



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ
ΥΠΟΛΟΓΙΣΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Συγκριτική Μελέτη εργαλείων Γενετικής Τεχνητής Νοημοσύνης
Comparative Study of Genetic Artificial Intelligence tools

ΒΙΚΤΩΡ ΛΕΜΠΩ
A.M. ice18390300

Εισηγητής: ΧΡΗΣΤΟΣ ΤΡΟΥΣΣΑΣ, ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΣΕΠΤΕΜΒΡΙΟΣ 2024

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Συγκριτική Μελέτη εργαλείων Γενετικής Τεχνητής Νοημοσύνης

**ΒΙΚΤΩΡ ΛΕΜΠΩ
Α.Μ. ice18390300**

Εισηγητής:

ΧΡΗΣΤΟΣ ΤΡΟΥΣΣΑΣ, ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ

Εξεταστική Επιτροπή:

**Π. ΤΣΕΛΕΝΤΗ, ΜΕΛΟΣ ΕΔΙΠ
Α. ΚΡΟΥΣΚΑ, ΜΕΛΟΣ ΕΔΙΠ**

Ημερομηνία εξέτασης

ΣΕΠΤΕΜΒΡΙΟΣ 2024

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Ο κάτωθι υπογεγραμμένος Βίκτωρ Λεμπώ με αριθμό μητρώου 18390300 φοιτητής του Τμήματος Μηχανικών Πληροφορικής και Υπολογιστών της Σχολής Μηχανικών του Πανεπιστημίου Δυτικής Αττικής, δηλώνω ότι:

«Βεβαιώνω ότι είμαι συγγραφέας αυτής της Διπλωματικής εργασίας και κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

«Βεβαιώνω ότι είμαι συγγραφέας της παρούσας διπλωματικής εργασίας και ότι έχω αναφέρει ή παραπέμψει σε αυτή, ρητά και συγκεκριμένα, όλες τις πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών, προτάσεων ή λέξεων, είτε αυτές μεταφέρονται επακριβώς (στο πρωτότυπο ή μεταφρασμένες) είτε παραφρασμένες. Επίσης βεβαιώνω ότι αυτή η εργασία προετοιμάστηκε από εμένα προσωπικά ειδικά για την συγκεκριμένη διπλωματική εργασία»

Ο Δηλών

Βίκτωρ Λεμπώ

Ευχαριστίες

Η παρούσα διπλωματική εργασία ολοκληρώθηκε μετά από επίπονες προσπάθειες σε ένα πολύ ενδιαφέρον γνωστικό αντικείμενο.

Την επιλογή και την προσπάθειά μου αυτή υποστήριξε ο επιβλέπων καθηγητής μου κ. Χρήστος Τρούσσας, Επίκουρος Καθηγητής στο Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών του Πανεπιστημίου Δυτικής Αττικής, στον οποίον εκφράζω ένα μεγάλο ευχαριστώ και την βαθιά ευγνωμοσύνη μου για την καλή του πάντα διάθεση, την κατανόηση και την πολύτιμη συμπαράστασή του στη διαδρομή των σπουδών μου που ολοκληρώνονται με την εκπόνηση αυτής της εργασίας.

Επίσης οφείλω ένα πολύ μεγάλο ευχαριστώ σε όλους τους καθηγητές μου στο Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών του Πανεπιστημίου Δυτικής Αττικής για τη βοήθεια και αμέριστη υποστήριξή τους χωρίς την οποία θα ήταν αδύνατη η ολοκλήρωση των σπουδών μου.

Ευχαριστώ ιδιαίτερα τους συμπαραστάτες καθηγητές μου κ. Χρήστο Σκουρλά και κ. Σταύρο Φατούρο για τη βοήθειά τους στα δύσκολα, το ειλικρινές ενδιαφέρον τους και το χρόνο τους.

Οπωσδήποτε ευχαριστώ βαθιά τους καλούς μου φίλους Παναγιώτη, Νίκο, Νικόλα, Γιάννη, Ορέστη, Ιωάννα και Κωσταντίνο για τον πολύτιμο χρόνο που διέθεσαν για να με βοηθήσουν καθ' όλη τη διάρκεια των σπουδών μου.

Ένα μεγάλο ευχαριστώ στους εκπαιδευτές μου που ακούραστα προσπάθησαν και με βοήθησαν να κατανοήσω τον κόσμο με έναν άλλο τρόπο διαφορετικό από τον δικό μου ιδιαίτερο τρόπο και να ανταποκριθώ στις απαιτήσεις του.

Δεν θα παραλείψω να θυμηθώ και να ευχαριστήσω με αγάπη τον δάσκαλό μου στο 14 ο Δημοτικό Σχολείο Αθηνών κ. Κώστα Πανούση και όλους τους καθηγητές μου στο Γυμνάσιο και στο Λύκειο.

Τέλος θα ήθελα να ευχαριστήσω την οικογένειά μου, την αδερφή μου Νεφέλη και τους γονείς μου, ανεκτίμητο στήριγμα στη διαδρομή των σπουδών μου και στη ζωή μου.

Περιεχόμενα

Ευχαριστίες.....	6
Κατάλογος Εικόνων	10
Περίληψη.....	11
Abstract	12
Κεφάλαιο 1. Εισαγωγή	13
Κεφάλαιο 2. Βιβλιογραφική Επισκόπηση.....	15
Κεφάλαιο 3. Νευρωνικά Δίκτυα	23
3.1 Εισαγωγή.....	23
3.2 Τεχνητοί Νευρώνες	24
3.3 Οργάνωση.....	24
3.4 Μάθηση	25
3.5 Στοχαστικά Νευρωνικά Δίκτυα	26
3.6 Εφαρμογές	26
Κεφάλαιο 4. Μηχανική Μάθηση (Machine Learning).....	27
4.1 Εισαγωγή.....	27
4.3 Κατηγοριοποίηση της Μηχανικής Μάθησης.....	27
4.3.1 Εποπτευόμενη Μηχανική Εκμάθηση.....	27
4.3.2 Μη εποπτευόμενη Μηχανική Μάθηση.....	28
4.3.3 Εκμάθηση Συμπερασμάτων.....	28
4.4 Υπάρχοντα Συστήματα Εκμάθησης	29
4.5 Εφαρμογές Μηχανικής Μάθησης	29
Κεφάλαιο 5. Βαθιά Μάθηση (Deep Learning)	31
5.1 Η βαθιά μάθηση ως συνέχιση των Νευρωνικών Δικτύων.....	31
5.2 Καινοτομίες στην αναγνώριση ομιλίας και αντικειμένων	31
5.3 Ήπια Προσοχή (Soft Attention) και η αρχιτεκτονική του μετασχηματιστή.	32
5.4 Χωρίς επίβλεψη και αυτοεπιβλέπουσα μάθηση.....	33
5.5 Αντιθετική μάθηση.....	34
5.6 Το μέλλον της βαθιάς μάθησης.....	35
Κεφάλαιο 6. Μεγάλα Μοντέλα Γλώσσας (Large Language Models - LLMs)	36
6.1 Εισαγωγή.....	36
6.2 Η εξάπλωση των Μεγάλων Μοντέλων Γλώσσας (LLM)	36
6.3 Κατηγορίες μεγάλων μοντέλων γλώσσας (LLM).....	37
6.4 Οι βελτιώσεις στο χώρο εξαιτίας των LLM.....	38
Κεφάλαιο 7. Τεχνητή Νοημοσύνη σε διάφορους τομείς	40

7.1 Αρχιτεκτονική	40
7.2 Τεχνητή Νοημοσύνη και Κυβερνοασφάλεια	41
7.2.1 Support Vector Machine	42
7.2.2 Δέντρο Απόφασης (Decision Tree)	43
7.2.3 Δίκτυο Βαθιάς Πεποίθησης - Deep Belief Network	43
7.2.4 Αναδρομικά Νευρωνικά Δίκτυα (Convolutional Neural Networks)	44
7.3 Κατηγοριοποίηση οδικών ατυχημάτων	46
Κεφάλαιο 8: Chat-GPT & DALL-E	47
8.1 Εισαγωγή	47
8.2 Στα ενδότερα του Chat-GPT	47
8.2.1 Εφαρμογές βασισμένες στο Chat-GPT	48
8.3 Στα ενδότερα του DALL-E	50
Κεφάλαιο 9. Συμπεράσματα	52
Βιβλιογραφία	54

Κατάλογος Εικόνων

Εικόνα 1. Η "ομπρέλα" της Τεχνητής Νοημοσύνης. (Πηγή: V7 labs)	13
Εικόνα 2. Απεικόνιση ενός τεχνητού νευρωνικού δικτύου δύο επιπέδων με επισημασμένους όρους βάρους και μεροληψίας. (Πηγή: Δωρεάν)	23
Εικόνα 3. Γραφική αναπαράσταση της μάθησης (Πηγή: Lorenzo Viscanti).....	25
Εικόνα 4. Ενδεικτική χρονοσειρά ανάπτυξης διάφορων MML. (Πηγή: Cascella et al., 2024)	39
Εικόνα 5. Παράδειγμα εξόδου εικόνων που παράχθηκαν από πλατφόρμες δημιουργίας εικόνων (Πηγή: Jo et al., 2024)	41
Εικόνα 6. Ένα παράδειγμα δέντρου αποφάσεων (Πηγή: Xin et al.).....	43
Εικόνα 7. Ένα παράδειγμα συνελκτικού νευρωνικού δικτύου (Πηγή: Xin et al.)	45

Περίληψη

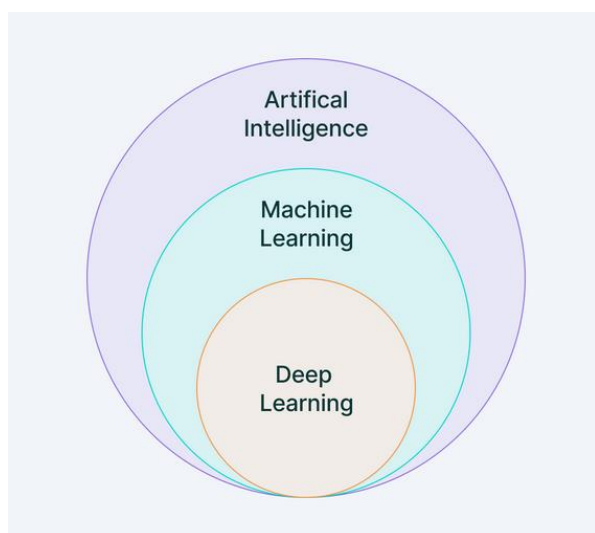
Η αύξηση της ταχύτητας της μετάδοσης της πληροφορίας, η δημιουργία συσκευών προσωπικής χρήσης προσβάσιμων σχεδόν σε οποιονδήποτε άνθρωπο στον κόσμο και κυρίως η ανάπτυξη της επεξεργαστικής ισχύος των συσκευών αυτών έχουν αποτελέσει κύριους παράγοντες για την ανάπτυξη εφαρμογών όπως το ChatGPT ή το DALL-E. Οι εφαρμογές αυτές είναι αποτελέσματα αυτής της ραγδαίας ανάπτυξης και αποτελούν σημεία καμπής στον τρόπο λειτουργίας των συστημάτων γενικότερα. Η εργασία αυτή έχει ως στόχο την επισκόπηση και την σύγκριση διάφορων τεχνικών που χρησιμοποιούνται για τη δημιουργία και χρήση της Τεχνητής Νοημοσύνης (AI). Αντλώντας πηγές μέσω του Scopus και του Google Scholar θα γίνει μία προσπάθεια επεξήγησης και κατανόησης δύσπεπτων και ιδιαίτερα περίπλοκων εννοιών όπως της μηχανικής εκμάθησης (Machine Learning) και βαθιάς μάθησης (Deep Learning). Επιπλέον θα δούμε πως αυτές διακλαδώνονται και αναπτύσσουν έννοιες όπως τα Μεγάλα Μοντέλα Γλώσσας (LLM - Large Language Models) και τέλος πως συγκλίνουν και υπάγονται στον ευρύτερο όρο Τεχνητή Νοημοσύνη. Επιπλέον θα δούμε παραδείγματα εφαρμογών πάνω σε τέτοια μοντέλα τα οποία χρησιμοποιούνται τόσο σε τομείς Ιατρικής, Κυβερνοασφάλειας αλλά και σε πιο δημιουργικούς τομείς..

