωρίζει τις σωστές έννοιες κατά τη διάρκεια της κατάρτισης. Αυτό είναι πάντα το τελευταίο στρώμα της κατάρτισης.

5.2.3 Εκτέλεση

Οι αλγόριθμοι που χρησιμοποιούνται στην κατάρτιση του CNN είναι ανάλογοι με τη μελέτη για τις εξετάσεις χρησιμοποιώντας κάρτες flash. Κατ' αρχάς, διαλέγεις διάφορα flash cards και ελέγχεις αν έχεις κατανοήσει τις έννοιες σε κάθε κάρτα. Για τις κάρτες με τις έννοιες που ήδη γνωρίζεις, τις απορρίπτεις.

Για αυτές τις κάρτες με έννοιες που δεν είσαι σίγουρος, τοποθέτησε τες πίσω στο σωρό. Επανάλαβε αυτή τη διαδικασία μέχρι να είσαι αρκετά σίγουρος ότι γνωρίζεις αρκετές έννοιες για να γράφεις καλά στις εξετάσεις. Αυτή η μέθοδος επιτρέπει να επικεντρωθείς στις λιγότερο γνωστές έννοιες διαβάζοντας τες συχνά. Τυπικά, αυτοί οι αλγόριθμοι ονομάζονται αλγόριθμοι καθοδικής κλίσης για προώθηση της μάθησης. Σύγχρονοι αλγόριθμοι βαθιάς μάθησης χρησιμοποιούν μια παραλλαγή όπου αντί για την επιλογή των flash cards διαδοχικά, η επιλογή γίνεται τυχαία. Η τυχαία προσέγγιση βοηθά στην ελαχιστοποίηση οποιασδήποτε μορφής μεροληψίας στην εκμάθηση των θεμάτων.

Οι αλγόριθμοι μάθησης απαιτούν ανατροφοδότηση. Αυτό γίνεται χρησιμοποιώντας ένα σύνολο επικύρωσης όπου το CNN κάνει προβλέψεις και να τις συγκρίνει με τις πραγματικές ετικέτες ή την αλήθεια. Οι προβλέψεις που έχουν λάθη τροφοδοτούνται πίσω στο CNN για να βελτιωθεί η μάθηση, σε μια λεγόμενη πίσω πέραση. Επίσημως, ο αλγόριθμος αυτός ονομάζεται πίσω διάδοση σφαλμάτων.

Οι CNNs είναι πάρα πολύ περίπλοκο να εφαρμοστούν από το μηδέν. Σήμερα, οι επαγγελματίες μηχανικής μάθησης χρησιμοποιούν συχνά βιβλιοθήκες που έχουν αναπτυχθεί όπως Caffe, Theano και Tensorflow για το έργο τους, αναλύουμε περισσότερα στο επόμενο κεφάλαιο.

5.2.4 Ερμηνεία

Παρόμοια με το πώς μερικές φορές είναι δύσκολο να εξορθολογήσουμε τις σκέψεις μας, το εμπνευσμένο από τον εγκέφαλο CNN είναι επίσης δύσκολο να ερμηνευθεί λόγω της πολυπλοκότητας του. Χρησιμοποιεί πολυάριθμα φίλτρα, σε συνδυασμό με διαφορετικούς τρόπους και διέπεται από διάφορες λειτουργίες ενεργοποίησης. Ωστόσο, μια τεχνική αναπτύχθηκε πρόσφατα για να μας βοηθήσει να παρατηρήσουμε πώς ένα νευρωνικό δίκτυο μαθαίνει να αναγνωρίζει τα σήματα.

Η μέθοδος DeConvolution Neural Net, όπως υποδηλώνει η ονομασία της, ανατέμνει την πρόοδο της κατάρτισης σε κάθε στρώμα. Εξετάζει το πρώτο στρώμα όπως επίσης και το τελευταίο πλήρως συνδεδεμένο στρώμα. Το πρώτο στρώμα αντιπροσωπεύει το σύνολο των χαρακτηριστικών που το CNN είναι εξοικειωμένο και καθορίζει ποια είναι τα βασικά αντικείμενα (χαρακτηριστικά χαμηλού επιπέδου) που το CNN θα είναι σε θέση να ανιχνεύσει. Από την άλλη πλευρά, το τελικό στρώμα αντιπροσωπεύει το πόσο καλά τα χαρακτηριστικά συνδέονται μαζί για να σχηματίσουν έννοιες - ένας σαφέστερος διαχωρισμός των εννοιών συνεπάγεται ότι το νευρωνικό δίκτυο μπορεί καλύτερα να διακρίνει τις διαφορετικές έννοιες χωρίς σύγχυση.

5.2.5 Συμπέρασμα

Τα Συνελεκτικά Νευρωνικά Δίκτυα είναι μια δημοφιλής τεχνική βαθιάς μάθησης για τις τρέχουσες εργασίες οπτικής αναγνώρισης. Όπως όλες οι τεχνικές βαθιάς μάθησης, το CNN εξαρτάται σε μεγάλο βαθμό από το μέγεθος και την ποιότητα των δεδομένων εκπαίδευσης. Λαμβάνοντας υπόψη ένα καλά προετοιμασμένο σύνολο δεδομένων, το CNN είναι σε θέση να ξεπεράσει τον άνθρωπο στην οπτική ικανότητα αναγνώρισης. Ωστόσο, δεν είναι ακόμα ισχυρά σε οπτικές ιδιαιτερότητες, όπως το έντονο φως και ο θόρυβος, τα οποία οι άνθρωποι είναι σε θέση να αντιμετωπίσουν. Η θεωρία του CNN εξακολουθεί να αναπτύσσεται και οι ερευνητές εργάζονται για να του προσδώσουν ιδιότητες, όπως η ενεργός προσοχή και η απευθείας σύνδεση με τη μνήμη.

5.2.6 Εφαρμογές

Νευρωνικά δίκτυα στην πράξη

Τα νευρωνικά δίκτυα έχουν ευρεία εφαρμογή σε πολλά επιχειρηματικά προβλήματα στον πραγματικό κόσμο. Στην πραγματικότητα, έχουν ήδη εφαρμοστεί με επιτυχία σε πολλούς βιομηχανικούς κλάδους.

Δεδομένου ότι τα νευρωνικά δίκτυα είναι κατάλληλα για αναγνώριση προτύπων και τάσεων δεδομένων, χρησιμοποιούνται για πρόβλεψη καταστάσεων όπως:

- Πρόβλεψη πωλήσεων
- Έλεγχο βιομηχανικών διεργασιών
- Έρευνα αγοράς
- Επιχύρωση δεδομένων
- Διαχείριση κινδύνου
- Στοχευμένο Μάρκετινγκ

Επίσης, τα ANN χρησιμοποιούνται επίσης στις ακόλουθες ειδικές περιπτώσεις: για την αναγνώριση των ομιλητών στις επικοινωνίες, για τη διάγνωση της ηπατίτιδας, την ανάκτηση των τηλεπικοινωνιών από ελαττωματικό λογισμικό, την ερμηνεία του δύσκολων κινεζικών λέξεων, την υποθαλάσσια ανίχνευση ναρκών, την ανάλυση υφής, την τρισδιάστατη αναγνώριση αντικειμένων, τη χειρόγραφη αναγνώριση λέξεων και την αναγνώριση προσώπου.

Νευρωνικά δίκτυα στην ιατρική

Τα ANN μονοπωλούν το ενδιαφέρον στο χώρο της έρευνας στην ιατρική και πιστεύεται ότι θα λάβουν εκτενή εφαρμογή σε βιοϊατρικά συστήματα μέσα στα επόμενα λίγα χρόνια. Προς το παρόν, η έρευνα είναι ως επί το πλείστον για τη μοντελοποίηση τμημάτων του ανθρώπινου σώματος και την αναγνώριση ασθενειών από διάφορες σαρώσεις (π.χ. καρδιογραφήματα, σαρώσεις CAT, υπέρηχους, κλπ).

Τα νευρωνικά δίκτυα είναι ιδανικά για την αναγνώριση ασθενειών χρησιμοποιώντας σαρώσεις αφού δεν υπάρχει ανάγκη για την παροχή ενός συγκεκριμένου αλγορίθμου για τον τρόπο προσδιορισμού της νόσου. Τα νευρωνικά δίκτυα μαθαίνουν από το παράδειγμα και έτσι δεν χρειάζονται λεπτομέρειες για το πώς να αναγνωρίζουν τη νόσο. Αυτό που χρειάζεται είναι μια σειρά από παραδείγματα που είναι αντιπροσωπευτικά του συνόλου των παραλλαγών της

νόσου. Η ποσότητα των παραδειγμάτων δεν είναι τόσο σημαντική όσο η ποιότητα. Τα παραδείγματα πρέπει να επιλεγούν πολύ προσεκτικά εάν θέλουμε το σύστημα να εκτελεί αξιόπιστα και αποτελεσματικά το σκοπό του.

Μοντελοποίηση και διάγνωση του καρδιαγγειακού συστήματος

Τα Νευρωνικά Δίκτυα χρησιμοποιούνται πειραματικά για να μοντελοποιήσουν το ανθρώπινο καρδιαγγειακό σύστημα. Η διάγνωση μπορεί να επιτευχθεί με τη δημιουργία ενός μοντέλου του καρδιαγγειακού συστήματος του ανθρώπου, συγκρίνοντας το σε πραγματικό χρόνο με τις φυσιολογικές μετρήσεις που λαμβάνονται από τον ασθενή. Αν αυτή η ρουτίνα γίνεται τακτικά, πιθανές βλαβερές ιατρικές παθήσεις μπορεί να ανιχνευθούν σε πρώιμο στάδιο και ως εκ τούτου να κάνουν τη διαδικασία της καταπολέμησης της νόσου πολύ πιο εύκολη.

Ένα μοντέλο του καρδιαγγειακού συστήματος του ατόμου πρέπει να μιμείται τη σχέση μεταξύ των φυσιολογικών μεταβλητών (δηλ., ο καρδιακός ρυθμός, η συστολική και διαστολική αρτηριακή πίεση, ο ρυθμός αναπνοής) σε διαφορετικά επίπεδα φυσικής δραστηριότητας. Αν ένα μοντέλο είναι προσαρμοσμένο σε ένα άτομο, τότε γίνεται ένα πρότυπο για τη φυσική κατάσταση αυτού του ατόμου. Ο προσομοιωτής θα πρέπει να είναι σε θέση να προσαρμοστεί στα χαρακτηριστικά του κάθε ατόμου χωρίς την επίβλεψη ενός εμπειρογνώμονα. Αυτό απαιτεί ένα νευρωνικό δίκτυο.

Ένας άλλος λόγος που δικαιολογεί τη χρήση της τεχνολογίας ANN, είναι η ικανότητα των ANN να παράσχουν σύντηξη αισθητήρων, δηλαδή του συνδυασμού των τιμών διαφορετικών αισθητήρων. Η σύντηξη αισθητήρων επιτρέπει στους ANN να μάθουν πολύπλοκες σχέσεις μεταξύ των επιμέρους τιμών των αισθητήρων, σχέσεις που διαφορετικά θα χανθούν αν οι τιμές αναλυθούν ατομικά. Στην ιατρική μοντελοποίηση και διάγνωση, αυτό σημαίνει ότι, ακόμα κι αν κάθε αισθητήρας σε ένα σύνολο μπορεί να είναι ευαίσθητος μόνο σε μία ειδική φυσιολογική μεταβλητή, το ANN είναι ικανό να ανιχνεύσει πολύπλοκες ιατρικές καταστάσεις με σύντηξη των δεδομένων από τους μεμονωμένους βιοϊατρικούς αισθητήρες.

Ηλεκτρονική μύτη

Τα ANN χρησιμοποιούνται πειραματικά για την υλοποίηση ηλεκτρονικών μυτών. Η ηλεκτρονική μύτη έχει πολλές πιθανές εφαρμογές στην τηλεϊατρική. Η τηλεϊατρική είναι η άσκηση της ιατρικής από απόσταση μέσω ενός συνδέσμου επικοινωνίας. Η ηλεκτρονική μύτη θα προσδιορίσει τις οσμές στο απομακρυσμένο χειρουργικό περιβάλλον. Αυτές οι οσμές στη συνέχεια θα διαβιβάζονται ηλεκτρονικά σε μια άλλη περιοχή όπου ένα σύστημα παραγωγής θα τις αναπαράγει. Η αίσθηση της όσφρησης μπορεί να είναι μια σημαντική έννοια για το χειρουργό.

Ιατρός άμεσα

Μια εφαρμογή που αναπτύχθηκε στα μέσα της δεκαετίας του 1980 που ονομάζεται «ιατρός άμεσα» εκπαιδευσε ένα αυτοσυσχετιζόμενο δίκτυο νευρωνικών δικτύων να αποθηκεύσει ένα μεγάλο αριθμό ιατρικών αρχείων, καθένα από τα οποία περιλαμβάνει πληροφορίες σχετικά με τα συμπτώματα, τη διάγνωση και τη θεραπεία για μια συγκεκριμένη υπόθεση. Μετά την εκπαίδευση, η είσοδος αποτελείται από ένα σύνολο συμπτωμάτων και το αποτέλεσμα παρουσιάζει το πρότυπο που αντιπροσωπεύει την καλύτερη διάγνωση και θεραπεία.

Νευρωνικά δίκτυα στις επιχειρήσεις

Οι επιχειρήσεις αποτελούν ένα γενικό πεδίο με διάφορους τομείς εξειδίκευσης όπως η λογιστική ή η χρηματοοικονομική ανάλυση. Σχεδόν κάθε νευρωνικό δίκτυο θα μπορούσε να εφαρμοστεί σε έναν επιχειρηματικό τομέα ή μια οικονομική ανάλυση. Υπάρχει τάση για τη χρήση των δικτύων νευρικών για επαγγελματικούς σκοπούς, συμπεριλαμβανομένης της κατανομής των πόρων και τον προγραμματισμό των εργασιών. Υπάρχει επίσης ένα ισχυρό δυναμικό για τη χρήση νευρωνικών δικτύων για την εξόρυξη δεδομένων. Η έρευνα σε αυτόν τον τομέα έχει χαρακτηριστεί ως βιομηχανική ιδιοκτησία. Έτσι, δεν είναι δυνατόν να γνωρίζουμε την πλήρη έκταση των εργασιών που βρίσκονται σε εξέλιξη.

Μάρκετινγκ

Υπάρχει μια εφαρμογή μάρκετινγκ η οποία έχει ενσωματωθεί με ένα σύστημα νευρωνικού δικτύου. Το Στρατηγικό Μάρκετινγκ Αεροπορικών Εταιρειών είναι ένα ηλεκτρονικό σύστημα από διάφορες έξυπνες τεχνολογίες, συμπεριλαμβανομένων των έμπειρων συστημάτων. Ένα νευρωνικό δίκτυο είναι ενσωματωμένο με το σύστημα και έχει εκπαιδευτεί χρησιμοποιώντας οπισθοδιάδοση να βοηθάει τον έλεγχο της κατανομής των θέσεων στα αεροπλάνα. Το περιβάλλον της εφαρμογής αλλάζει ταχέως και συνεχώς, πράγμα που απαιτούσε συνεχώς προσαρμοστική λύση. Το σύστημα χρησιμοποιείται για να παρακολουθεί και να συστήνει συμβουλές κρατήσεων για κάθε αεροδρόμιο αναχώρησης. Οι πληροφορίες αυτές έχουν άμεσο αντίκτυπο στην κερδοφορία της αεροπορικής εταιρείας και μπορεί να προσφέρουν ένα τεχνολογικό πλεονέκτημα στους χρήστες του συστήματος.

Ενώ είναι σημαντικό το γεγονός ότι τα νευρωνικά δίκτυα έχουν εφαρμοστεί σε αυτό το πρόβλημα, είναι επίσης σημαντικό να δούμε ότι αυτή η έξυπνη τεχνολογία μπορεί να ενσωματωθεί με έμπειρα συστήματα και άλλες προσεγγίσεις για να δημιουργήσει ένα λειτουργικό σύστημα.

5.2.7 Επικρίσεις και κριτικές

Θέματα εκπαίδευσης

Μια κοινή κριτική των νευρωνικών δικτύων, ιδιαίτερα στη ρομποτική, είναι ότι απαιτούν μεγάλη ποικιλία κατάρτισης για πραγματικού κόσμου λειτουργίες. Αυτό δεν αποτελεί έκπληξη, δεδομένου ότι οποιαδήποτε μηχανική μάθηση χρειάζεται επαρκή αντιπροσωπευτικά παραδείγματα, προκειμένου να συλλάβει την υποκείμενη δομή που επιτρέπει να γενικεύσουμε σε νέες περιπτώσεις.

Θεωρητικά ζητήματα

Ο AK Dewdney, μαθηματικός και ο επιστήμονας στο Πανεπιστήμιο του Δυτικού Οντάριο, έγραψε το 1997, «Παρά το γεγονός ότι τα νευρωνικά δίκτυα μπορούν να λύσουν μερικά προβλήματα παιχνιδιών, η δύναμη των υπολογισμών είναι τόσο περιορισμένη που με εκπλήσσει που κάποιοι τους παίρνουν στα σοβαρά ως ένα γενικό εργαλείο επίλυσης προβλημάτων». Κανένα νευρωνικό δίκτυο δεν έχει ποτέ αποδειχθεί ότι επιλύει δύσκολα υπολογιστικά προβλήματα, όπως το πρόβλημα n-Queens, το πρόβλημα του περιπλανώμενου πωλητή, ή το πρόβλημα της παραγοντοποίησης μεγάλων ακεραίων.

Εκτός από τη χρησιμότητά τους, μια θεμελιώδης αντίρρηση για τα τεχνητά νευρωνικά δίκτυα είναι ότι δεν αντικατοπτρίζουν το πώς πραγματικά λειτουργούν οι νευρώνες. Η πίσω διάδοση βρίσκεται στην καρδιά όλων των τεχνητών νευρωνικών δικτύων αλλά δεν υπάρχει απόδειξη ενός τέτοιου μηχανισμού σε φυσικά νευρωνικά δίκτυα, και μάλιστα φαίνεται να έρχεται σε αντίθεση με τη θεμελιώδη αρχή της πραγματικής λειτουργίας των νευρώνων όπου οι πληροφορίες μπορεί να ρέουν μόνο προς τα εμπρός, κατά μήκος του άξονα.

Το κίνητρο πίσω από τεχνητά νευρωνικά δίκτυα δεν είναι απαραίτητα η αναπαραγωγή της πραγματικής νευρικής λειτουργίας, αλλά το να χρησιμοποιηθούν τα φυσικά νευρωνικά δίκτυα ως πηγή έμπνευσης για μια προσέγγιση στην πληροφορική. Ένα γενικό πλεονέκτημα των τεχνητών νευρωνικών δικτύων είναι, επομένως, ότι ενσαρκώνουν κάποια νέα και ισχυρή γενική αρχή για την επεξεργασία των πληροφοριών. Δυστυχώς, αυτές οι γενικές αρχές είναι ασαφείς και συχνά ισχυρίστηκαν ότι είναι αναδυόμενες από το ίδιο το νευρωνικό δίκτυο. Αυτό επιτρέπει την απλή στατιστική συσχέτιση (η βασική λειτουργία των τεχνητών νευρωνικών δικτύων) να περιγραφεί ως μάθηση ή αναγνώριση.

Θέματα Hardware

Για την υλοποίηση μεγάλων και αποτελεσματικών νευρωνικών δικτύων, πρέπει να δεσμευτούν σημαντικοί πόροι επεξεργασίας και αποθήκευσης. Ενώ ο εγκέφαλος έχει υλικό

προσαρμοσμένο στο έργο των σημάτων επεξεργασίας, προσομοιώνοντας ακόμα και μια πιο απλοποιημένη μορφή μπορεί να υποχρεώσει έναν σχεδιαστή νευρωνικών δικτύων να γεμίσει πολλές εκατομμύρια γραμμές βάσης δεδομένων για τις συνδέσεις του (γεγονός που μπορεί να καταναλώσει τεράστιες ποσότητες της μνήμης του υπολογιστή και χώρο στο σκληρό δίσκο). Επιπλέον, ο σχεδιαστής των νευρωνικών δικτύων θα πρέπει συχνά να προσομοιώνει τη μετάδοση των σημάτων μέσω πολλών συνδέσεων και συναφών νευρώνων – γεγονός που συνδυάζεται με απίστευτες ποσότητες επεξεργαστικής ισχύς της CPU αλλά και χρόνου.

Ο Jürgen Schmidhuber σημειώνει ότι η αναβίωση των νευρωνικών δικτύων στον εικοστό πρώτο αιώνα καθώς και η ανανέωση της επιτυχίας τους στις εργασίες αναγνώρισης εικόνας οφείλονται σε μεγάλο βαθμό στην πρόοδο του ηαρδWARE: από το 1991 έως το 2015, η υπολογιστική ισχύ, ειδικά όπως παραδίδεται από GPUs, έχει αυξηθεί περίπου ένα εκατομμύριο φορές, κάνοντας το πρότυπο αλγόριθμο ανάστροφης διάδοσης εφικτό για τα δίκτυα κατάρτισης. Η χρήση των GPUs, αντί των συνήθους επεξεργαστών μπορεί να μειώσει τους χρόνους κατάρτισης για ορισμένα δίκτυα από μήνες σε λίγες ημέρες. Η υπολογιστική ισχύ συνεχίζει να αυξάνεται σύμφωνα με το νόμο του Moore, πράγμα που σημαίνει ότι μπορεί να παρέχει επαρκείς πόρους για την επίτευξη νέων καθηκόντων.