Abstract

The increase in the speed of information transmission, the creation of devices for personal use accessible to almost any person in the world and above all the development of the processing power of these devices, have been the main factors for the development of applications such as ChatGPT or DALL-E. These applications are the result of this rapid development and are turning points in the way systems operate in general. This paper aims to review and compare various techniques used to create and use Artificial Intelligence (AI). Drawing sources through Scopus and Google Scholar, an attempt will be made to explain and understand difficult and particularly complex concepts such as machine learning (Machine Learning) and deep learning (Deep Learning). In addition, we will see how these branch out and develop concepts such as Large Language Models (LLM - Large Language Models) and finally how they converge and come under the broader term Artificial Intelligence. In addition, we will see examples of applications based on such models, which are used both in the fields of Medicine, Cybersecurity and also in more creative fields.

Κεφάλαιο 1. Εισαγωγή

Αρχικά επινοήθηκε τη δεκαετία του 1950 ο όρος «τεχνητή νοημοσύνη». Ξεκίνησε αρχικά ως η απλή θεωρία της ανθρώπινης νοημοσύνης που παρουσιάζεται από μηχανές. Το 1976, ο Jerrold S. Maxmen προέβλεψε ότι η τεχνητή νοημοσύνη (AI) θα επέφερε τη «μετα-ιατρική εποχή» στον εικοστό πρώτο αιώνα. Στη σημερινή εποχή της ραγδαίας τεχνολογικής προόδου και με εκθετικές αυξήσεις σε εξαιρετικά μεγάλα σύνολα δεδομένων ("μεγάλα δεδομένα") η τεχνητή νοημοσύνη έχει περάσει από απλή θεωρία σε απτή εφαρμογή σε άνευ προηγουμένου κλίμακα. Από την αξιολόγηση εξαιρετικά μεγάλων συνόλων δεδομένων σε σχεδόν πραγματικό χρόνο, αυτόνομης οδήγησης αυτοκινήτων και προτάσεων προβολής βίντεο που επηρεάζονται από το ιστορικό ροής (Netflix, Los Gatos, CA, ΗΠΑ) έως προτάσεις αγορών στο διαδίκτυο, διαφημίσεις και εντοπισμό απάτης (Amazon, Seattle, WA, ΗΠΑ) η τεχνητή νοημοσύνη έχει ριζώσει θεμελιωδώς σε πολλές πτυχές της κοινωνίας και συχνά λειτουργεί αόρατα στο φόντο των προσωπικών μας ηλεκτρονικών συσκευών. Τα εκθέματα μηχανικής μάθησης (ML – Machine Learning) θεωρούνται υποσύνολο της AI, τη βιωματική «μάθηση» που σχετίζεται με την ανθρώπινη νοημοσύνη ενώ έχει επίσης την ικανότητα να μαθαίνει και να βελτιώνει τις αναλύσεις της μέσω της χρήσης υπολογιστικών αλγορίθμων. Αυτοί οι αλγόριθμοι χρησιμοποιούν μεγάλα σύνολα εισόδων και εξόδων δεδομένων για να αναγνωρίσουν μοτίβα και να «μάθουν» αποτελεσματικά, προκειμένου να εκπαιδεύσουν το μηχάνημα να κάνει αυτόνομες συστάσεις ή αποφάσεις. Μετά από επαρκείς επαναλήψεις και τροποποίηση του αλγορίθμου το μηχάνημα μπορεί να λάβει μια είσοδο και να προβλέψει μια έξοδο. Στη συνέχεια οι έξοδοι συγκρίνονται με ένα σύνολο γνωστών αποτελεσμάτων προκειμένου να κριθεί η ακρίβεια του αλγορίθμου ο οποίος στη συνέχεια προσαρμόζεται επαναληπτικά για να τελειοποιήσει την ικανότητα πρόβλεψης περαιτέρω αποτελεσμάτων.



Εικόνα 1. Η "ομπρέλα" της Τεχνητής Νοημοσύνης. (Πηγή: V7 labs)

Τα "βαθιά νευρωνικά δίκτυα" (Neural Networks) είναι πιο σύνθετες εκδόσεις αυτών των μοντέλων που χρησιμοποιούν ιεραρχικά επίπεδα για τον διαχωρισμό και τη διαχείριση της τελικής παραγωγής. Το δίκτυο ξεκινά με ένα επίπεδο εισόδου που στη συνέχεια εξελίσσεται σε έναν αριθμό «κρυφών επιπέδων» που το καθένα ανταποκρίνεται σε διαφορετικά χαρακτηριστικά της εισόδου. Αυτά τα επίπεδα επιτρέπουν την αύξηση της κατανόησης καθώς η είσοδος ανεβαίνει «βαθύτερα», επιτρέποντας την ανάπτυξη μοντέλων χωρίς ρητά προγραμματισμένες κατευθύνσεις. Καθώς το μηχάνημα μελετά μια συγκεκριμένη ιδέα σε πολλαπλά επίπεδα, ο υπάρχων αλγόριθμος μπορεί στη συνέχεια να βελτιωθεί διαδοχικά καθώς είναι διαθέσιμα νέα

δεδομένα. Παρόμοια με τον τρόπο που λειτουργεί ο ανθρώπινος εγκέφαλος το μηχάνημα είναι σε θέση να κάνει «νευρωνικές» συνδέσεις από «δενδριτικές» συνδέσεις σε πολλαπλά ιεραρχικά επίπεδα δεδομένων. Αυτά τα δίκτυα οδήγησαν έτσι σε μια νέα μορφή τεχνητής νοημοσύνης γνωστή ως «βαθιά μάθηση» (Deep Learning). Είναι σημαντικό να κατανοηθούν κάποιες έννοιες όπως Τεχνητή Νοημοσύνη (Artificial Intelligence), Μηχανική Μάθηση (Machine Learning) και Βαθιά Μάθηση (Deep Learning), διότι χωρίς αυτό το υπόβαθρο είναι δύσκολο να υπάρξει μια ξεκάθαρη εικόνα πάνω σε αυτές τις τεχνολογίες.

Κεφάλαιο 2. Βιβλιογραφική Επισκόπηση

TITLE	AUTHORS	TYPE	JOURNAL	ABSTRACT
A DIST BIG DATA ANALYTICS MODEL FOR TRAFFIC ACCIDENTS CLASSIFICATION AND RECOGNITION BASED ON SPARKMLLIB CORES	Imad El Mallahi, Jamal Riffi, Hamid Tairi, Abderrahamane Ez-Zahout, Mohamed Adnane Mahraz	Journal Article	Journal of Automation, Mobile Robotics and Intelligent Systems VOLUME 16, N° 4 2022	Αυτή η εργασία επικεντρώνεται στο ζήτημα της ανάλυσης μεγάλων δεδομένων για την πρόβλεψη τροχαίων ατυχημάτων με βάση τους πυρήνες SparkMllib.
AI-GAS: AI-GENERATING ALGORITHMS, AN ALTERNATE PARADIGM FOR PRODUCING GENERAL ARTIFICIAL INTELLIGENCE	(Clune, 2020)	Preprint	Uber AI Labs, University of Wyoming	Αυτό το έγγραφο περιγράφει ένα άλλο συναρπαστικό μονοπάτι που τελικά μπορεί να είναι πιο επιτυχημένο στην παραγωγή γενικής τεχνητής νοημοσύνης.
NOW YOU SEE ME, NOW YOU DON'T: AN EXPLORATION OF RELIGIOUS EXNOMINATION IN DALL-E	(Alfano et al., 2024)	Journal Article	Ethics and Information Technology (2024) 26:27	Η μελέτη περίπτωσης του άρθρου είναι το DALL-E, ένα εργαλείο που δημιουργεί εικόνες από προτροπές φυσικής γλώσσας. Χρησιμοποιώντας το DALL-E mini, δημιουργούμε εικόνες από γενικές προτροπές όπως

				"θρησκευόμενο άτομο".
MACHINE LEARNING OVERVIEW	(Ayodele, n.d.)	Journal Article	New Advances in Machine Learning	Σκοπός του παρακάτω κεφαλαίου είναι η εξήγηση του τι είναι το Machine Learning.
PROGRAMMING IS HARD - OR AT LEAST IT USED TO BE: EDUCATIONAL OPPORTUNITIES AND CHALLENGES OF AI CODE GENERATION	(Becker et al., 2023)	Conference Paper	SIGCSE 2023: The 54th ACM Technical Symposium on Computer Science Education	Αυτό το Conference paper υποστηρίζει ότι η κοινότητα πρέπει να δράσει γρήγορα για να αποφασίσει ποιες πιθανές ευκαιρίες μπορούν και πρέπει να αξιοποιηθούν και πώς (σχετικά με τις τεχνικές AI), ενώ παράλληλα εργάζεται για να ξεπεράσει τις πιθανές προκλήσεις.
DEEP LEARNING FOR AI	(Bengio, 2009)	Journal Article	Communications of the ACM, Vol 64, Issue 7, pages 58-65	Πώς μπορούν τα νευρωνικά δίκτυα να μάθουν τις πλούσιες εσωτερικές αναπαραστάσεις που απαιτούνται για δύσκολες εργασίες όπως η αναγνώριση αντικειμένων ή η κατανόηση της γλώσσας;
THE DEFINITIVE	(Boonstra, 2021)	Book	ISBN: 978-1-4842-7013-4 978-	

<p>GUIDE TO CONVERSATIONAL AI WITH DIALOGFLOW AND GOOGLE CLOUD: BUILD ADVANCED ENTERPRISE CHATBOTS, VOICE, AND TELEPHONY AGENTS ON GOOGLE CLOUD</p>	<p>1-4842-7014-1</p>			
<p>IMPLEMENTING AI PRINCIPLES: FRAMEWORKS, PROCESSES, AND TOOLS</p>	<p>(Boza & Evgeniou, 2021)</p>	<p>Journal Article</p>	<p>SSRN Electronic Journal, 2021</p>	<p>Το άρθρο αυτό παρατηρεί και σχολιάζει την παρούσα κατάσταση στο πεδίο της χρήσης ΑΙ. Δίνεται έμφαση στον τρόπο εφαρμογής των προτεινόμενων αρχών τεχνητής νοημοσύνης και των επικείμενων κανονισμών που θα τα διέπουν. Επιπλέον προσπαθεί να απαντήσει στο πως θα αναπτυχθούν κατάλληλα εργαλεία για την επίτευξη αυτού του στόχου.</p>
<p>THE BREAKTHROUGH OF LARGE</p>	<p>(Casella et al., 2024)</p>	<p>Journal Article</p>	<p>Journal of Medical Systems, Vol. 48, Issue 1</p>	<p>Αυτό το άρθρο στοχεύει να παρέχει μια</p>