Κεφάλαιο 6

Βιβλιοθήκες μηχανικής μάθησης στην γλώσσα Python

Χωρίς αμφιβολία να νευρωνικά δίκτυα και η μηχανική μάθηση είναι από τα πιο δημοφιλή θέματα στο χώρο της πληροφορικής τα τελευταία χρόνια. Σ' αυτή την ενότητα θα αναλυθούν τα πιο δημοφιλή εργαλεία της μηχανικής αναζήτησης στη γλώσσα προγραμματισμού Python. Πρώτα θα δωθεί μια μικρή περιγραφή για το κάθε εργαλείο και στη συνέχεια θα τα συγκρίνουμε τις βιβλιοθήκες.

6.1 Tensorflow



Σχήμα 6.1: Tensorflow λογότυπο

Η βιβλιοθήκη Tensorflow¹ είναι η νεότερη βιβλιοθήκη νευρωνικών δικτύων και μηχανικής

¹<https://www.tensorflow.org/>

μάθησης. Είναι μια υψηλού επιπέδου βιβλιοθήκη που βοηθά τον υψηλού επιπέδου προγραμματισμό αποφεύγοντας λεπτομέρειες χαμηλού επιπέδου. Η εστίαση είναι περισσότερο στην έκφραση των υπολογισμών, όπως μένα διάγραμμα ροής δεδομένων, γεγονός που την κατατάσσει κατάλληλη για την επίλυση σύνθετων προβλημάτων.

Η βιβλιοθήκη είναι κυρίως γραμμένη στη γλώσσα προγραμματισμού C++ αλλά περιλαμβάνει τις συνδέσεις (bindings) στη Python, έτσι ώστε να μην θυσιάζεται η απόδοση των προγραμμάτων. Ένα από τα χαρακτηριστικά της είναι η ευέλικτη αρχιτεκτονική που της επιτρέπει να αναπτυχθεί σε μία ή περισσότερες υπολογιστικές μονάδες (CPUs) ή γραφικές υπολογιστικές μονάδες (GPUs) σε ένα desktop, server, ή μια κινητή συσκευή, όλα με το ίδιο API². Δεν υπάρχουν άλλες βιβλιοθήκες μέχρι στιγμής με αυτή τη δυνατότητα.

Η βιβλιοθήκη αυτή αρχικά αναπτύχθηκε για το έργο Google Brain³ και τώρα χρησιμοποιείται από εκατοντάδες μηχανικούς σε όλη την εταιρεία, οπότε αποδεδειγμένα είναι ικανή να δημιουργήσει ενδιαφέρουσες λύσεις.

Σε αντίθεση με άλλες αριθμητικές βιβλιοθήκες η TensorFlow έχει σχεδιαστεί για χρήση στον τομέα της έρευνας και ανάπτυξης, των συστημάτων παραγωγής αλλά και στην RankBrain στην αναζήτηση Google και στο έργο DeepDream.

6.2 Scikit-learn



Σχήμα 6.2: Scikit-learn λογότυπο

Η βιβλιοθήκη Scikit-Learn⁴ αναπτύχθηκε αρχικά από τον David Cournapeau στη Google.

Αργότερα ο Matthieu Brucher εντάχθηκε στο έργο και άρχισε να τη χρησιμοποιεί εκτός της διπλωματικής εργασίας του. Το 2010 το INRIA πήρε μέρος και η πρώτη δημόσια έκδοση (v0.1 beta) δόθηκε στη δημοσιότητα στα τέλη Ιανουαρίου του 2010. Το έργο έχει τώρα πάνω από 30 ενεργούς συμμετέχοντες και έχει καταβληθεί χορηγία από το INRIA, το Google, Tinyclues και το Ίδρυμα Ελεύθερου Λογισμικού της Python. Έχει αδειοδοτηθεί από μια ανεκτική απλουστευμένη άδεια BSD και διανέμεται υπό πολλούς διανομές Linux, ενθαρρύνοντας την

²Το API αλλιώς και διεπαφή προγραμματισμού εφαρμογών (για συντομία: Διεπαφή ή API από το αγγλικό API, Application Programming Interface)

³<https://research.google.com/teams/brain/>

⁴<https://scikit-learn.org/>

ακαδημαϊκή αλλά και την επαγγελματική χρήση.

Η βιβλιοθήκη αυτή παρέχει μια σειρά από επιβλέψιμους και μη επιβλέψιμους αλγορίθμους μάθησης μέσω μιας συνεπούς διεπαφής στην Python.

Η βιβλιοθήκη είναι χτισμένη με SciPy⁵ (Επιστημονική Python) που πρέπει να εγκατασταθεί για να μπορέσει να χρησιμοποιηθεί. Αυτή η στοίβα περιλαμβάνει:

- NumPy: n διαστάσεων πακέτο
- SciPy: Θεμελιώδεις βιβλιοθήκη για την επιστημονική υπολογιστική
- Matplotlib: Ολοκληρωμένη 2D / 3D σχεδίαση
- IPython: Ενισχυμένη διαδραστική κονσόλα
- SymPy: Συμβολικά μαθηματικά
- Pandas: Δομές Δεδομένων και ανάλυση

Το όραμα για τη βιβλιοθήκη είναι να μπορέσει να παρέχει τη δυναμική και την υποστήριξη που χρειάζονται τα συστήματα παραγωγής. Αυτό σημαίνει βαθιά εστίαση σε ζητήματα όπως η ευκολία χρήσης, η ποιότητα κώδικα, η συνεργασία, η τεκμηρίωση και η απόδοση.

Παρά το γεγονός ότι το περιβάλλον είναι Python, οι βιβλιοθήκες C είναι απαραίτητες για την απόδοση.

Η συγκεκριμένη βιβλιοθήκη επικεντρώνεται σε στοιχεία μοντελοποίησης και δεν επικεντρώνεται στην φόρτωση, το χειρισμό και την περίληψη των δεδομένων. Για αυτά τα χαρακτηριστικά, υπάρχουν άλλες βιβλιοθήκες.

Μερικά δημοφιλή μοντέλα που παρέχονται από την συγκεκριμένη βιβλιοθήκη περιλαμβάνουν:

- Ομαδοποίηση: για την ομαδοποίηση μη επισημασμένων δεδομένων όπως ο KMeans.
- Cross Validation: για την εκτίμηση της απόδοσης των εποπτευόμενων μοντέλων σε άορα δεδομένα.
- Σύνολα δεδομένων: για σύνολα δεδομένων δοκιμών και για την δημιουργία συνόλων δεδομένων με συγκεκριμένες ιδιότητες για τη διερεύνηση της συμπεριφοράς του μοντέλου.
- Μείωσης διάστασης: για τη μείωση του αριθμού των χαρακτηριστικών στα δεδομένα για περιλήψεις, για την οπτικοποίηση και επιλογή χαρακτηριστικών, όπως η ανάλυση βασικών συστατικών.

⁵<https://www.scipy.org/>

- Ensemble μεθόδους: για το συνδυασμό των προβλέψεων πολλαπλών εποπτευόμενων μοντέλων.
- Εξαγωγή χαρακτηριστικών: για τον καθορισμό των ιδιοτήτων στα δεδομένα εικόνας και κειμένου.
- Επιλογής χαρακτηριστικών: για τον εντοπισμό ουσιαστικών χαρακτηριστικών από για την δημιουργία εποπτευόμενων μοντέλων.
- Συντονισμός παραμέτρων: για να πάρει τα μέγιστα από εποπτευόμενα μοντέλα.
- Πολλαπλή Μάθηση: Για να συνοψίζει και να απεικονίζει συγχροτήματα πολλαπλών διαστάσεων των δεδομένων.
- Εποπτευόμενα Μοντέλα: μια τεράστια ποικιλία που δεν περιορίζεται σε γενικευμένα γραμμικά μοντέλα, αναλύσεις, Naive Bayes, νευρωνικά δίκτυα, μηχανές διανυσμάτων υποστήριξης και δένδρα αποφάσεων.

Τέλος χρησιμοποιείται από εταιρείες και οργανισμούς όπως: Inria, Mendeley, *wise.io*, Evernote, Telecom ParisTech και Aweber.

6.3 Theano



Σχήμα 6.3: Theano λογότυπο

Η βιβλιοθήκη αυτή αποτελεί ένα έργο ανοικτού πηγαίου κώδικα που διατίθεται βάσει της άδειας BSD. Αναπτύχθηκε από την ομάδα LISA (τώρα MILA) στο Πανεπιστήμιο του Μόντρεαλ στον Καναδά. Πήρε το όνομά του από την Θεανώ, Ελληνίδα μαθηματικό και σύζυγο του Πυθαγόρα.

Είναι μια βιβλιοθήκη Python που επιτρέπει τον ορισμό, τη βελτιστοποίηση, και την αξιολόγηση μαθηματικών εκφράσεων, ειδικά αυτών με πολυδιάστατες-συστοιχίες (numpy, ndarray). Χρησιμοποιώντας την, είναι δυνατόν να επιτευχθούν ταχύτητες που συναγωνίζονται εφαρμογές C για προβλήματα που αφορούν μεγάλες ποσότητες δεδομένων. Μπορεί επίσης να ξεπεράσει τη C σε CPU κατά πολλές τάξεις μεγέθους με την αξιοποίηση των πρόσφατων GPUs.

Η Theano⁶ συνδυάζει πτυχές ενός συστήματος άλγεβρας υπολογιστών ("AS) με πτυχές μεταγλωττιστή βελτιστοποίησης. Μπορεί επίσης να δημιουργήσει προσαρμοσμένο κώδικα " για πολλές μαθηματικές πράξεις. Αυτός ο συνδυασμός των CAS με την βελτιστοποίηση είναι ιδιαίτερα χρήσιμος για εργασίες στις οποίες περιπλέκονται μαθηματικές εκφράσεις που αξιολογούνται επανειλημμένα και η ταχύτητα αξιολόγησης είναι κρίσιμη. Για καταστάσεις όπου πολλές διαφορετικές εκφράσεις αξιολογούνται κάθε φορά, η βιβλιοθήκη αυτή μπορεί να ελαχιστοποιήσει το ποσό της σύνταξης/ανάλυσης, εξακολουθώντας να παρέχει συμβολικά χαρακτηριστικά, όπως η αυτόματη διαφοροποίηση.

Ο μεταγλωττιστής της βιβλιοθήκης εφαρμόζει πολλές βελτιστοποιήσεις ποικίλης πολυπλοκότητας σε αυτές τις συμβολικές εκφράσεις. Αυτές οι βελτιώσεις περιλαμβάνουν, αλλά δεν περιορίζονται σε:

- Χρήση της GPU για υπολογισμούς
- Σταθερή αναδίπλωση
- Συγχώνευση ομοειδών υπογράφων, για να αποφευχθεί περιττός υπολογισμός
- Αριθμητική απλούστευση
- Εισάγοντας αποτελεσματική BLAS λειτουργία (π.χ. GEMM) σε μια ποικιλία περιβαλόντων
- Βελτιώσεις στην αριθμητική σταθερότητα

6.4 Caffe

Η βιβλιοθήκη Caffe⁷ είναι μια βιβλιοθήκη για μηχανική μάθηση σε εφαρμογές όρασης. Μπορεί να χρησιμοποιηθεί για τη δημιουργία βαθιών νευρωνικών δικτύων που αναγνωρίζουν αντικείμενα σε εικόνες ή ακόμα μπορούν και αναγνωρίζουν ένα οπτικό στυλ. Παρά το γεγονός ότι αυτή η βιβλιοθήκη φαίνεται να είναι ως επί το πλείστον για τους ακαδημαϊκούς και τους ερευνητές, θα μπορούσε να χρησιμοποιηθεί και άλλους τομείς όπως στην παραγωγή και στις βιομηχανίες.

6.5 Lasagne

Δεδομένου ότι η βιβλιοθήκη Theano έχει ως πρωταρχικό στόχο να είναι μια βιβλιοθήκη για συμβολικά μαθηματικά, η Lasagne⁸ προσφέρει εφαρμογές πάνω από το Theano που το

⁶<http://deeplearning.net/software/theano/>

⁷<http://caffe.berkeleyvision.org/>

⁸<https://lasagne.readthedocs.io/en/latest/>

καθιστούν πιο κατάλληλο για βαθιά εκμάθηση. Η βιβλιοθήκη είναι γραμμένη και συντηρείται κυρίως από τον Sander Dieleman, έναν σημερινό ερευνητή της DeepMind. Αντί να καθορίζει μοντέλα δικτύου όσον αφορά τις σχέσεις λειτουργίας μεταξύ των συμβολικών μεταβλητών, η Lasagne επιτρέπει στους χρήστες να σκέφτονται στο επίπεδο του Layer, προσφέροντας δομικά στοιχεία όπως "Conv2DLayer" και "DropoutLayer". Η Lasagne απαιτεί λίγη θυσία από την άποψη της ευελιξίας, παρέχοντας παράλληλα έναν πλούτο κοινών συστατικών για να βοηθήσει με τον ορισμό του στρώματος, την προετοιμασία των επιπέδων, την ταξτοποίηση των μοντέλων, την παρακολούθηση μοντέλων και την κατάρτιση μοντέλων.

6.6 Blocks

Η Blocks⁹ είναι μία βιβλιοθήκη που χρησιμοποιεί την Theano για καθαρότεροι, απλούστεροι και πιο τυποποιημένοι ορισμοί των μοντέλων βαθιάς μάθησης. Αναπτύχθηκε από το εργαστήριο του Πανεπιστημίου του Μόντρεαλ, MILA - μερικοί από τους ίδιους ανθρώπους που συνέβαλαν στο κτίριο του Theano και της πρώτης υψηλού επιπέδου διεπαφής με ορισμούς νευρωνικού δικτύου, PyLearn2. Είναι λίγο πιο ευέλικτο από τη Lasagne με το κόστος μιας ελαφρώς πιο δύσκολης καμπύλης μάθησης για αποτελεσματική χρήση. Μεταξύ άλλων, η Blocks έχει εξαιρετική υποστήριξη για επαναλαμβανόμενες αρχιτεκτονικές νευρωνικών δικτύων.

6.7 Keras



Σχήμα 6.4: Keras λογότυπο

Η Keras¹⁰ η πιο φιλική προς το χρήστη βιβλιοθήκη του τσαμπιού. Είναι γραμμένο και συντηρείται από τον Francis Chollet, άλλο μέλος της ομάδας Google Brain. Επιτρέπει στους χρήστες να επιλέξουν αν τα μοντέλα που κατασκευάζουν εκτελούνται στο συμβολικό γράφημα του Theano ή του TensorFlow. Η διεπαφή χρήστη του Keras είναι εμπνευσμένη από τη Torch. Λόγω της εξαιρετικής τεκμηρίωσης και της σχετικής ευκολίας χρήσης της, η κοινότητα

⁹<http://blocks.readthedocs.io/en/latest/>

¹⁰<https://keras.io/>

Keras είναι αρκετά μεγάλη και πολύ ενεργή. Πρόσφατα, η ομάδα TensorFlow ανακοίνωσε ότι σχεδιάζει να αποσταλεί με υποστήριξη Keras ενσωματωμένο, έτσι σύντομα Keras θα είναι ένα υποσύνολο του έργου TensorFlow.

6.8 MXNet



Σχήμα 6.5: MXNet λογότυπο

Η MXNet¹¹ είναι η βιβλιοθήκη της Amazon¹² για βαθιά εκμάθηση και είναι ίσως η πιο αποδοτική βιβλιοθήκη ανοιχτού κώδικα. Έχει ένα γράφημα ροής δεδομένων παρόμοιο με τη Theano και τη TensorFlow, προσφέρει καλή υποστήριξη για διαμορφώσεις πολλαπλών GPU, έχει δομικά στοιχεία υψηλότερου επιπέδου όπως της Lasagne και της Blocks και μπορεί να τρέχει σε σχεδόν οποιοδήποτε υλικό που μπορείτε να φανταστείτε (συμπεριλαμβανομένων των κινητών τηλεφώνων). Η υποστήριξη Python είναι ακριβώς η κορυφή του παγόβουνου. Η MXNet προσφέρει επίσης διασυνδέσεις με R, Julia, C ++, Scala, Matlab Javascript.

6.9 PyTorch



Σχήμα 6.6: Pytorch λογότυπο

Η PyTorch¹³ είναι σχετικά νέα βιβλιοθήκη βαθιάς μάθησης για την Python. Είναι μια επέκταση της βιβλιοθήκης Torch της Lua στην Python και είναι αξιοσημείωτη γιατί υποστηρίζεται από την ερευνητική ομάδα του Facebook Artificial Intelligence Research (FAIR) και επειδή έχει σχεδιαστεί για να χειρίζεται δυναμικά γράμματα υπολογισμού, ένα στοιχείο που απουσιάζει από τους Theano, TensorFlow. Είναι πολύ νωρίς για να φανεί πιο ρόλο θα διαδραματίσει η PyTorch στο οικοσύστημα βαθιάς μάθησης της Python, αλλά όλα τα σημάδια δείχνουν ότι το PyTorch είναι μια πολύ αξιολογική εναλλακτική λύση.

¹¹<http://mxnet.io/api/python/index.htmlpython-api-reference>

¹²<http://amazon.com>

¹³<http://pytorch.org/docs/>

Κεφάλαιο 7

Αναγνώριση αντικειμένων στο CDS

7.1 Το πρόβλημα

Μέχρι σήμερα το CDS έχει πάνω από 200.000 φωτογραφίες τις οποίες οι χρήστες μπορούν να αναζητήσουν μόνο εφόσον γνωρίζουν τον τίτλο της συλλογής. Αυτό πολλές φορές καθιστά πολύ δύσκολη έως αδύνατη την αναζήτηση φωτογραφιών με συγκεκριμένα χαρακτηριστικά, όπως για παράδειγμα φωτογραφίες που έχουν φορητούς υπολογιστές ή αυτοκίνητα.

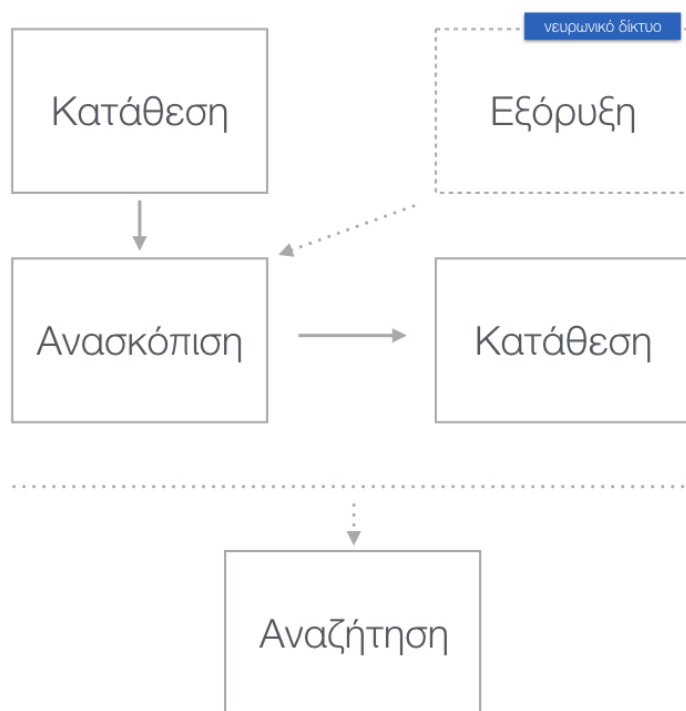
Για την επίλυση αυτού του προβλήματος θα χρησιμοποιηθούν νευρωνικά δίκτυα έτσι ώστε κάθε φωτογραφία που ανεβαίνει στο CDS να αναλύεται, να εξάγονται τα αντικείμενα και να αποθηκεύονται σαν ετικέτες στο τελικό αντικείμενο.

Με αυτό τον τρόπο, με την βοήθεια της ElasticSearch¹ που ενσωματώνεται στο Invenio, οι χρήστες θα μπορούν όχι μόνο να αναζητήσουν φωτογραφίες με συγκεκριμένο αντικείμενο αλλά και να συνδυάσουν αντικείμενα ταυτόχρονα.

7.2 Σχεδίαση

Για τη λύση του προβλήματος, όπως φαίνεται και στο σχήμα 7.1, ο χρήστης θα ανεβάζει μία φωτογραφία, μόλις ανεβεί θα στέλνεται στο νευρωνικό δίκτυο για την εξόρυξη των ετικετών των αντικειμένων και έπειτα ο χρήστης θα πλοηγείται στην σελίδα της ανασκόπησης. Από αυτή τη σελίδα θα μπορεί να κάνει κατάθεση των δεδομένων και στην συνέχεια να τα αναζητήσει βάση των ετικετών.