<p>LANGUAGE MODELS RELEASE FOR MEDICAL APPLICATIONS: 1-YEAR TIMELINE AND PERSPECTIVES</p>		<p>17/02/2024</p>	<p>συνοπτική επισκόπηση των πρόσφατα δημοσιευμένων LLMs, δίνοντας έμφαση στην πιθανή χρήση τους στον τομέα της ιατρικής.</p>
<p>KATTIS VS CHATGPT: ASSESSMENT AND EVALUATION OF PROGRAMMING TASKS IN THE AGE OF ARTIFICIAL INTELLIGENCE</p>	<p>(Dunder et al., 2024) Conference Paper</p>	<p>LAK '24: The 14th Learning Analytics and Knowledge Conference</p>	<p>Η παρούσα μελέτη εξετάζει την ικανότητα του ChatGPT να δημιουργεί λύσεις κώδικα σε διαφορετικά επίπεδα δυσκολίας για εισαγωγικά μαθήματα προγραμματισμού.</p>
<p>AUTOMATED MACHINE LEARNING: METHODS, SYSTEMS, CHALLENGES</p>	<p>(Hutter et al., 2019) Book</p>	<p>ISBN: 978-3-030-05317-8 978-3-030-05318-5</p>	
<p>GENERATIVE ARTIFICIAL INTELLIGENCE AND BUILDING DESIGN: EARLY PHOTOREALISTIC RENDER VISUALIZATION OF FAÇADES USING LOCAL IDENTITY-</p>	<p>(Jo et al., 2024) Journal Article</p>	<p>Journal of Computational Design and Engineering Vol. 11, Issue 2</p>	<p>Αυτή η εργασία αποσαφηνίζει μια προσέγγιση που χρησιμοποιεί τη γενετική τεχνητή νοημοσύνη (AI) για την ανάπτυξη εναλλακτικών επιλογών αρχιτεκτονικού σχεδιασμού με βάση την τοπική</p>

TRAINED MODELS				ταυτότητα.
<p>STRUCTURAL AND FUNCTIONAL PREDICTION, EVALUATION, AND VALIDATION IN THE POST-SEQUENCING ERA</p>	(Li et al., 2024)	Journal Article	<p>Computational and Structural Biotechnology Journal Vol. 23</p>	<p>Στο άρθρο αυτό υπάρχει μια επισκόπηση έρευνας σχετικά με την πρόβλεψη παραλλαγών που βασίζονται στην τεχνητή νοημοσύνη και αποσαφηνίζει τις προκλήσεις και τις ευκαιρίες για την πρόβλεψη παραλλαγών που βασίζονται στη δομή των πρωτεϊνών στην εποχή μετά την αλληλούχιση.</p>
<p>AN OVERVIEW OF THE SUPERVISED MACHINE LEARNING METHODS</p>	(Nasteski, 2017)	Journal Article	<p>HORIZONS.B Vol 4.</p>	<p>Αυτή η εργασία συνοψίζει τις θεμελιώδεις πτυχές δύο εποπτευόμενων μεθόδων. Ο κύριος στόχος και η συμβολή αυτού του εγγράφου ανασκόπησης είναι να παρουσιάσει την επισκόπηση της μηχανικής μάθησης και να παρέχει τεχνικές μηχανικής μάθησης.</p>
<p>A SELF-REGULATED LEARNING</p>	(Prasad & Sane, 2024)	Conference Paper	<p>Proceedings of the 55th ACM Technical</p>	<p>Σε αυτό το έγγραφο θέσης, συζητάμε τις</p>

<p>FRAMEWORK USING GENERATIVE AI AND ITS APPLICATION IN CS EDUCATIONAL INTERVENTION DESIGN</p>		<p>Symposium on Computer Science Education V. 1</p>	<p>προκλήσεις και τις ευκαιρίες που θέτουν οι γενετικές τεχνολογίες AI για τις στρατηγικές αυτορρύθμισης των αρχαρίων στο πλαίσιο του προγραμματισμού επίλυσης προβλημάτων.</p>
<p>AI FOR LOW-CODE FOR AI</p>	<p>(Rao et al., 2024) Conference Paper</p>	<p>Proceedings of the 29th International Conference on Intelligent User Interfaces</p>	<p>Το άρθρο αυτό υπογραμμίζει τα οφέλη του συνδυασμού της δύναμης της τεχνητής νοημοσύνης με τον προγραμματισμό χαμηλού κώδικα.</p>
<p>DEEP LEARNING WITH AZURE</p>	<p>(Salvaris et al., 2018)</p>	<p>Book ISBN: 978-1-4842-3678-9 978-1-4842-3679-6</p>	
<p>INSTRUCTOR PERCEPTIONS OF AI CODE GENERATION TOOLS – A MULTI-INSTITUTIONAL INTERVIEW STUDY</p>	<p>(Sheard et al., 2024) Conference Paper</p>	<p>SIGCSE 2024: The 55th ACM Technical Symposium on Computer Science Education</p>	<p>Σε αυτό το άρθρο, αναφέρονται τα αποτελέσματα μιας μελέτης συνέντευξης στην οποία συμμετείχαν 12 εκπαιδευτές από την Αυστραλία, τη Φινλανδία και τη Νέα Ζηλανδία, στην οποία διερευνούνται οι τρέχουσες πρακτικές, οι ανησυχίες και οι</p>

<p>APPLICATION OF BIGML IN THE CLASSIFICATION ON EVALUATION OF TOP COAL CAVING</p>	<p>(Wang et al., 2021)</p>	<p>Journal Article</p>	<p>Shock and Vibration Vol. 2021</p>	<p>προγραμματισμένες προσαρμογές των εκπαιδευτικών σχετικά με αυτά τα εργαλεία. Προκειμένου να πραγματοποιηθεί η αξιολόγηση ταξινόμησης της κορυφαίας σπηλαίωσης άνθρακα, αυτό το άρθρο εισάγει τη μέθοδο χρήσης του BigML για την καθιέρωση του μοντέλου αξιολόγησης ταξινόμησης της κορυφαίας σπηλαίωσης άνθρακα.</p>
<p>MACHINE LEARNING AND DEEP LEARNING METHODS FOR CYBERSECURITY</p>	<p>(Xin et al., 2018)</p>	<p>Journal Article</p>	<p>IEEE Access Vol. 6</p>	<p>Αυτή η αναφορά έρευνας περιγράφει βασικές βιβλιογραφικές έρευνες σχετικά με τις μεθόδους μηχανικής μάθησης (ML) και βαθιάς μάθησης (DL) για την ανάλυση δικτύου ανίχνευσης εισβολών και παρέχει μια σύντομη περιγραφή εκμάθησης κάθε</p>

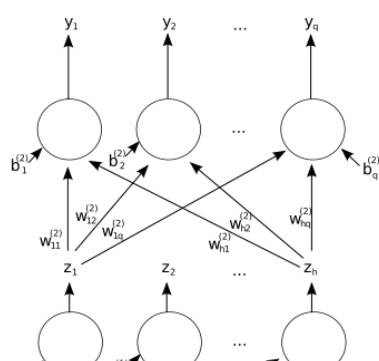
<p>MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE: DEFINITIONS, APPLICATIONS, AND FUTURE DIRECTIONS</p>	<p>(Helm et al., 2020)</p>	<p>Journal Article</p>	<p>Current Reviews in Musculoskeletal Medicine Vol 13, Issue 1</p>	<p>μεθόδου ML/DL. Ο σκοπός αυτής της ανασκόπησης είναι να αξιολογήσει κριτικά την πρόσφατη και νέα βιβλιογραφία σχετικά με το ML στον τομέα της ορθοπεδικής και να αντιμετωπίσει τις πιθανές επιπτώσεις της στο μέλλον της μυοσκελετικής φροντίδας.</p>
--	----------------------------	------------------------	--	---

Κεφάλαιο 3. Νευρωνικά Δίκτυα

3.1 Εισαγωγή

Η έρευνα στα τεχνητά νευρωνικά δίκτυα (Artificial Neural Network ή εν συντομία ANN) υποκινήθηκε από την παρατήρηση ότι η ανθρώπινη νοημοσύνη αναδύεται από πολύπλοκα παράλληλα δίκτυα σχετικά απλών, μη γραμμικών νευρώνων που μαθαίνουν προσαρμόζοντας τις δυνάμεις των συνδέσεών τους. Αυτή η παρατήρηση οδηγεί σε ένα κεντρικό υπολογιστικό ερώτημα: Πώς είναι δυνατόν τα δίκτυα αυτού του γενικού είδους να μάθουν τις περίπλοκες εσωτερικές αναπαραστάσεις που απαιτούνται για δύσκολες εργασίες όπως η αναγνώριση αντικειμένων ή η κατανόηση της γλώσσας;

Ένα τεχνητό νευρωνικό δίκτυο αποτελείται από συνδεδεμένες μονάδες ή κόμβους που ονομάζονται τεχνητοί νευρώνες οι οποίοι μοντελοποιούν κατά προσέγγιση τους



Εικόνα 2. Απεικόνιση ενός τεχνητού νευρωνικού δικτύου δύο επιπέδων με επισημασμένους όρους βάρους και μεροληψίας. (Πηγή: Δωρεάν)

νευρώνες στον εγκέφαλο. Αυτά συνδέονται με άκρες οι οποίες διαμορφώνουν τις συνάψεις στον εγκέφαλο. Κάθε τεχνητός νευρώνας λαμβάνει σήματα από συνδεδεμένους νευρώνες, στη συνέχεια τα επεξεργάζεται και στέλνει ένα σήμα σε άλλους συνδεδεμένους νευρώνες. Το «σήμα» είναι ένας πραγματικός αριθμός και η έξοδος κάθε νευρώνα υπολογίζεται από κάποια μη-γραμμική συνάρτηση του αθροίσματος των εισόδων του που ονομάζεται συνάρτηση ενεργοποίησης. Η ισχύς του σήματος σε κάθε σύνδεση καθορίζεται από ένα βάρος το οποίο προσαρμόζεται κατά τη διαδικασία εκμάθησης.

Τυπικά οι νευρώνες συγκεντρώνονται σε στρώματα (layers). Διαφορετικά επίπεδα μπορεί να εκτελούν διαφορετικούς μετασχηματισμούς στις εισόδους τους. Τα σήματα ταξιδεύουν από το πρώτο στρώμα (το στρώμα εισόδου ή ευρέως γνωστό ως Input Layer) στο τελευταίο στρώμα (το επίπεδο εξόδου – Output Layer) πιθανώς περνώντας από πολλαπλά ενδιάμεσα στρώματα (κρυμμένα στρώματα – Hidden Layers). Ένα δίκτυο συνήθως ονομάζεται βαθύ νευρωνικό δίκτυο εάν έχει τουλάχιστον δύο κρυφά επίπεδα.

Τα τεχνητά νευρωνικά δίκτυα χρησιμοποιούνται για διάφορες εργασίες συμπεριλαμβανομένης της προγνωστικής μοντελοποίησης, του προσαρμοστικού ελέγχου και της επίλυσης προβλημάτων στην τεχνητή νοημοσύνη. Μπορούν να μάθουν από την εμπειρία και μπορούν να βγάλουν συμπεράσματα από ένα περίπλοκο και φαινομενικά άσχετο σύνολο πληροφοριών.

Ιστορικά οι ψηφιακοί υπολογιστές εξελίχθηκαν από το μοντέλο von Neumann¹ και λειτουργούν μέσω της εκτέλεσης ρητών εντολών με πρόσβαση στη μνήμη από έναν αριθμό επεξεργαστών. Τα νευρωνικά δίκτυα από την άλλη πλευρά προέκυψαν από τις προσπάθειες για τη μοντελοποίηση της επεξεργασίας πληροφοριών σε βιολογικά συστήματα μέσω του πλαισίου της συνδετικότητας. Σε αντίθεση με το μοντέλο von Neumann ο συνδετικός υπολογισμός δεν διαχωρίζει τη μνήμη και την επεξεργασία.

3.2 Τεχνητοί Νευρώνες

Τα Τεχνητά Νευρωνικά Δίκτυα αποτελούνται από τεχνητούς νευρώνες οι οποίοι εννοιολογικά προέρχονται από βιολογικούς νευρώνες. Κάθε τεχνητός νευρώνας έχει εισόδους και παράγει μια ενιαία έξοδο που μπορεί να σταλεί σε πολλούς άλλους νευρώνες. Οι εισοδοί μπορεί να είναι οι τιμές χαρακτηριστικών ενός δείγματος εξωτερικών δεδομένων όπως εικόνες ή έγγραφα ή μπορεί να είναι οι εξοδοί άλλων νευρώνων. Οι εξοδοί των τελικών νευρώνων εξόδου του νευρικού δικτύου ολοκληρώνουν την εργασία όπως η αναγνώριση ενός αντικειμένου σε μια εικόνα.