¹Open Source Search & Analytics <https://www.elastic.co/>



Σχήμα 7.1: Ροή εργασίας μονάδας

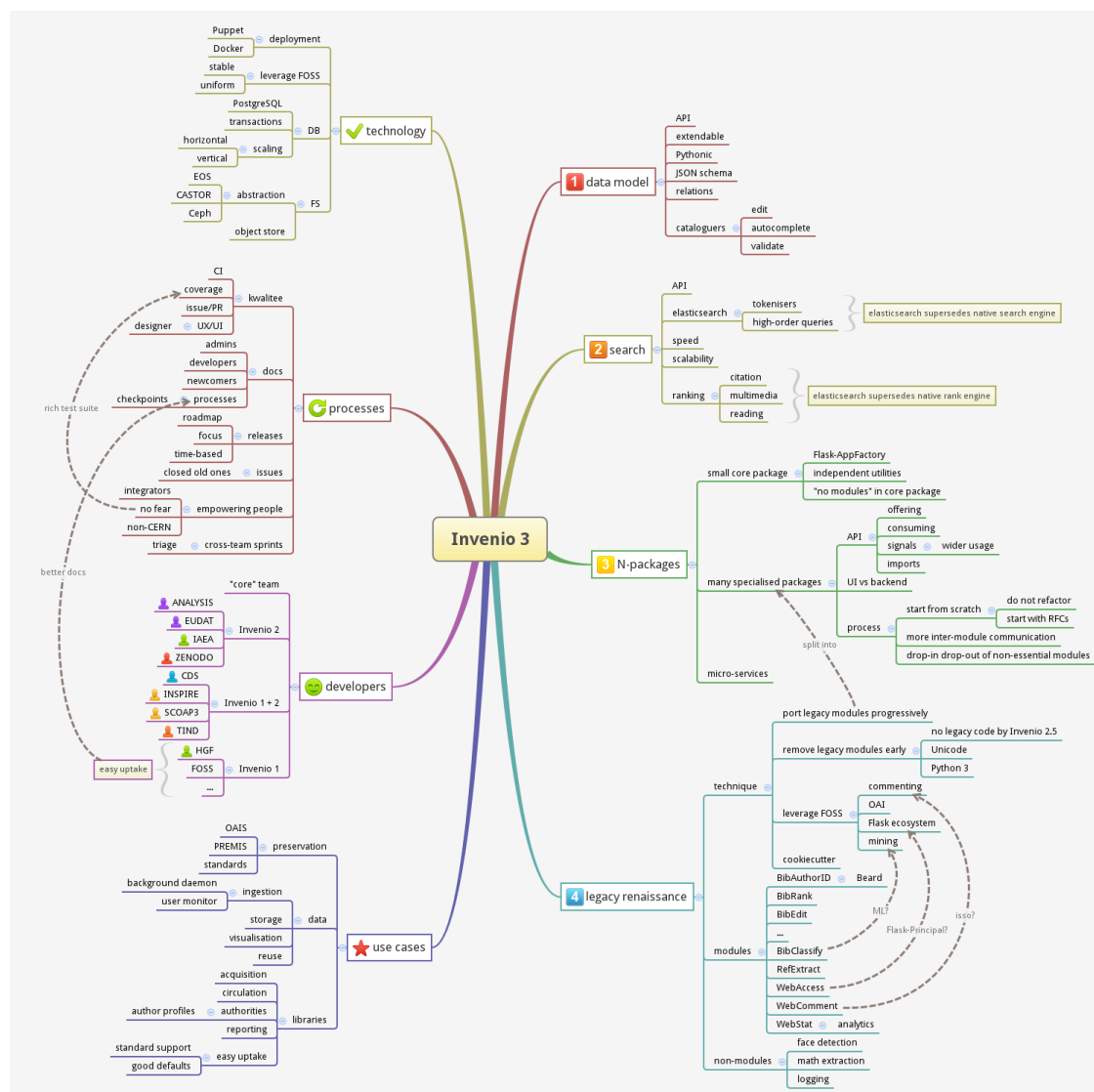
Θα δημιουργηθεί μια νέα μονάδα η οποία θα υλοποιεί το νευρωνικό δίκτυο καθώς και το RESTful API για την επικοινωνία της διεπαφής των χρηστών με αυτό. Το Invenio Software v3, όπως φαίνεται από το σχήμα 7.2, αποτελείται από πάρα πολλές μονάδες, όμως για την καινούρια μονάδα χρειάζονται μόνο τρεις.

7.2.1 Μονάδα αναζήτησης

Η μονάδα αυτή επιτρέπει την αναζήτηση δεδομένων στο Invenio Software v3. Η μηχανή αναζήτησης βασίζεται στην τεχνολογία Elasticsearch και παρέχει μια κατανεμημένη μηχανή αναζήτησης πλήρους κειμένου που διαθέτει κατανεμημένη δυνατότητα πολλαπλών λειτουργιών, με διεπαφή ιστού HTTP και έγγραφα JSON χωρίς σχήματα.

7.2.2 Μονάδα καταθέσεων

Η μονάδα αυτή επιτρέπει την καταχώρηση νέων δεδομένων και εγγράφων στη βάση δεδομένων. Τα μεταδεδομένα καταχωρούνται σε μορφή JSON και στέλνονται έπειτα από κάθε αποθήκευση στη μονάδα της Αναζήτησης όπου γίνονται άμεσα διαθέσιμα για αναζήτηση.



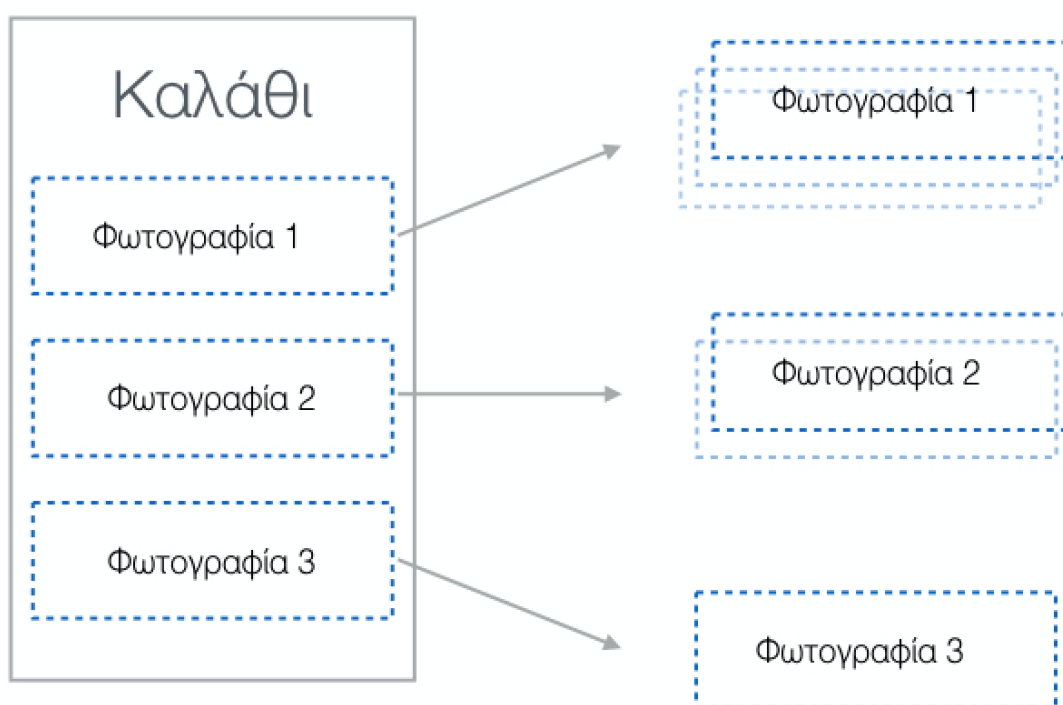
Σχήμα 7.2: Όλες οι μονάδες του Invenio Software v3

7.2.3 Μονάδα αρχείων

Η μονάδα αποθήκευσης αρχείων του Invenio Software v3 διαθέτει μια απλή διεπαφή υπηρεσιών ιστού που μπορεί να χρησιμοποιηθεί για την αποθήκευση και την ανάκτηση οποιουδήποτε όγκου αρχείων ανά πάσα στιγμή από οπουδήποτε στον ιστό.

Οι δύο βασικές έννοιες της μονάδας αρχείων είναι:

- Καλάθι (Bucket). Είναι μία λογική μονάδα αποθήκευσης η οποία χρησιμοποιείται για την αποθήκευση αντικειμένων και των μεταδεδομένων τους. Επίσης ορίζει τα δικαιώματα πρόσβασης για τα αντικείμενα που ανήκουν σ' αυτή.
- Αντικείμενο (Object). Είναι μια λογική μονάδα η οποία χρησιμοποιείται για να πε-



Σχήμα 7.3: Καλάθι και αντικείμενα στη μονάδα αρχείων

ριγράφει τα αρχεία και τα μεταδεδομένα τους. Κάθε αντικείμενο αντιστοιχεί σε ένα αρχείο.

Στην μονάδα του νευρωνικού δικτύου θα χρησιμοποιηθεί η βιβλιοθήκη Tensorflow, για τους παρακάτω λόγους:

- Παραλληλισμός σε μεμονωμένες λειτουργίες. Οι λειτουργίες αυτές έχουν αποδοτικές παράλληλες υλοποιήσεις για επεξεργαστές πολλαπλών πυρήνων και GPU και το Tensorflow χρησιμοποιεί αυτές τις εφαρμογές όπου είναι διαθέσιμες.
- Παραλληλισμός μεταξύ πράξεων. Το Tensorflow χρησιμοποιεί μια παράσταση γραφικής παράστασης δεδομένων για το μοντέλο όπου υπάρχουν δύο κόμβοι που δεν συνδέονται με μια κατευθυνόμενη διαδρομή στο γράφημα ροής δεδομένων, οι πράξεις αυτές μπορούν να εκτελούνται παράλληλα. Για παράδειγμα, το μοντέλο αναγνώρισης εικόνων Inception έχει πολλούς παράλληλους κλάδους στο γράφημα ροής δεδομένων του και το Tensorflow μπορεί να το εκμεταλλευτεί για να τρέξει πολλές λειτουργίες ταυτόχρονα. Στη μελέτη του AlexNet² περιγράφεται επίσης ο τρόπος χρήσης του «παραλληλισμού μοντέλου» για την εκτέλεση παράλληλων λειτουργιών σε διάφορα μέρη του μοντέλου και το Tensorflow υποστηρίζει ότι χρησιμοποιεί τον ίδιο μηχανισμό.

²<https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks>

- Παραλληλισμός μεταξύ αντιγράφων μοντέλων. Το Tensorflow σχεδιάστηκε επίσης για κατανεμημένη εκτέλεση. Ένα κοινό σύστημα παράλληλης εκπαίδευσης («Παραλληλισμός των δεδομένων») περιλαμβάνει τη συσσώρευση του συνόλου δεδομένων σε ένα σύνολο πανομοιότυπων εργασιών, την εκτέλεση των ίδιων υπολογισμών σε κάθε μία από αυτές για διαφορετικά δεδομένα και την κατανομή των παραμέτρων του μοντέλου μεταξύ των εργασιών.

Επίσης λόγω της δημοφιλότητας της συγκεκριμένης βιβλιοθήκης υπάρχει μια ενεργή κοινότητα η οποία αποτελεί κύριο πυλώνα της μακροβιότητας της.

7.3 Εκπαίδευση του νευρωνικού δικτύου

Πριν γίνει η προγραμματιστική υλοποίηση της λύσης, το νευρωνικό δίκτυο θα πρέπει να εκπαιδευτεί για όσο τον δυνατόν περισσότερα αντικείμενα. Σύμφωνα με τον πίνακα, υπάρχουν πολλές συλλογές δεδομένων οι οποίες μπορούν να χρησιμοποιηθούν 7.1

Συλλογή	Αντικείμενα	Κλάσεις
Flowers	2500	5
Cifar10	60.000	10
MNIST	60.000	10
ImageNet	12.000.000	1000

Πίνακας 7.1: Συλλογές δεδομένων

Για παράδειγμα, η συλλογή Flowers³ αποτελείται από 2500 διαφορετικά είδη λουλουδιών όπου κάθε είδος έχει 5 διαφορετικές φωτογραφίες (κλάσεις) οι οποίες χρησιμεύουν στην αναγνώριση του συγκεκριμένου είδους.

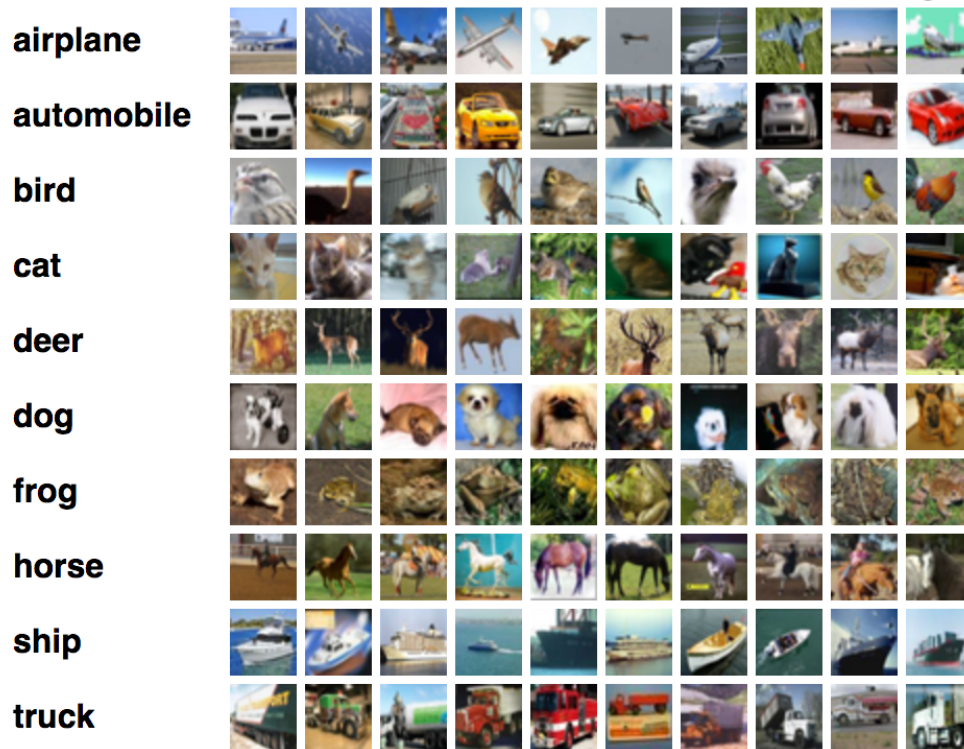
Λόγω του εύρους φωτογραφιών στο CDS, χρησιμοποιήθηκε η μεγαλύτερη συλλογή ImageNet⁴, η οποία περιέχει 12 εκατομμύρια αντικείμενα και 1000 κλάσεις που βασίζονται στην ιεραρχία του WordNet⁵.

Το ImageNet είναι μια μεγάλη οπτική βάση δεδομένων σχεδιασμένη για χρήση στην έρευνα λογισμικού αναγνώρισης οπτικών αντικειμένων. Από το 2016, πάνω από δέκα εκατομμύρια διευθύνσεις URL εικόνων έχουν επισημανθεί από χρήστες για να υποδείξουν τα αντικείμενα που απεικονίζονται. Σε τουλάχιστον ένα εκατομμύριο εικόνες παρέχεται επίσης οριοθέτηση. Η βάση δεδομένων των σχολιασμών είναι ελεύθερα διαθέσιμη απευθείας από τη ιστοσελίδα. Ωστόσο, οι πραγματικές εικόνες δεν ανήκουν στο ImageNet.

³http://download.tensorflow.org/example_images/flower_photos.tgz

⁴ImageNet <http://www.image-net.org/>

⁵WordNet <https://wordnet.princeton.edu/>



Σχήμα 7.4: Παράδειγμα αντικειμένων και κλάσεων

Για την εκπαίδευση ακολουθήθηκαν τα παρακάτω βήματα:

```

1 # http://image-net.org/download
2 DATA_DIR=$HOME/imagenet-data
3 cd tensorflow-models/inception
4 bazel build //inception:download_and_preprocess_imagenet
5 bazel-bin/inception/download_and_preprocess_imagenet "${DATA_DIR}"

```

Αλγόριθμος 7.1: Εκπαίδευση

Ο χρόνος εκπαίδευσης κυμάνθηκε στα 30 λεπτά και ουσιαστικά κατά την εκπαίδευση η βιβλιοθήκη δημιούργησε το νευρωνικό δίκτυο. Έπειτα από αυτό το βήμα μπορεί να γίνει η υπόλοιπη υλοποίηση της μονάδας.

7.4 Υλοποίηση

Για τη καλύτερη ενσωμάτωση με το CDS δημιουργήθηκε η μονάδα invenio-objectrecognition Invenio που περιέχει τις διεπαφές που δημιουργήσαμε. Η καινούργια μονάδα δημιουργήθηκε με

τη χρήση του cookiecutter⁶ το οποίο είναι μία command-line βιβλιοθήκη που δημιουργεί πακέτα μέσω διαφόρων προτύπων. Για τη συγκεκριμένη περίπτωση χρησιμοποιήθηκε το πρότυπο cookiecutter-invenio-module⁷. Η διαδικασία για τη δημιουργία της μονάδας είναι πολύ απλή:

```
1 cookiecutter https://github.com/inveniosoftware/cookiecutter-invenio-module
```

Αλγόριθμος 7.2: Δημιουργία μονάδας

Στη συνέχεια συνδέθηκε η μονάδα με το CDS χρησιμοποιώντας τα σημεία εισόδου entry points της Python. Το Invenio προσφέρει πολλά σημεία εισόδου όπου το καθένα αντιστοιχεί σε κάποια διαφορετική λειτουργία του. Για παράδειγμα το *invenio_base.blueprints* αντιστοιχεί στα τελικά σημεία (endpoints) των μονάδων. Για τη μονάδα invenio-objectrecognition προστέθηκαν τα παρακάτω:

```
1 entry_points={
2     'invenio_base.apps': [
3         'invenio_objectrecognition = invenio_objectrecognition:
4         InvenioObjectRecognition',
5     ],
6     'invenio_base.api_blueprints': [
7         'invenio_objectrecognition = invenio_objectrecognition.views:
8         blueprint',
9     ],
10    'invenio_i18n.translations': [
11        'messages = invenio_objectrecognition',
12    ],
13    'invenio_search.mappings': [
14        'images = invenio_objectrecognition.mappings',
15    ],
16    'invenio_jsonschemas.schemas': [
17        'images = invenio_objectrecognition.jsonschemas',
18    ],
19 }
```

Αλγόριθμος 7.3: Σημεία Εισόδου

Για την επίλυση, το πρόβλημα χωρίστηκε σε τρία μέρη, το πρώτο είναι η διεπαφή προγραμματισμού εφαρμογών (API), το δεύτερο είναι η διεπαφή ανεβάσματος φωτογραφίας και τέλος είναι η διεπαφή αναζήτησης.

Οι διεπαφές αναζήτησης και ανεβάσματος φωτογραφίας είναι σχεδιασμένες για να παρέχουν όσο το δυνατόν περισσότερη ευκολία στη χρήση καθώς αποτελούν τις βασικές σελίδες αλληλεπίδρασης με την εφαρμογή. Σε πολύ γρήγορο χρόνο ο χρήστης μπορεί να βρει τις διαθέσιμες

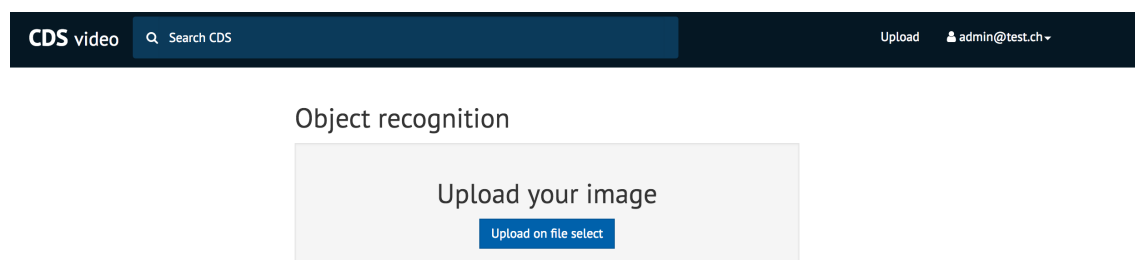
⁶<https://github.com/audreyr/cookiecutter>

⁷<https://github.com/inveniosoftware/cookiecutter-invenio-module>

ενέργειες για την κάθε σελίδα. Επιπρόσθετα, η διεπαφή προγραμματισμού εφαρμογών σχεδιάστηκε έτσι ώστε οι προγραμματιστές να μπορούν να χρησιμοποιήσουν και να επεκτείνουν τη διεπαφή στις διαφορετικές πλατφόρμες και εφαρμογές.

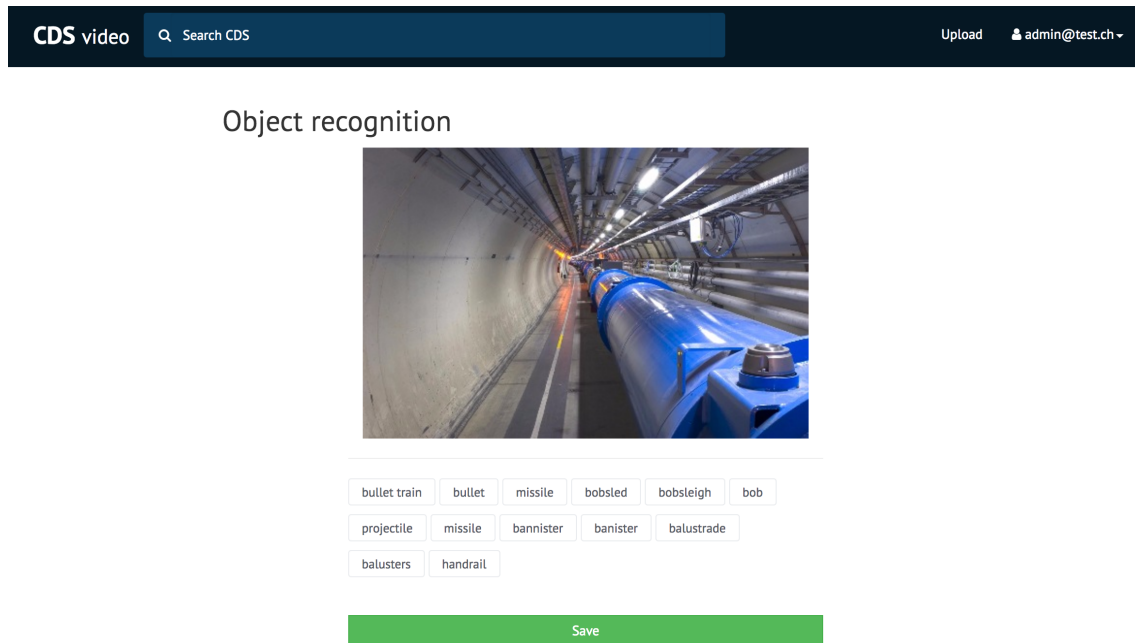
7.4.1 Διεπαφή ανεβάσματος φωτογραφίας

Σ' αυτή τη διεπαφή ο χρήστης ανεβάζει φωτογραφίες από τον τοπικό υπολογιστή. Στην συνέχεια, το σύστημα ανεβάζει τη φωτογραφία στο διακομιστή και επικοινωνεί με τη διεπαφή προγραμματισμού εφαρμογών ώστε να αρχίσει η εξόρυξη αντικειμένων από αυτή.