Για να βρούμε την έξοδο του νευρώνα παίρνουμε το σταθμισμένο (ζυγισμένο) άθροισμα όλων των εισόδων, σταθμισμένο με τα βάρη των συνδέσεων από τις εισόδους στον νευρώνα. Προσθέτουμε έναν όρο μεροληψίας (bias term) σε αυτό το άθροισμα. Αυτό το σταθμισμένο άθροισμα μερικές φορές ονομάζεται ενεργοποίηση. Αυτό το σταθμισμένο άθροισμα στη συνέχεια διέρχεται μέσω μιας (συνήθως μη-γραμμικής) συνάρτησης ενεργοποίησης για να παραχθεί η έξοδος. Οι αρχικές εισοδοί είναι εξωτερικά δεδομένα όπως εικόνες και έγγραφα. Οι τελικές εξοδοί ολοκληρώνουν την εργασία όπως η αναγνώριση ενός αντικειμένου σε μια εικόνα.

3.3 Οργάνωση

Οι νευρώνες οργανώνονται συνήθως σε πολλαπλά στρώματα ειδικά στη βαθιά μάθηση. Οι νευρώνες ενός στρώματος συνδέονται μόνο με τους νευρώνες των αμέσως προηγούμενων και αμέσως επόμενων στιβάδων. Το επίπεδο που λαμβάνει εξωτερικά δεδομένα είναι το επίπεδο εισόδου. Το επίπεδο που παράγει το τελικό αποτέλεσμα είναι το επίπεδο εξόδου. Ανάμεσά τους υπάρχουν μηδέν ή περισσότερα κρυφά στρώματα. Χρησιμοποιούνται επίσης δίκτυα μεμονωμένων επιπέδων και μη-επιπέδων. Μεταξύ δύο στρωμάτων είναι δυνατά πολλαπλά μοτίβα σύνδεσης. Μπορούν να είναι «πλήρως συνδεδεμένοι», με κάθε νευρώνα σε ένα στρώμα να συνδέεται με κάθε νευρώνα στο επόμενο

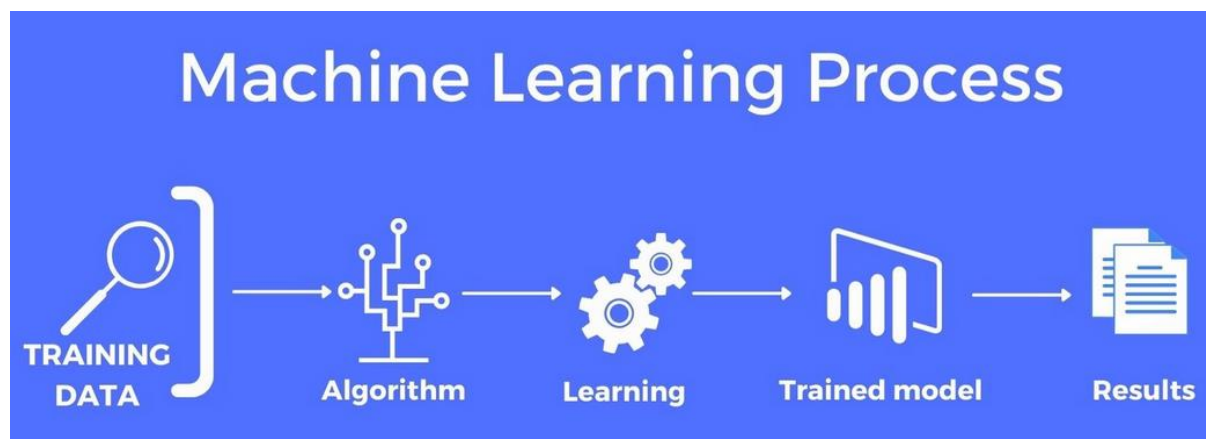
¹ Η αρχιτεκτονική von Neumann είναι μια αρχιτεκτονική υπολογιστή που βασίζεται σε μια περιγραφή του 1945 από τον John von Neumann και από άλλους, στο Πρώτο Προσχέδιο Έκθεσης για το EDVAC. Το έγγραφο περιγράφει μια αρχιτεκτονική σχεδίασης για έναν ηλεκτρονικό ψηφιακό υπολογιστή με μια μονάδα επεξεργασίας με αριθμητική λογική μονάδα και καταχωρητές επεξεργαστή, μια μονάδα ελέγχου που περιλαμβάνει έναν καταχωρητή εντολών και έναν μετρητή προγράμματος, μνήμη που αποθηκεύει δεδομένα και εντολές, εξωτερική μαζική αποθήκευση, και μηχανισμούς εισόδου και εξόδου.

στρώμα. Μπορούν να συνενώνονται, όπου μια ομάδα νευρώνων σε ένα στρώμα συνδέεται με έναν μόνο νευρώνα στο επόμενο στρώμα, μειώνοντας έτσι τον αριθμό των νευρώνων σε αυτό το στρώμα. Οι νευρώνες με μόνο τέτοιες συνδέσεις σχηματίζουν ένα κατευθυνόμενο μη κυκλικό γράφημα και είναι γνωστά ως δίκτυα τροφοδοσίας. Εναλλακτικά τα δίκτυα που επιτρέπουν συνδέσεις μεταξύ νευρώνων στο ίδιο ή προηγούμενα στρώματα είναι γνωστά ως επαναλαμβανόμενα δίκτυα.

Μια υπερπαράμετρος είναι μια σταθερή παράμετρος της οποίας η τιμή τίθεται πριν ξεκινήσει η διαδικασία εκμάθησης. Οι τιμές των παραμέτρων προέρχονται από τη μάθηση. Παραδείγματα υπερπαραμέτρων περιλαμβάνουν το ρυθμό εκμάθησης, τον αριθμό των κρυφών επιπέδων και το μέγεθος παρτίδας. Οι τιμές ορισμένων υπερπαραμέτρων μπορεί να εξαρτώνται από αυτές άλλων υπερπαραμέτρων. Για παράδειγμα το μέγεθος ορισμένων στρωμάτων μπορεί να εξαρτάται από τον συνολικό αριθμό των επιπέδων.

3.4 Μάθηση

Μάθηση είναι η προσαρμογή του δικτύου για να χειριστεί καλύτερα μια εργασία λαμβάνοντας υπόψη δειγματοληπτικές παρατηρήσεις. Η μάθηση περιλαμβάνει την προσαρμογή των βαρών (και των προαιρετικών ορίων) του δικτύου για τη βελτίωση της ακρίβειας του αποτελέσματος. Αυτό γίνεται ελαχιστοποιώντας τα παρατηρούμενα σφάλματα. Η μάθηση είναι πλήρης όταν η εξέταση πρόσθετων παρατηρήσεων δεν μειώνει επωφελώς το ποσοστό σφάλματος. Ακόμη και μετά την εκμάθηση το ποσοστό σφάλματος συνήθως δεν φτάνει το 0. Εάν μετά την εκμάθηση το ποσοστό σφάλματος είναι πολύ υψηλό, το δίκτυο συνήθως πρέπει να επανασχεδιαστεί.



Εικόνα 3. Γραφική αναπαράσταση της μάθησης (Πηγή: Lorenzo Viscanti)

Πρακτικά αυτό γίνεται με τον καθορισμό μιας συνάρτησης κόστους που αξιολογείται περιοδικά κατά τη διάρκεια της εκμάθησης. Όσο η παραγωγή του συνεχίζει να μειώνεται η μάθηση συνεχίζεται. Το κόστος ορίζεται συχνά ως ένα στατιστικό στοιχείο του οποίου η αξία μπορεί να υπολογιστεί μόνο κατά προσέγγιση. Οι έξοδοι είναι στην πραγματικότητα αριθμοί, οπότε όταν το σφάλμα είναι χαμηλό, η διαφορά μεταξύ της εξόδου (για παράδειγμα σχεδόν σίγουρα μια γάτα) και της σωστής απάντησης (γάτα) είναι μικρή. Η μάθηση προσπαθεί να

μειώσει το σύνολο των διαφορών μεταξύ των παρατηρήσεων. Τα περισσότερα μοντέλα μάθησης μπορούν να θεωρηθούν ως μια απλή εφαρμογή της θεωρίας βελτιστοποίησης και της στατιστικής εκτίμησης.

3.5 Στοχαστικά Νευρωνικά Δίκτυα

Τα στοχαστικά νευρωνικά δίκτυα είναι ένας τύπος τεχνητού νευρωνικού δικτύου που δημιουργείται με την εισαγωγή τυχαίων παραλλαγών στο δίκτυο, είτε δίνοντας στους τεχνητούς νευρώνες του δικτύου συναρτήσεις στοχαστικής μεταφοράς είτε δίνοντάς τους στοχαστικά βάρη. Αυτό τα καθιστά χρήσιμα εργαλεία για προβλήματα βελτιστοποίησης, καθώς οι τυχαίες διακυμάνσεις βοηθούν το δίκτυο να ξεφύγει από τα τοπικά ελάχιστα. Τα στοχαστικά νευρωνικά δίκτυα που εκπαιδεύονται χρησιμοποιώντας μια Μπεϋζιανή προσέγγιση είναι γνωστά ως Μπεϋζιανά νευρωνικά δίκτυα.

3.6 Εφαρμογές

Λόγω της ικανότητάς τους να αναπαράγουν και να μοντελοποιούν μη γραμμικές διεργασίες, τα τεχνητά νευρωνικά δίκτυα έχουν βρει εφαρμογές σε πολλούς κλάδους. Αυτά περιλαμβάνουν:

- Ανάλυση προσέγγισης συναρτήσεων ή παλινδρόμησης
- Επεξεργασία δεδομένων
- Μη γραμμική αναγνώριση και έλεγχος συστήματος (συμπεριλαμβανομένου του ελέγχου οχήματος, της πρόβλεψης τροχιάς, προσαρμοστικού ελέγχου, ελέγχου διεργασιών και διαχείρισης φυσικών πόρων)
- Αναγνώριση προτύπων (συμπεριλαμβανομένων συστημάτων ραντάρ, αναγνώρισης προσώπου, ταξινόμησης σημάτων, ανίχνευσης νεωτερισμού, τρισδιάστατης ανακατασκευής, αναγνώρισης αντικειμένων και διαδοχικής λήψης αποφάσεων)
- Αναγνώριση ακολουθίας (συμπεριλαμβανομένης της χειρονομίας, της ομιλίας και της αναγνώρισης χειρόγραφου και έντυπου κειμένου)
- Ανάλυση δεδομένων αισθητήρα (συμπεριλαμβανομένης της ανάλυσης εικόνας)
- Ρομποτική (συμπεριλαμβανομένης της διεύθυνσης χειριστών και προσθετικών)
- Εξόρυξη δεδομένων (συμπεριλαμβανομένης της ανακάλυψης γνώσης σε βάσεις δεδομένων)
- Χρηματοοικονομικά (όπως εκ των προτέρων μοντέλα για συγκεκριμένες χρηματοοικονομικές μακροπρόθεσμες προβλέψεις και τεχνητές χρηματοπιστωτικές αγορές)
- Γενετική Τεχνητή Νοημοσύνη (Generative AI)
- Οπτικοποίηση δεδομένων
- Μηχανική μετάφραση
- Φιλτράρισμα κοινωνικών δικτύων
- Φιλτράρισμα ανεπιθύμητων μηνυμάτων ηλεκτρονικού ταχυδρομείου

- Ιατρική διάγνωση

Κεφάλαιο 4. Μηχανική Μάθηση (Machine Learning)

4.1 Εισαγωγή

Η κύρια εστίαση της μηχανικής μάθησης είναι η αυτόματη παραγωγή μοντέλων και ένα μοντέλο είναι ένα μοτίβο, σχέδιο, αναπαράσταση ή περιγραφή που έχει σχεδιαστεί για να δείξει την κύρια λειτουργία ενός συστήματος ή έννοιας. Για παράδειγμα κανόνες που καθορίζουν τον κανόνα για την εκτέλεση μιας μαθηματικής πράξης και την απόκτηση ενός συγκεκριμένου αποτελέσματος, μια συνάρτηση από σύνολα τύπων σε τύπους και μοτίβα (μοντέλο που μπορεί να χρησιμοποιηθεί για τη δημιουργία πραγμάτων ή τμημάτων ενός πράγματος από δεδομένα).

Οι μαθησιακές διαδικασίες περιλαμβάνουν την απόκτηση νέας γνώσης, την ανάπτυξη κινητικών και γνωστικών δεξιοτήτων μέσω διδασκαλίας ή πρακτικής, την οργάνωση της νέας γνώσης σε γενικές, αποτελεσματικές αναπαραστάσεις και την ανακάλυψη νέων γεγονότων και θεωριών μέσω παρατήρησης και πειραματισμού. Η μελέτη και η υπολογιστική μοντελοποίηση των μαθησιακών διαδικασιών, στις πολλαπλές εκφάνσεις τους, αποτελεί το αντικείμενο της μηχανικής μάθησης. Αν και η μηχανική μάθηση ήταν κεντρικό μέλημα της τεχνητής νοημοσύνης από τις πρώτες μέρες που ήταν δημοφιλής, η ιδέα των «αυτο-οργανωμένων συστημάτων», οι περιορισμοί που υπήρξαν στις πρώτες προσεγγίσεις νευρωνικών δικτύων οδήγησαν σε προσωρινή μείωση του όγκου της έρευνας. Πιο πρόσφατα νέες συμβολικές μέθοδοι και τεχνικές έντασης γνώσης έχουν δώσει πολλά υποσχόμενα αποτελέσματα και αυτά με τη σειρά τους έχουν οδηγήσει στην τρέχουσα αναβίωση της έρευνας μηχανικής μάθησης.

4.3 Κατηγοριοποίηση της Μηχανικής Μάθησης

4.3.1 Εποπτευόμενη Μηχανική Εκμάθηση

Η εποπτευόμενη μάθηση είναι μια τεχνική μηχανικής εκμάθησης για την εκμάθηση μιας συνάρτησης από δεδομένα εκπαίδευσης. Τα δεδομένα εκπαίδευσης αποτελούνται από ζεύγη αντικειμένων εισόδου (συνήθως διανύσματα) και επιθυμητές εξόδους. Η έξοδος της συνάρτησης μπορεί να είναι μια συνεχής τιμή (που ονομάζεται παλινδρόμηση) ή μπορεί να προβλέψει μια ετικέτα κλάσης του αντικειμένου εισόδου (ονομάζεται ταξινόμηση). Το καθήκον του εποπτευόμενου εκπαιδευόμενου είναι να προβλέψει την τιμή της συνάρτησης για οποιοδήποτε έγκυρο αντικείμενο εισόδου αφού έχει δει έναν αριθμό παραδειγμάτων εκπαίδευσης (δηλαδή ζεύγη εισόδου και εξόδου στόχου). Για να επιτευχθεί αυτό, ο εκπαιδευόμενος πρέπει να γενικεύσει από τα παρουσιαζόμενα δεδομένα σε αόρατες καταστάσεις με «λογικό» τρόπο. Η εποπτευόμενη μάθηση λοιπόν είναι μια τεχνική μηχανικής μάθησης κατά την οποία ο αλγόριθμος παρουσιάζεται αρχικά με δεδομένα

εκπαίδευσης που αποτελούνται από παραδείγματα που περιλαμβάνουν τόσο τις εισόδους όσο και τις επιθυμητές εξόδους, επιτρέποντάς του έτσι να μάθει μια λειτουργία. Ο μαθητής θα πρέπει στη συνέχεια να είναι σε θέση να γενικεύσει από τα παρουσιαζόμενα δεδομένα σε μη ορατά παραδείγματα.

4.3.2 Μη εποπτευόμενη Μηχανική Μάθηση

Η μάθηση χωρίς επίβλεψη είναι ένας τύπος μηχανικής εκμάθησης όπου δεν χρησιμοποιούνται χειροκίνητες ετικέτες εισόδων. Διακρίνεται από τις εποπτευόμενες προσεγγίσεις μάθησης που μαθαίνουν πώς να εκτελούν μια εργασία, όπως η ταξινόμηση ή η παλινδρόμηση, χρησιμοποιώντας ένα σύνολο παραδειγμάτων προετοιμασμένων από τον άνθρωπο. Η μάθηση χωρίς επίβλεψη σημαίνει ότι μας δίνονται μόνο τα X και κάποια (τελική) λειτουργία ανατροφοδότησης για την απόδοσή μας. Έχουμε απλώς ένα εκπαιδευτικό σύνολο διανυσμάτων χωρίς τιμές συνάρτησης αυτών.

Με απλούστερα λόγια, η μάθηση χωρίς επίβλεψη είναι μια μέθοδος μηχανικής μάθησης όπου ο αλγόριθμος τροφοδοτείται με παραδείγματα μόνο από την είσοδο και ένα μοντέλο ταιριάζει αυτές τις παρατηρήσεις. Για παράδειγμα, ένας αλγόριθμος ομαδοποίησης θα ήταν μια μορφή μάθησης χωρίς επίβλεψη. Η μάθηση χωρίς επίβλεψη είναι μια μέθοδος μηχανικής μάθησης όπου ένα μοντέλο είναι κατάλληλο για παρατηρήσεις. Διακρίνεται από την εποπτευόμενη μάθηση από το γεγονός ότι δεν υπάρχει από πριν κάποια έξοδος.

4.3.3 Εκμάθηση Συμπερασμάτων

Αυτή είναι μια μορφή ταξινόμησης με βάση την υποκείμενη στρατηγική που εμπλέκεται. Αυτές οι στρατηγικές εξαρτώνται από το μέγεθος των συμπερασμάτων που εκτελεί το σύστημα εκμάθησης σχετικά με τις πληροφορίες που παρέχονται στο σύστημα. Οι στρατηγικές μάθησης διακρίνονται από το μέγεθος των συμπερασμάτων που εκτελεί ο εκπαιδευόμενος σχετικά με τις παρεχόμενες πληροφορίες. Έτσι εάν ένα σύστημα υπολογιστή πραγματοποιεί ταξινόμηση email για παράδειγμα, οι γνώσεις του αυξάνονται αλλά αυτό μπορεί να μην παράγει κανένα συμπέρασμα για τις νέες πληροφορίες. Αυτό σημαίνει ότι όλες οι γνωστικές προσπάθειες είναι από την πλευρά του αναλυτή ή του προγραμματιστή. Αλλά, εάν ο ταξινομητής μηχανικής μάθησης ανακαλύψει ανεξάρτητα νέες θεωρίες ή υιοθετήσει νέες έννοιες, αυτό θα δώσει ένα πολύ σημαντικό συμπέρασμα. Αυτό ονομάζεται απόκτηση γνώσης από παράδειγμα ή πειράματα ή με παρατήρηση. Ένα ακόμα πιο απλό παράδειγμα είναι: Όταν ένας μαθητής θέλει να λύσει στατιστικά προβλήματα σε ένα εγχειρίδιο αυτή είναι μια διαδικασία που περιλαμβάνει συμπέρασμα αλλά η λύση δεν είναι να ανακαλύψει μια ολοκαίνουργια εξίσωση χωρίς καθοδήγηση από έναν δάσκαλο ή ένα βιβλίο. Έτσι, καθώς αυξάνεται ο όγκος των συμπερασμάτων που μπορεί να κάνει ο εκπαιδευόμενος, μειώνονται τα βάρη που επιβαρύνουν τον δάσκαλο ή το εξωτερικό περιβάλλον.

4.4 Υπάρχοντα Συστήματα Εκμάθησης

Παρακάτω αναφέρονται διάφορα συστήματα εκμάθησης.

- Εκμάθηση μέσω προγραμματισμού

Όταν ένας αλγόριθμος ή ένας κώδικας είναι γραμμένος για την εκτέλεση συγκεκριμένης εργασίας. Για παράδειγμα ένας κωδικός γράφεται ως παιχνίδι εικασίας για τον τύπο του ζώου. Ένα τέτοιο πρόγραμμα θα μπορούσε να τροποποιηθεί από εξωτερική οντότητα.

- Μάθηση μέσω απομνημόνευσης

Αυτό γίνεται με την απομνημόνευση δεδομένων γεγονότων ή δεδομένων χωρίς να προκύψουν συμπεράσματα από τις εισερχόμενες πληροφορίες ή δεδομένα.

- Μάθηση από παραδείγματα

Πρόκειται για μια ειδική περίπτωση επαγωγικής μάθησης. Δεδομένου ενός συνόλου παραδειγμάτων και αντιπαραδειγμάτων μιας έννοιας, ο εκπαιδευόμενος προκαλεί μια γενική περιγραφή έννοιας που περιγράφει όλα τα θετικά παραδείγματα και κανένα από τα αντιπαραδείγματα. Η μάθηση από παραδείγματα είναι μια μέθοδος που έχει διερευνηθεί σε μεγάλο βαθμό στον τομέα της τεχνητής νοημοσύνης. Ο όγκος των συμπερασμάτων που εκτελεί ο μαθητής είναι πολύ μεγαλύτερος από ό,τι όταν μαθαίνει από οδηγίες.

- Μάθηση από την παρατήρηση

Αυτή είναι μια προσέγγιση μάθησης χωρίς επίβλεψη και είναι μια πολύ γενική μορφή επαγωγικής μάθησης που περιλαμβάνει συστήματα ανακάλυψης, εργασίες σχηματισμού θεωρίας, δημιουργία κριτηρίων ταξινόμησης για τη διαμόρφωση ταξινομικών ιεραρχιών και παρόμοια εργασία που πρέπει να εκτελείται χωρίς το όφελος ενός εξωτερικού δασκάλου. Η μάθηση χωρίς επίβλεψη απαιτεί από τον μαθητή να κάνει περισσότερα συμπεράσματα από οποιαδήποτε προσέγγιση, όπως εξηγήθηκε προηγουμένως. Δεν παρέχεται στον εκπαιδευόμενο ένα σύνολο δεδομένων ή παράδειγμα μιας συγκεκριμένης έννοιας.

4.5 Εφαρμογές Μηχανικής Μάθησης

Παρακάτω αναφέρονται τομείς στους οποίους έχουν εφαρμοστεί διάφορα υπάρχοντα συστήματα μάθησης. Αυτοί είναι:

- Προγραμματισμός Η/Υ
- Παιχνίδι (σκάκι, πόκερ και ούτω καθεξής)
- Αναγνώριση εικόνας, Αναγνώριση ομιλίας
- Ιατρική διάγνωση
- Διαχείριση email, Ρομποτική

Συγκριτική Μελέτη εργαλείων Γενετικής Τεχνητής Νοημοσύνης

- Μαθηματικά
- Επεξεργασία Φυσικής Γλώσσας και πολλά άλλα.