Σχήμα 7.5: Διεπαφή για το ανέβασμα των φωτογραφιών

Μόλις η διαδικασία ολοκληρωθεί, εμφανίζονται στην οθόνη μια προβολή της εικόνας που ανέβηκε καθώς και οι ετικέτες που επέστρεψε η διεπαφή προγραμματισμού εφαρμογών έπειτα από ανάλυση στο νευρωνικό δίκτυο.



Σχήμα 7.6: Διεπαφή για προβολή ετικετων

Στη διεπαφή υπάρχει επίσης η επιλογή για αποθήκευση της φωτογραφίας με τις ετικέτες όπου αν πατηθεί αυτομάτως η φωτογραφία θα γίνει διαθέσιμη και τα φίλτρα αναζήτησης θα ανανεωθούν με τις τυχόν καινούργιες ετικέτες. Η αποθήκευση γίνεται μέσω της διεπαφής προγραμματισμού εφαρμογών του Invenio που επιτρέπει την επικοινωνία της διεπαφής με την βάση δεδομένων και την Elasticsearch⁸ που χρησιμοποιείται από το Invenio για την αναζήτηση δεδομένων.

7.4.2 Διεπαφή προγραμματισμού εφαρμογών

Γι' αυτή τη διεπαφή χρησιμοποιήθηκε RESTful API ώστε η επικοινωνία με την διεπαφή ανεβάσματος φωτογραφίας να γίνεται μέσω HTTP πρωτόκολλου. Σ αυτή τη διεπαφή χρειάζεται να δοθεί το όνομα της φωτογραφίας ώστε να βρεθεί η τοποθεσία στο διακομιστή και έπειτα να επεξεργαστεί και αναλυθεί από τη βιβλιοθήκη Tensorflow στο νευρωνικό δίκτυο.

```

1 @staticmethod
2 def get_tags(file_path):
3     """Get tags for image."""
4     if not tf.gfile.Exists(file_path):
5         raise ObjectRecognitionDoesNotExist

```

⁸ <http://elastic.co/>

```

6
7 image_data = tf.gfile.FastGFile(file_path, 'rb').read()
8 create_graph()
9
10 with tf.Session() as sess:
11     softmax_tensor = sess.graph.get_tensor_by_name('softmax:0')
12     predictions = sess.run(
13         softmax_tensor, {'DecodeJpeg/contents:0': image_data}
14     )
15     predictions = np.squeeze(predictions)
16     node_lookup = NodeLookup()
17     # sort the predictions
18     top_k = predictions.argsort()[-5:][: -1]
19     # map to the friendly names and return the tuples
20     return [
21         node_lookup.id_to_string(node_id) for node_id in top_k
22     ]
23
24 @error_handler
25 def get(self, bucket_id=None, filename=None):
26     """Handle GET request."""
27     try:
28         file_path = ObjectVersion.get(bucket_id, filename).file.uri
29     except AttributeError:
30         raise ObjectRecognitionDoesNotExist
31
32     tags = self.get_tags(file_path)
33
34     return jsonify(
35         dict(
36             tags=tags,
37             filename=filename
38         )
39     )

```

Αλγόριθμος 7.4: Διεπαφή Προγραμματισμού Εφαρμογών

Η παραπάνω διεπαφή απαντάει σε ερωτήματα GET του πρωτοκόλλου HTTP μέσω της διεύθυνσης `/api/image/recognition/` < path : image_image >

Η απάντηση της διεπαφής έχει τις παρακάτω μορφές που εξαρτώνται από τα HTTP Status Code:

- **200** Το όνομα της φωτογραφίας βρέθηκε και επιστρέφει τα αντικείμενα που βρέθηκαν σε μορφή ετικετών.
- **404** Δεν βρέθηκε η ζητούμενη φωτογραφία.
- **500** Υπήρξε κάποιο τεχνικό πρόβλημα στη διαδικασία.

Για παράδειγμα, εάν χρησιμοποιηθεί στη διεπαφή μία φωτογραφία από ένα ζώο Πάντα όπως είναι η παρακάτω, το σύστημα θα πρέπει να επιστρέφει τουλάχιστον μια ετικέτα με το όνομα Panda.



Σχήμα 7.7: Πάντα στο πάρκο

Πράγματι, καλώντας την διεπαφή το αποτέλεσμα είναι το ακόλουθο:

```
1 {  
2   "filename": "panda.jpg",  
3   "tags": [  
4     "giant panda, panda, panda bear, coon bear",  
5     "lesser panda, red panda, panda, bear cat",  
6     "brown bear, bruin, Ursus arctos",  
7     "earthstar",  
8     "ice bear, polar bear, Ursus Maritimus"  
9   ]  
10 }
```

Αλγόριθμος 7.5: Απάντηση ερωτήματος

Είναι εμφανές ότι περιέχει παραπάνω από μια ετικέτες που περιγράφει το Panda.

Όπως είναι φυσικό η βιβλιοθήκη Tensorflow επιστρέφει αποτελέσματα από το νευρωνικό δίκτυο τα οποία πρέπει μετά να μετατραπούν σε ανθρωπίνως αναγνώσιμα. Γί αυτό το λόγο,

δημιουργήθηκε μία κλάση η οποία αποτελεί τον μεταφραστή μεταξύ του νευρωνικού δικτύου και του ανθρώπου.

```

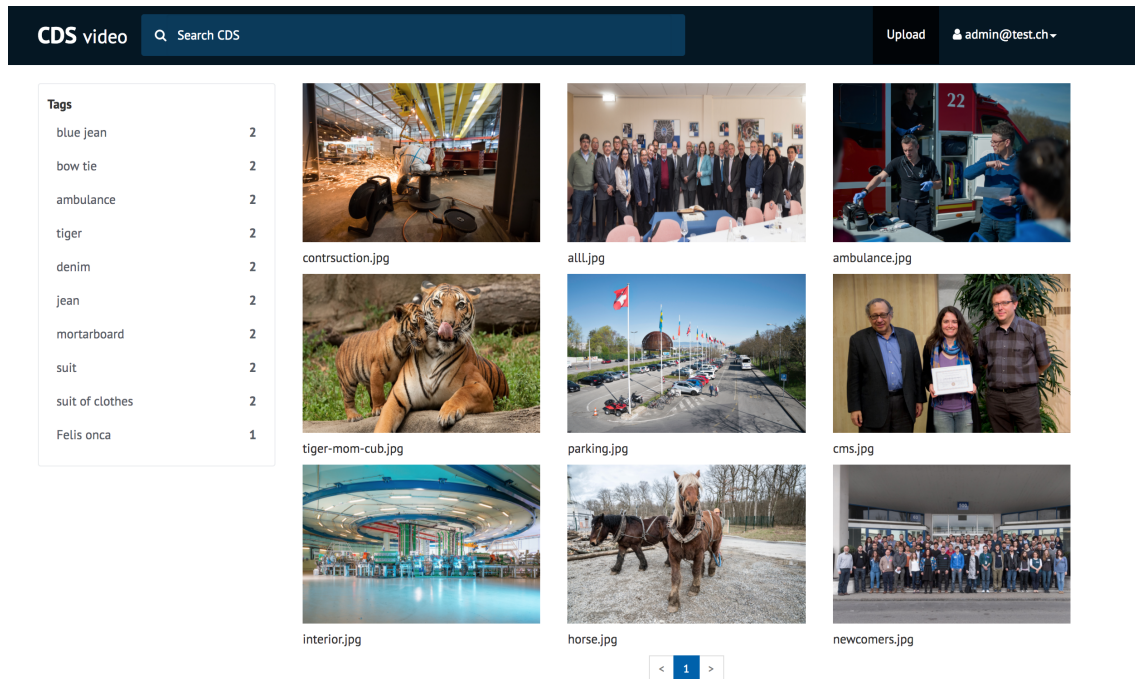
1 class Node(object):
2     """Converts nodes to labels."""
3
4     def __init__(self, label_lookup_path, uid_lookup_path):
5         """Initialize the lookup."""
6         self.node_lookup = self.load(label_lookup_path, uid_lookup_path)
7
8     def load(self, label_lookup_path, uid_lookup_path):
9         """Loads the matching classes"""
10        proto_as_ascii_lines = tf.gfile.GFile(uid_lookup_path).readlines()
11        uid_to_human = {}
12        p = re.compile(r'[n\d]*[ \S,]*')
13        for line in proto_as_ascii_lines:
14            parsed_items = p.findall(line)
15            uid = parsed_items[0]
16            human_string = parsed_items[2]
17            uid_to_human[uid] = human_string
18
19        node_id_to_uid = {}
20        proto_as_ascii = tf.gfile.GFile(label_lookup_path).readlines()
21        for line in proto_as_ascii:
22            if line.startswith('  target_class:'):
23                target_class = int(line.split(':')[1])
24            if line.startswith('  target_class-string:'):
25                target_class_string = line.split(':')[1]
26                node_id_to_uid[target_class] = target_class_string[1:-2]
27
28        node_id_to_name = {}
29        for key, val in node_id_to_uid.items():
30            if val in uid_to_human:
31                name = uid_to_human[val]
32                node_id_to_name[key] = name
33
34        return node_id_to_name
35
36    def id_to_string(self, node_id):
37        """Node id to matched human string."""
38        return self.node_lookup[node_id]

```

Αλγόριθμος 7.6: Κλάση μετάφρασης

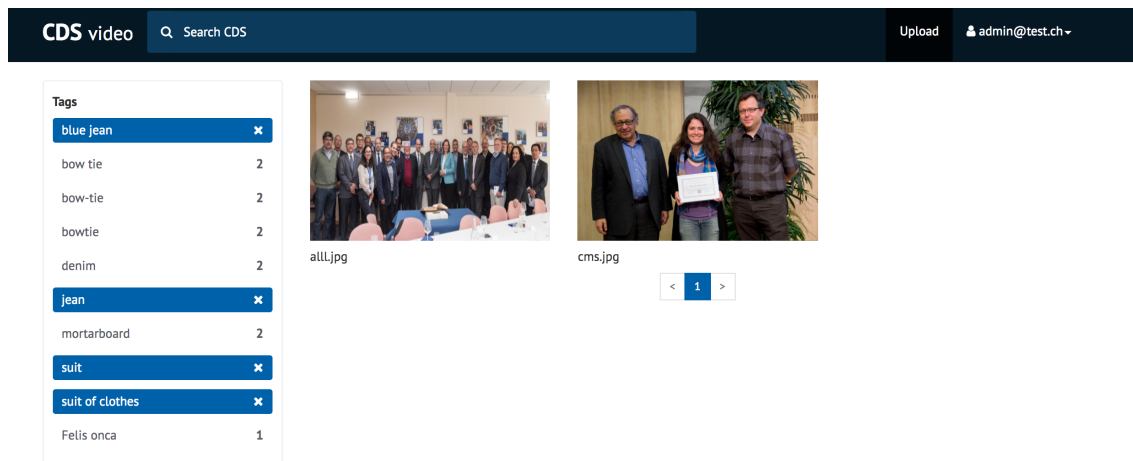
7.4.3 Διεπαφή της αναζήτησης

Σ' αυτή τη διεπαφή ο χρήστης μπορεί να αναζητήσει φωτογραφίες βάση του περιεχομένου τους (αρχικός στόχος της εργασίας). Στον αριστερό τομέα του γραφικού περιβάλλοντος έχει τοποθετηθεί ένα εργαλείο το οποίο, επιλέγοντας κάποια ετικέτα, ψάχνει σε όλες τις φωτογραφίες και βρίσκει αυτές που την περιέχουν.