Κεφάλαιο 5. Βαθιά Μάθηση (Deep Learning)

Οι μέθοδοι βαθιάς μάθησης (Βαθιά Μάθηση – Deep Learning) στοχεύουν στη μάθηση ιεραρχιών χαρακτηριστικών με χαρακτηριστικά από υψηλότερα επίπεδα της ιεραρχίας που σχηματίζονται από τη σύνθεση χαρακτηριστικών χαμηλότερου επιπέδου. Με πιο απλά λόγια, διάφορα χαρακτηριστικά (features) χαμηλότερου επιπέδου συνθέτουν χαρακτηριστικά ανωτέρων επιπέδων. Οι λειτουργίες αυτόματης εκμάθησης σε πολλαπλά επίπεδα αφαίρεσης (abstraction) επιτρέπουν σε ένα σύστημα να μαθαίνει πολύπλοκες λειτουργίες αντιστοιχίζοντας την είσοδο στην έξοδο απευθείας από δεδομένα, χωρίς να εξαρτάται πλήρως από χαρακτηριστικά που έχουν δημιουργηθεί από τον άνθρωπο. Αυτό είναι ιδιαίτερα σημαντικό για αφαιρέσεις υψηλότερου επιπέδου τις οποίες οι άνθρωποι συχνά δεν ξέρουν πώς να προσδιορίσουν ρητά από την άποψη της ακατέργαστης αισθητηριακής εισροής. Η δυνατότητα αυτόματης εκμάθησης ισχυρών λειτουργιών θα γίνεται όλο και πιο σημαντική καθώς ο όγκος των δεδομένων και το εύρος των εφαρμογών στις μεθόδους μηχανικής εκμάθησης συνεχίζει να αυξάνεται. (Bengio, 2009)

5.1 Η βαθιά μάθηση ως συνέχιση των Νευρωνικών Δικτύων

Η βαθιά μάθηση αναζωογόνησε την έρευνα νευρωνικών δικτύων στις αρχές της δεκαετίας του 2000, εισάγοντας μερικά στοιχεία που διευκόλυναν την εκπαίδευση βαθύτερων δικτύων. Η εμφάνιση των γραφικών επεξεργαστικών μονάδων (GPU – Graphics Processor Unit) και η διαθεσιμότητα μεγάλων συνόλων δεδομένων ήταν βασικοί παράγοντες για τη βαθιά μάθηση και ενισχύθηκαν σημαντικά από την ανάπτυξη ανοιχτού κώδικα ευέλικτων πλατφόρμων λογισμικού με αυτόματη διαφοροποίηση όπως Theano, Torch, Caffe, TensorFlow και PyTorch. Αυτό διευκόλυνε την εκπαίδευση πολύπλοκων βαθιών δικτύων και την επαναχρησιμοποίηση των πιο πρόσφατων μοντέλων και των δομικών στοιχείων τους. Αλλά η σύνθεση περισσότερων επιπέδων είναι αυτό που επέτρεψε πιο σύνθετες μηχανισμικότητες και πέτυχε εκπληκτικά καλά αποτελέσματα στις εργασίες αντίληψης.

5.2 Καινοτομίες στην αναγνώριση ομιλίας και αντικειμένων

Ένα ακουστικό μοντέλο μετατρέπει μια αναπαράσταση του ηχητικού κύματος σε κατανομή πιθανοτήτων σε θραύσματα φωνημάτων. Το 2009, δύο μεταπτυχιακοί φοιτητές² που χρησιμοποιούσαν GPU της Nvidia έδειξαν ότι τα προ-εκπαιδευμένα βαθιά νευρωνικά δίκτυα θα μπορούσαν να ξεπεράσουν ελαφρώς το SOTA (State Of The Art) στο σύνολο δεδομένων TIMIT³. Αυτό το αποτέλεσμα αναζωογόνησε το ενδιαφέρον πολλών κορυφαίων

² Mohamed, A., Dahl, G., and Hinton, G. Deep belief networks for phone recognition. In Proceedings of NIPS Workshop on Deep Learning for Speech Recognition and Related Applications. (Vancouver, Canada, 2009).

³ Το σύνολο δεδομένων TIMIT Acoustic-Phonetic Continuous Speech Corpus είναι ένα τυπικό σύνολο δεδομένων που χρησιμοποιείται για την αξιολόγηση συστημάτων αυτόματης αναγνώρισης ομιλίας. Περιέχει ηχογραφήσεις 630 ομιλούντων. Επίσης, οι ηχογραφήσεις περιλαμβάνουν οκτώ διαλέκτους της αμερικανικής

ομάδων ομιλίας στα νευρωνικά δίκτυα. Το 2010, ουσιαστικά το ίδιο δίκτυο σε βάθος αποδείχθηκε ότι ξεπερνά το SOTA για την αναγνώριση ομιλίας μεγάλου λεξιλογίου χωρίς να απαιτείται εκπαίδευση εξαρτώμενη από τον ομιλητή και μέχρι το 2012 η Google είχε σχεδιάσει μια έκδοση παραγωγής που βελτίωσε σημαντικά τη φωνητική αναζήτηση στο Android. Αυτή ήταν μια πρώιμη επίδειξη της ανατρεπτικής δύναμης της βαθιάς μάθησης.

Την ίδια περίπου εποχή, η βαθιά εκμάθηση σημείωσε μια δραματική νίκη στον διαγωνισμό ImageNet του 2012, μειώνοντας σχεδόν στο μισό το ποσοστό σφάλματος για την αναγνώριση χιλίων διαφορετικών κατηγοριών αντικειμένων σε φυσικές εικόνες. Τα κλειδιά αυτής της νίκης ήταν η μεγάλη προσπάθεια της Fei-Fei Li και των συνεργατών της να συλλέξουν περισσότερες από ένα εκατομμύριο ετικέτες εικόνων⁴ για το σετ εκπαίδευσης και η πολύ αποτελεσματική χρήση πολλαπλών GPU από τον Alex Krizhevsky. Το τρέχον υλικό, συμπεριλαμβανομένων των GPU, ενθαρρύνει τη χρήση μεγάλων μινι-συνόλων δεδομένων (mini-batches), προκειμένου να αποσβεστεί το κόστος ανάκτησης ενός βάρους από τη μνήμη σε πολλές χρήσεις αυτού του βάρους. Καθαρή διαδικτυακή στοχαστική κλίση που χρησιμοποιεί κάθε βάρους, μόλις συγκλίνει γρηγορότερα και το μελλοντικό υλικό μπορεί απλώς να χρησιμοποιεί βάρη στη θέση του αντί να τα παίρνει από τη μνήμη. Το βαθύ συνελκτικό νευρωνικό δίκτυο περιείχε μερικές καινοτομίες όπως τη χρήση των διορθωμένων συναρτήσεων γραμμικής ενεργοποίησης (ReLU⁵) για ταχύτερη εκμάθηση και τη χρήση της εγκατάλειψης για την αποφυγή υπερβολικής προσαρμογής, αλλά βασικά ήταν απλώς ένα συνελκτικό νευρωνικό δίκτυο με ανατροφοδότηση του είδους που ανέπτυξαν ο Yann LeCun και οι συνεργάτες του για πολλά χρόνια. Η ανταπόκριση της κοινότητας της υπολογιστικής όρασης σε αυτή την ανακάλυψη ήταν αξιοθαύμαστη. Δεδομένης αυτής της αδιαμφισβήτητης απόδειξης της ανωτερότητας των συνελκτικών νευρωνικών δικτύων, η κοινότητα εγκατέλειψε γρήγορα προηγούμενες χειροκίνητες προσεγγίσεις και στράφηκε στη βαθιά μάθηση.

5.3 Ήπια Προσοχή (Soft Attention) και η αρχιτεκτονική του μετασχηματιστή.

Μια σημαντική εξέλιξη στη βαθιά μάθηση, ειδικά όταν πρόκειται για διαδοχική επεξεργασία, είναι η χρήση πολλαπλασιαστικών αλληλεπιδράσεων, ιδιαίτερα με τη μορφή της ήπιας προσοχής. αλλάζει τα νευρωνικά δίκτυα από μηχανές μετασχηματισμού καθαρών διανυσμάτων σε αρχιτεκτονικές που μπορούν να επιλέξουν δυναμικά ποιες εισόδους θα λειτουργήσουν και μπορούν να αποθηκεύσουν πληροφορίες σε διαφοροποιήσιμες συσχετιστικές μνήμες. Μια βασική ιδιότητα τέτοιων αρχιτεκτονικών είναι ότι μπορούν να

αγγλικής γλώσσας. Κάθε ομιλητής στο σύνολο δεδομένων διαβάζει 10 φωνητικά πλούσιες προτάσεις. Το σώμα TIMIT περιλαμβάνει χρονικά ευθυγραμμισμένες ορθογραφικές, φωνητικές και μεταγραφές λέξεων. Περιλαμβάνει επίσης ένα αρχείο κυματομορφής ομιλίας 16 bit, 16 kHz για κάθε φράση που λέγεται.

⁴ . Deng, J., Dong, W., Socher, R., Li, L., Li, K., and Fei-Fei, L. ImageNet: A large-scale hierarchical image database. In Proceedings of 2009 IEEE Conf. Computer Vision and Pattern Recognition, 248–255

⁵ Η διορθωμένη συνάρτηση γραμμικής ενεργοποίησης ή ReLU για συντομία είναι μια τμηματικά γραμμική συνάρτηση που θα εξάγει απευθείας την είσοδο εάν είναι θετική, διαφορετικά, θα βγάξει μηδέν. Έχει γίνει η προεπιλεγμένη λειτουργία ενεργοποίησης για πολλούς τύπους νευρωνικών δικτύων, επειδή ένα μοντέλο που το χρησιμοποιεί είναι πιο εύκολο να εκπαιδευτεί και συχνά επιτυγχάνει καλύτερη απόδοση.

λειτουργήσουν αποτελεσματικά σε διαφορετικά είδη δομών δεδομένων, συμπεριλαμβανομένων συνόλων και γραφημάτων.

Η ήπια προσοχή (soft attention) μπορεί να χρησιμοποιηθεί από δομοστοιχεία (modules) σε ένα επίπεδο για να επιλέξουν δυναμικά ποια διανύσματα από το προηγούμενο επίπεδο θα συνδυάσουν για να υπολογίσουν τις εξόδους τους. Αυτό μπορεί να χρησιμεύσει για να γίνει η έξοδος ανεξάρτητη από τη σειρά με την οποία παρουσιάζονται οι είσοδοι (μεταχείρισή τους ως σύνολο) ή για τη χρήση σχέσεων μεταξύ διαφορετικών εισόδων (μεταχείρισή τους ως γράφημα).

Η αρχιτεκτονική του μετασχηματιστή, η οποία έχει γίνει η κυρίαρχη αρχιτεκτονική σε πολλές εφαρμογές, στοιβάξει πολλά στρώματα μονάδων «αυτοπροσοχής». Κάθε ενότητα σε ένα επίπεδο χρησιμοποιεί ένα βαθμωτό γινόμενο για να υπολογίσει την αντιστοίχιση μεταξύ του διανύσματος ερωτήματός του και των διανυσμάτων κλειδιών άλλων λειτουργικών μονάδων σε αυτό το επίπεδο. Οι αντιστοιχίσεις κανονικοποιούνται στο άθροισμα στο 1 και οι προκύπτοντες βαθμωτοί συντελεστές χρησιμοποιούνται στη συνέχεια για να σχηματίσουν έναν κυρτό συνδυασμό των διανυσματικών τιμών που παράγονται από τις άλλες μονάδες στο προηγούμενο επίπεδο. Το διάνυσμα που προκύπτει σχηματίζει μια είσοδο για μια ενότητα του επόμενου σταδίου υπολογισμού. Οι ενότητες μπορούν να γίνουν πολλαπλές, έτσι ώστε κάθε ενότητα να υπολογίζει πολλά διαφορετικά διανύσματα ερωτημάτων, κλειδιών και τιμών, καθιστώντας έτσι δυνατό για κάθε ενότητα να έχει πολλές διαφορετικές θέσεις, καθεμία επιλεγμένη από τις ενότητες του προηγούμενου σταδίου με διαφορετικό τρόπο. Η σειρά και ο αριθμός των μονάδων δεν έχει σημασία σε αυτή τη λειτουργία, καθιστώντας δυνατή τη λειτουργία σε σύνολα διανυσμάτων και όχι μεμονωμένα διανύσματα, όπως στα παραδοσιακά νευρωνικά δίκτυα. Για παράδειγμα, ένα σύστημα μετάφρασης γλώσσας, όταν παράγει μια λέξη στην πρόταση εξόδου, μπορεί να επιλέξει να δώσει προσοχή στην αντίστοιχη ομάδα λέξεων στην πρόταση εισαγωγής, ανεξάρτητα από τη θέση τους στο κείμενο. Ενώ η πολλαπλασιαστική πύλη είναι μια παλιά ιδέα για πράγματα όπως οι μετασχηματισμοί συντεταγμένων και οι ισχυρές μορφές επαναλαμβανόμενων δικτύων, οι πρόσφατες μορφές της το έχουν κάνει κυρίαρχο ρεύμα (mainstream). Ένας άλλος τρόπος να σκεφτούμε τους μηχανισμούς προσοχής είναι ότι καθιστούν δυνατή τη δυναμική δρομολόγηση πληροφοριών μέσω κατάλληλα επιλεγμένων ενοτήτων και συνδυάζουν αυτές τις ενότητες με δυναμικά νέους τρόπους για βελτιωμένη γενίκευση εκτός διανομής.

Οι μετασχηματιστές έχουν δημιουργήσει δραματικές βελτιώσεις στην απόδοση που έχουν φέρει επανάσταση στην επεξεργασία της φυσικής γλώσσας και πλέον χρησιμοποιούνται συστηματικά στη βιομηχανία. Αυτά τα συστήματα είναι όλα προεκπαιδευμένα με τρόπο αυτοεποπτευόμενο για την πρόβλεψη λέξεων που λείπουν σε ένα τμήμα του κειμένου.

5.4 Χωρίς επίβλεψη και αυτοεπιβλέπουσα μάθηση

Η εποπτευόμενη μάθηση, ενώ είναι επιτυχής σε μια ευρεία ποικιλία εργασιών, απαιτεί συνήθως μεγάλο αριθμό δεδομένων με επισήμανση ανθρώπου. Ομοίως, όταν η

ενισχυτική μάθηση βασίζεται μόνο σε ανταμοιβές, απαιτεί πολύ μεγάλο αριθμό αλληλεπιδράσεων. Αυτές οι μέθοδοι μάθησης τείνουν να παράγουν εξειδικευμένα συστήματα ειδικά για κάθε εργασία, που είναι συχνά μη αποδοτικά έξω από τον στενό τομέα στον οποίο έχουν εκπαιδευτεί. Η μείωση του αριθμού των δειγμάτων με ανθρώπινη ετικέτα ή των αλληλεπιδράσεων με τον κόσμο που απαιτούνται για την εκμάθηση μιας εργασίας και η αύξηση της ευρωστίας εκτός τομέα είναι ζωτικής σημασίας για εφαρμογές όπως η μετάφραση γλώσσας χαμηλών πόρων, η ανάλυση ιατρικής εικόνας, η αυτόνομη οδήγηση, και φιλτράρισμα περιεχομένου.

Στην εποπτευόμενη μάθηση, μια ετικέτα για μία από τις κατηγορίες N μεταφέρει, κατά μέσο όρο, το πολύ $\log_2(N)$ bits πληροφοριών για τον κόσμο. Στην ενισχυτική μάθηση χωρίς μοντέλα μια ανταμοιβή μεταφέρει παρομοίως μόνο μερικά κομμάτια πληροφοριών. Αντίθετα ο ήχος, οι εικόνες και το βίντεο είναι τρόποι υψηλού εύρους ζώνης, που μεταφέρουν έμμεσα μεγάλες ποσότητες πληροφοριών σχετικά με τη δομή του κόσμου. Αυτό παρακινεί μια μορφή πρόβλεψης ή ανασυγκρότησης που ονομάζεται αυτό-εποπτευόμενη μάθηση η οποία είναι η εκπαίδευση για τη «συμπλήρωση των κενών» με την πρόβλεψη καλυμμένων ή κατεστραμμένων τμημάτων των δεδομένων. Η αυτό-εποπτευόμενη μάθηση ήταν πολύ επιτυχημένη για τους μετασχηματιστές εκπαίδευσης να εξάγουν διανύσματα που συλλαμβάνουν την εξαρτώμενη από το πλαίσιο έννοια μιας λέξης ή τμήματος λέξης.

Για κείμενο ο μετασχηματιστής εκπαιδεύεται να προβλέπει λέξεις που λείπουν από ένα διακριτό σύνολο δυνατοτήτων. Αλλά σε συνεχείς τομείς υψηλών διαστάσεων, όπως το βίντεο, όπου το σύνολο των εύλογων συνεχειών ενός συγκεκριμένου τμήματος βίντεο είναι μεγάλο και πολύπλοκο και αντιπροσωπεύει σωστά την κατανομή των εύλογων συνεχειών είναι ουσιαστικά ένα άλυτο πρόβλημα.

5.5 Αντιθετική μάθηση

Ένας τρόπος προσέγγισης αυτού του προβλήματος είναι μέσω μοντέλων λανθάνουσας μεταβλητής που αποδίδουν μια ενέργεια (δηλαδή ένα κακό) σε παραδείγματα βίντεο και μια πιθανή συνέχεια. Δεδομένου ενός βίντεο εισόδου X και μιας προτεινόμενης συνέχειας Y θέλουμε ένα μοντέλο να υποδεικνύει εάν το Y είναι συμβατό με το X χρησιμοποιώντας μια ενεργειακή συνάρτηση $E(X, Y)$ που παίρνει χαμηλές τιμές όταν τα X και Y είναι συμβατά και υψηλότερες τιμές διαφορετικά.

Το $E(X, Y)$ μπορεί να υπολογιστεί από ένα βαθύ νευρωνικό δίκτυο το οποίο, για ένα δεδομένο X , εκπαιδεύεται με αντιθετικό τρόπο ώστε να δίνει χαμηλή ενέργεια σε τιμές Y που είναι συμβατές με το X (όπως παραδείγματα (X, Y) ζεύγη από ένα σετ προπόνησης) και υψηλή ενέργεια σε άλλες τιμές του Y που δεν είναι συμβατές με το X . Για ένα δεδομένο X το συμπέρασμα συνίσταται στην εύρεση ενός Y^* που ελαχιστοποιεί το $E(X, Y)$ ή ίσως τη δειγματοληψία από τα Y που έχουν χαμηλές τιμές $E(X, Y)$. Αυτή η προσέγγιση βασισμένη στην ενέργεια για την αναπαράσταση του τρόπου με τον οποίο το Y εξαρτάται από το X καθιστά δυνατή τη μοντελοποίηση ενός διαφορετικού, πολυτροπικού συνόλου εύλογων συνεχειών.

Η βασική δυσκολία με την αντιθετική μάθηση είναι να επιλέξετε καλά «αρνητικά» δείγματα: κατάλληλα σημεία Y των οποίων η ενέργεια θα ωθηθεί προς τα πάνω. Όταν το σύνολο των πιθανών αρνητικών παραδειγμάτων δεν είναι πολύ μεγάλο, μπορούμε απλώς να τα εξετάσουμε όλα. Αυτό κάνει μια κανονικοποιημένη εκθετική συνάρτηση (softmax⁶), οπότε σε αυτή την περίπτωση η αντιθετική μάθηση μειώνεται σε τυπική εποπτευόμενη ή αυτό-εποπτευόμενη μάθηση σε ένα πεπερασμένο διακριτό σύνολο συμβόλων. Αλλά σε έναν χώρο υψηλών διαστάσεων πραγματικής αξίας, υπάρχουν πάρα πολλοί τρόποι με τους οποίους ένα διάνυσμα Y' θα μπορούσε να είναι διαφορετικό από το Y και για να βελτιώσουμε το μοντέλο πρέπει να εστιάσουμε σε εκείνα τα Y που θα έπρεπε να έχουν υψηλή ενέργεια αλλά επί του παρόντος έχουν χαμηλή ενέργεια.

5.6 Το μέλλον της βαθιάς μάθησης

Η απόδοση των συστημάτων βαθιάς μάθησης μπορεί συχνά να βελτιωθεί δραματικά με την απλή κλιμάκωση τους. Με πολλά περισσότερα δεδομένα και πολύ περισσότερους υπολογισμούς γενικά λειτουργούν πολύ καλύτερα. Το γλωσσικό μοντέλο GPT-3 με 175 δισεκατομμύρια παραμέτρους (που είναι ακόμα μικρό σε σύγκριση με τον αριθμό των συνάψεων στον ανθρώπινο εγκέφαλο) δημιουργεί αισθητά καλύτερο κείμενο από το GPT-2 με μόνο 1,5 δισεκατομμύρια παραμέτρους. Τα γλωσσικά μοντέλα (chatbot) Meena και BlenderBot συνεχίζουν να βελτιώνονται καθώς γίνονται μεγαλύτερα. Τώρα καταβάλλεται τεράστια προσπάθεια για την κλιμάκωση και θα βελτιώσει πολύ τα υπάρχοντα συστήματα, αλλά υπάρχουν θεμελιώδεις ελλείψεις της τρέχουσας βαθιάς μάθησης που δεν μπορούν να ξεπεραστούν μόνο με την κλιμάκωση. (Bengio et al., 2021)

⁶ Η συνάρτηση softmax, επίσης γνωστή ως softargmax ή κανονικοποιημένη εκθετική συνάρτηση, μετατρέπει ένα διάνυσμα K πραγματικών αριθμών σε κατανομή πιθανότητας K πιθανών αποτελεσμάτων. Είναι μια γενίκευση της λογιστικής συνάρτησης σε πολλαπλές διαστάσεις και χρησιμοποιείται στην πολυωνυμική λογιστική παλινδρόμηση.

Κεφάλαιο 6. Μεγάλα Μοντέλα Γλώσσας (Large Language Models - LLMs)

6.1 Εισαγωγή

Στον τομέα της Επεξεργασίας Φυσικής Γλώσσας (NLP) τα Μεγάλα Μοντέλα Γλωσσών (LLM) αντιπροσωπεύουν εξελιγμένα μοντέλα σχεδιασμένα για να κατανοούν, να δημιουργούν και να χειρίζονται κείμενο που μοιάζει με ανθρώπινη γλώσσα σε εκτεταμένη κλίμακα. Είναι αρχιτεκτονικές βαθιάς μάθησης (Deep Learning) που βασίζονται σε μετασχηματιστές, που λαμβάνονται μέσω της κλιμάκωσης του μεγέθους του μοντέλου, της προ-εκπαίδευσης των σωμάτων και των υπολογιστικών πόρων.

Τα μεγάλα μοντέλα γλώσσας (Large Language Models - LLM) είναι προηγμένα μοντέλα φυσικής επεξεργασίας γλώσσας (NLP - Natural Language Processing) στην κατηγορία των προ-εκπαιδευμένων γλωσσικών μοντέλων (PLM – Pre-trained Language Models), τα οποία επιτυγχάνονται μέσω της κλιμάκωσης του μεγέθους του μοντέλου, του προκαταρκτικού σώματος και των υπολογιστικών πόρων. Εν συντομία, τα LLM αναπτύσσονται μέσω μεθοδολογιών βαθιάς μάθησης, ιδιαίτερα χρησιμοποιώντας αρχιτεκτονικές μετασχηματιστών. Είναι μοντέλα νευρωνικών δικτύων που εφαρμόζουν μηχανισμούς αυτο-επίβλεψης που επιτρέπουν στο μοντέλο να εξετάζει ολόκληρο το πλαίσιο αντί να περιορίζεται σε παράθυρα σταθερού μεγέθους στην καταγραφή σχέσεων συμφραζομένων σε ακολουθίες εισόδου. Σε αυτή τη διαδικασία δεν απαιτούνται επαναλαμβανόμενα επίπεδα και στρώματα συνέλιξης. Άλλα κρίσιμα στοιχεία της αρχιτεκτονικής του μετασχηματιστή περιλαμβάνουν δομές κωδικοποιητή και αποκωδικοποιητή για να επεξεργαστούν αντίστοιχα την ακολουθία εισόδου και να δημιουργήσουν την ακολουθία εξόδου. Ωστόσο η αρχιτεκτονική ενός μετασχηματιστή μπορεί να ποικίλλει ανάλογα με το συγκεκριμένο έργο και το σχεδιασμό του. Ορισμένοι μετασχηματιστές έχουν σχεδιαστεί μόνο με δομή κωδικοποιητή ενώ άλλοι μπορεί να περιλαμβάνουν στοιχεία κωδικοποιητή και αποκωδικοποιητή. Για παράδειγμα, σε εργασίες όπως η μετάφραση γλώσσας, όπου εμπλέκονται αλληλουχίες εισόδου και εξόδου, απαιτούνται και οι δύο μονάδες κωδικοποιητή και αποκωδικοποιητή. Αντίθετα για μοντελοποίηση γλώσσας ή ταξινόμηση κειμένου, μπορεί να χρησιμοποιηθεί μόνο ένας κωδικοποιητής. Άλλα βασικά στοιχεία μιας αρχιτεκτονικής μετασχηματιστή περιλαμβάνουν νευρωνικά δίκτυα τροφοδοσίας για την καταγραφή πολύπλοκων, μη-γραμμικών σχέσεων στα δεδομένα και κωδικοποιήσεις θέσεων για την παροχή πληροφοριών σχετικά με τις θέσεις των διακριτικών στην ακολουθία.