Σχήμα 7.8: Διεπαφή για προβολή ετικετών

Παράλληλα, δίνεται η δυνατότητα να επιλεχτούν πολλές ετικέτες ταυτόχρονα. Το σύστημα εκτελεί μια λογική αναζήτηση και επιστρέφει φωτογραφίες που περιέχουν κάποια από τις επιλεγμένες ετικέτες. Στο παρακάτω παράδειγμα έχει γίνει επιλογή φωτογραφιών που περιέχουν **blue jeans**, **jean**, **suit**, **suit of clothes** και το σύστημα επιστρέφει φωτογραφίες από ανθρώπους που φοράνε αυτό το είδος ρούχων.



Σχήμα 7.9: Διεπαφή για προβολή ετικετων

Δεδομένου ότι το σύστημα θα πρέπει να είναι εύκολα επεκτάσιμο, η διεπαφή δίνει την δυνατότητα σε άλλους προγραμματιστές να έχουν πρόσβαση σ' αυτή τη μεγάλη βάση μέσω RESTful API και να μπορούν να κάνουν ερωτήματα για συγκεκριμένες ετικέτες. Η απάντηση που δίνει το σύστημα είναι σε μορφή JSON και έχει την παρακάτω μορφή:

```

1 {
2   "results": [
3     {
4       "tags": [
5         "carousel",
6         "carrousel",
7         "merry-go-round",
8         "roundabout",
9         "whirligig",
10        "West Highland white terrier",
11        "palace",
12        "stage",
13        "steel arch bridge"
14      ],
15      "title": "interior.jpg"
16    },
17    {
18      "tags": [

```

```
19         "horse cart" ,  
20         "horse-cart" ,  
21         "sorrel" ,  
22         "plow" ,  
23         "plough" ,  
24         "cowboy hat" ,  
25         "ten-gallon hat" ,  
26         "barrel" ,  
27         "cask"  
28     ] ,  
29     "title": "horse.jpg"  
30 }  
31 ]  
32 }  
33 ]  
34 }
```

Αλγόριθμος 7.7: Απάντηση ερωτήματος αναζήτησης

Κεφάλαιο 8

Επίλογος

Με βάση τις ιδιότητες και τον τρόπο λειτουργίας του λογισμικού CDS που αναπτύσσεται στο CERN, αναλύθηκε η σχεδίαση και η υλοποίηση νευρωνικού δικτύου το οποίο αναγνωρίζει αντικείμενα και δημιουργεί ετικέτες όταν ο χρήστης ανεβάζει φωτογραφίες στο σύστημα. Επίσης παρουσιάστηκαν τρόποι με τους οποίους άλλοι προγραμματιστές μπορούν να χρησιμοποιήσουν το CDS σύστημα. Με τη δημιουργία της μονάδας νενιο-οβιδεστρεσογνιτιον, η κοινότητα ελεύθερου λογισμικού και ψηφιακών βιβλιοθηκών θα μπορεί να επεκτείνει και να δημιουργήσει νέες υπηρεσίες με στόχο την εξέλιξη της.

8.1 Συμπεράσματα

Είμαστε στη μέση μιας ιστορικής στιγμής. Παλαιότερα ένας υπολογιστής θα έπρεπε να προγραμματιστεί έτσι ώστε να γνωρίζει το πως να εκτελέσει συγκεκριμένες εργασίες. Τώρα οι υπολογιστές μπορούν να μάθουν από την εμπειρία. Η μηχανική μάθηση είναι αφάνταστα σημαντική για την κατανόηση του που βαδίζει η τεχνολογία, καθώς και η κοινωνία χρησιμοποιώντας την.

Οι ταινίες που προτείνει το Netflix, οι συστάσεις προϊόντων της Amazon, η ικανότητά του Facebook να εντοπίζει τους φίλους ως πρόσωπα, οι εφαρμογές που υπόσχονται να βρουν το κατάλληλο ταίρι - όλα αυτά είναι πρώιμα παραδείγματα της μηχανικής μάθησης.

Αλλά και το αυτόνομο αυτοκίνητο της Google αποτελεί πλέον ένα κλασσικό παράδειγμα. «Ένα αυτόνομο αυτοκίνητο δεν είναι προγραμματισμένο να οδηγεί μόνο του» αναφέρει ο επιστήμονας Pedro Domingos¹ από το Πανεπιστήμιο της Ουάσιγκτον.

«Κανείς δεν ξέρει πραγματικά πώς να προγραμματίσει ένα αυτοκίνητο ώστε να μπορεί να

¹<https://homes.cs.washington.edu/~pedrod/>

οδηγήσει», λέει. «Γνωρίζουμε πώς να οδηγούμε, αλλά δεν μπορούμε καν να το εξηγήσουμε στον εαυτό μας. Το Google αυτοκίνητο μαθαίνει οδηγώντας εκατομμύρια μίλια και παρατηρώντας τους ανθρώπους εν ώρα οδήγησης.»

Αυτό είναι το κλειδί: η μηχανική μάθηση επιτρέπει στους αλγόριθμους να μάθουν μέσα από την εμπειρία, και να κάνουν πράγματα που τα προγράμματα δεν γνωρίζουν πως να τα κάνουν. Για τον Domingos, η μηχανική μάθηση είναι τόσο σημαντική όσο και η εφεύρεση των προσωπικών υπολογιστών, το διαδίκτυο, ή την ίδια την ηλεκτρική ενέργεια.

«Υπήρχαν δύο στάδια στην εποχή της πληροφορίας», λέει ο Domingos. «Ένα στάδιο όπου πρέπει να προγραμματίζουμε υπολογιστές, και το δεύτερο στάδιο, το οποίο αρχίζει τώρα, είναι αυτό όπου οι υπολογιστές μπορούν να προγραμματίσουν με την εξέταση των δεδομένων.»

Από τα παραπάνω, συμπεραίνουμε ότι η μηχανική Μάθηση δεν είναι μια νέα προσέγγιση, αλλά ένας ταχέως αναπτυσσόμενος κλάδος. Αρωγός αυτής της ανάπτυξης αποτελεί ένας συνδυασμός παραγόντων – οι βελτιωμένες μέθοδοι κατάρτισης, η αυξανόμενη ισχύς των υπολογιστών, τα τεράστια σύνολα δεδομένων που εκθέτουν τις αδυναμίες μας, η βελτιωμένη χρηστικότητα των εργαλείων Μηχανικής Μάθησης.

Η πραγματική αξία της Μηχανικής Μάθησης είναι ότι ανοίγει μια εντελώς νέα τεχνική για την επίλυση των προβλημάτων που έχουν αποδειχθεί δυσεπίλυτα σε πολλές περιπτώσεις. Μπορούμε πλέον να επανεξετάσουμε προκλήσεις που προηγουμένως φαινόταν άπιαστες.

8.2 Μελλοντικές επεκτάσεις

Η παρούσα εργασία αποτελεί μια αξιοποιήσιμη λύση για την αναζήτηση φωτογραφιών μέσω των αντικειμένων που απεικονίζουν, όμως επιδέχεται βελτιώσεις και επεκτάσεις. Μία μεγαλύτερη συλλογή θα βοηθούσε για ακόμα πιο ακριβή αποτελέσματα όπου οι ετικέτες στα φίλτρα θα είναι ευκολότερες στη χρήση. Μία κεντρική βιβλιοθήκη φωτογραφιών για όλα τα πειράματα που πραγματοποιούνται στο CERN θα βοηθούσε στην εύκολη εύρεση εξαρτημάτων και εξοπλισμού ανεβάζοντας απλά φωτογραφία από το τομέα του επιταχυντή που πρέπει επισκευαστεί.

Ακόμα η επέκταση του αλγορίθμου για την αναγνώριση αντικειμένων θα μπορούσε να γίνει και σε βίντεο, όπου βάση των κυρίων στιγμιοτύπων θα μπορούσαν να αναγνωρισθούν τα αντικείμενα και μέσω της διεπαφή της αναζήτησης να μπορούν να αναζητηθούν βάση ετικετών.

Βιβλιογραφία

- [1] DSpace. <http://www.dspace.org/>
- [2] Fedora <http://www.fedora-commons.org/>
- [3] Greenstone <http://www.greenstone.org/>
- [4] Eprint <http://www.eprints.org/>
- [5] Tensorflow <https://www.tensorflow.org/>
- [6] Scikit-learn <http://scikit-learn.org/stable/>
- [7] Theano <http://deeplearning.net/software/theano/>
- [8] Caffe <http://caffe.berkeleyvision.org/>
- [9] <http://el.wikipedia.org/wiki/CERN>
- [10] https://en.wikipedia.org/wiki/Particle_physics
- [11] https://en.wikipedia.org/wiki/Digital_library
- [12] <http://cds.cern.ch/help/>
- [13] <http://invenio-software.org/>

Γλωσσάριο

Ελληνικός όρος

μηχανική μάθηση
βαθιά μάθηση
αποτίμηση ερωτημάτων
δειγματοληψία
δεικτοδότηση
ερώτημα διαρκείας
ερώτημα εγγύτερου γείτονα
ιδιωτικότητα
κάνναβος
κινούμενο αντικείμενο
παράθυρο
πολυπλεξία
ρεύμα δεδομένων
σημειακή εστία
συνάθροιση
σύνδεση
φιλτράρισμα
χρονόσημο

Αγγλικός όρος

machine learning
deep learning
query evaluation
sampling
indexing
continuous query
nearest-neighbor query
privacy
grid
moving object
window
multiplexing
data stream
focal point
aggregation
join
filtering
timestamp