6.2 Η εξάπλωση των Μεγάλων Μοντέλων Γλώσσας (LLM)

Ένα από τα βασικά χαρακτηριστικά των μεγάλων μοντέλων γλώσσας (LLM) είναι η ικανότητά τους να μαθαίνουν πληροφορίες σχετικά με τα συμφραζόμενα από μεγάλα σύνολα δεδομένων, επιτρέποντάς τους να κατανοούν πολύπλοκες γλωσσικές δομές και αποχρώσεις. Ως εκ τούτου οι εφαρμογές μεγάλων μοντέλων γλώσσας (LLM) χρησιμοποιούνται

αποτελεσματικά για την κατανόηση κειμένου, την αναγνώριση ομιλίας, τη δημιουργία και μετάφραση γλώσσας, τα γλωσσικά μοντέλα (chatbots) και την εικονική βοήθεια, τη συναισθηματική ανάλυση και άλλους σκοπούς.

Η ευρεία ενσωμάτωση του μεγάλου μοντέλου γλώσσας Chat-GPT (Chat Generative Pre-Trained Transformer) του OpenAI έχει προκαλέσει μεγάλο ενθουσιασμό από το ντεμπούτο του τον Νοέμβριο του 2022⁷. Μετά από αυτήν την κυκλοφορία, ο πολλαπλασιασμός νέων εργαλείων καθ' όλη τη διάρκεια του 2023, οδήγησε σε ένα δυναμικό τοπίο τεχνολογικής προόδου. Οι αρχιτεκτονικές και οι μέθοδοι εκπαίδευσης αυτών των οργάνων διαφέρουν και η λειτουργικότητά τους είναι εν μέρει κατανοητή όσον αφορά την ερμηνευτικότητα του μοντέλου, των εισόδων/χαρακτηριστικών/εξόδων, τη διαφάνεια της αρχιτεκτονικής του και τις μεθόδους εκπαίδευσης. Σε ορισμένες περιπτώσεις τα βάρη, δηλαδή οι παράμετροι που μαθαίνει το μοντέλο κατά τη φάση εκπαίδευσης και χρησιμοποιεί για αποφάσεις (έξοδος).

Παραμερίζοντας τις τεχνικές λεπτομέρειες, η φιλική προς τον χρήστη διεπαφή και η διαθεσιμότητα ανοιχτών αδειών, ιδιαίτερα για βασικά πλαίσια, υπήρξαν βασικοί παράγοντες για τη γρήγορη εξάπλωση αυτών των συστημάτων. Υπόσχονται εφαρμογές υγειονομικής περίθαλψης, ιδιαίτερα στην ανάπτυξη chatbot, συστημάτων αλληλεπίδρασης για τη διαχείριση κλινικής τεκμηρίωσης και τη σύνοψη της ιατρικής βιβλιογραφίας.

6.3 Κατηγορίες μεγάλων μοντέλων γλώσσας (LLM)

Η καινοτόμος αρχιτεκτονική μετασχηματιστών έχει ανοίξει το δρόμο για την ανάπτυξη διαφόρων LLM, καθένα από τα οποία διακρίνεται για τα μοναδικά χαρακτηριστικά του. Οι πρόσφατες εξελίξεις στη μοντελοποίηση γλώσσας οδήγησαν στην εμφάνιση τριών κυρίαρχων κατηγοριών, που ταξινομούνται με βάση τις θεμελιώδεις ενότητες που χρησιμοποιούνται στην κατασκευή τους. Πρώτον, υπάρχουν LLM μόνο για κωδικοποιητές που παραδειγματίζονται από το BERT (Bidirectional Encoder Representations from Transformers - Αμφίδρομες Αναπαραστάσεις Κωδικοποιητή από τους Μετασχηματιστές) και τις διάφορες επαναλήψεις του. Αυτά τα μοντέλα υπερέρχουν στην αποτύπωση πληροφοριών συμφραζομένων αμφίδρομα, ενισχύοντας μια ολοκληρωμένη κατανόηση της γλωσσικής σημασιολογίας. Δεύτερον, τα μοντέλα γλώσσας μόνο για αποκωδικοποιητές, όπως επισημάνθηκαν από τα μέλη της οικογένειας GPT, τονίζουν τη δημιουργία συνεκτικών και σχετικών με τα συμφραζόμενα ακολουθιών. Αξιοποιώντας μονοκατευθυντικά μπλοκ προσοχής, αυτά τα μοντέλα έχουν επιδείξει επάρκεια σε εργασίες που απαιτούν διαδοχική κατανόηση και δημιουργία. Τέλος, τα μοντέλα γλώσσας κωδικοποιητή-αποκωδικοποιητή, όπως το T5 (Text-to-Text Transfer Transformer - Μετασχηματιστής μεταφοράς κειμένου σε κείμενο) και το BART (Bidirectional and AutoRegressive Transformers - Μετασχηματιστές διπλής κατεύθυνσης και αυτόματης παλινδρόμησης), αντιπροσωπεύουν μια συγχώνευση αμφίδρομων και μονοκατευθυντικών μηχανισμών προσοχής. Αυτή η υβριδική προσέγγιση

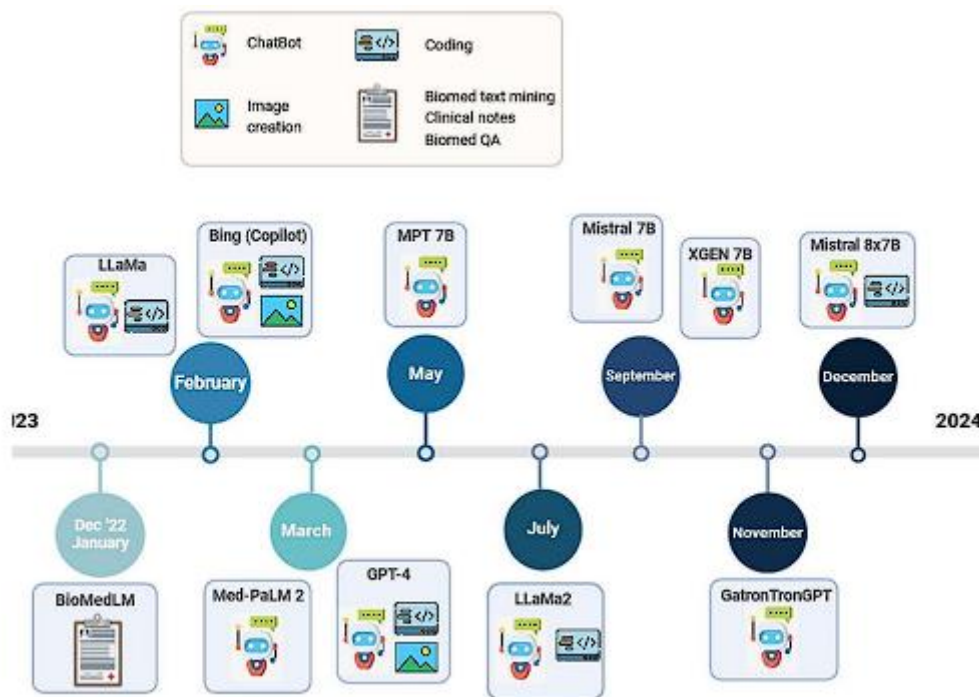
⁷ Open AI. ChatGPT release note. Available at: https://help.openai.com/en/articles/6825453-chatgpt-release-notes#h_4799933861 Last Accessed: December 22, 2023.

επιτρέπει ευέλικτες εφαρμογές, που κυμαίνονται από τη σύνοψη κειμένου έως τη μετάφραση γλώσσας, όπου η κατανόηση του πλαισίου και η δημιουργία συνεκτικών απαντήσεων είναι και τα δύο ζωτικής σημασίας. Η εφαρμογή διαφορετικών διαδικασιών και συνόλων δεδομένων επιτρέπει την παροχή ενός φάσματος εργαλείων προσαρμοσμένων στις εξελισσόμενες απαιτήσεις της κατανόησης και παραγωγής φυσικής γλώσσας.

6.4 Οι βελτιώσεις στο χώρο εξαιτίας των LLM

Πριν από την άνοδο των LLM, τα παραδοσιακά μοντέλα βαθιάς μάθησης αντιμετώπιζαν πολλές τεχνικές προκλήσεις, συμπεριλαμβανομένης της ανεπαρκούς αλληλουχίας και της σημασιολογικής κατανόησης, μαζί με την υπολογιστική πολυπλοκότητα. Αυτή η πολυπλοκότητα απαιτούσε έναν σημαντικό αριθμό παραμέτρων, όπως είναι εμφανές στα συνελκτικά νευρωνικά δίκτυα, για να επιτευχθούν ικανοποιητικά αποτελέσματα. Επιπλέον ζητήματα όπως οι διαβαθμίσεις που εξαφανίζονται (π.χ. σε επαναλαμβανόμενα νευρωνικά δίκτυα) θέτουν προκλήσεις στην αποτύπωση μακροπρόθεσμων εξαρτήσεων ενώ ο διαδοχικός υπολογισμός εμπόδιζε την αποτελεσματικότητα της εκπαίδευσης και των διαδικασιών εξαγωγής συμπερασμάτων, ιδιαίτερα για εκτεταμένες ακολουθίες. Η πραγματική καινοτομία προήλθε από τη βελτιστοποίηση προ-εκπαιδευμένων μοντέλων γλώσσας για να ανταποκρίνονται στις συγκεκριμένες απαιτήσεις των εργασιών που προσανατολίζονται στη συνομιλία, επιτυγχάνοντας έτσι βελτιωμένη απόδοση σε εφαρμογές που περιλαμβάνουν εκδόσεις με λεπτομέρεια διαλόγου για αλληλεπιδράσεις συνομιλίας. Αυτή η πολύπλοκη διαδικασία παρέχει τη χρήση διαφορετικών προσεγγίσεων. Η λεπτομέρεια που βασίζεται σε συνομιλία, για παράδειγμα, αναφέρεται στον τρόπο βελτίωσης ενός προ-υπάρχοντος LLM μέσω πρόσθετης εκπαίδευσης ειδικά προσαρμοσμένης σε περιβάλλοντα συνομιλίας. Σε αυτή την προσέγγιση, το μοντέλο ρυθμίζεται με ακρίβεια χρησιμοποιώντας σύνολα δεδομένων που αποτελούνται από διαλόγους ή αλληλεπιδράσεις, συχνά με τη μορφή ζευγών μηνύματος-απόκρισης. Επομένως το μοντέλο μπορεί να μάθει τις αποχρώσεις των αλληλεπιδράσεων φυσικής γλώσσας, συμπεριλαμβανομένης της ροής της συνομιλίας, του χειρισμού του περιβάλλοντος και της δημιουργίας κατάλληλων απαντήσεων. Μπορεί να κατανοεί τα ερωτήματα των χρηστών, να διατηρεί το πλαίσιο σε όλες τις στροφές και να παρέχει συναφείς και συνεκτικές απαντήσεις. Το προ-εκπαιδευμένο μοντέλο μπορεί επίσης να παρέχεται με πρόσθετα δεδομένα εκπαίδευσης ειδικά προσαρμοσμένα σε οδηγίες ή κατευθυντήριες γραμμές. Αυτή η μέθοδος μικρορύθμισης βοηθά στη δημιουργία εξόδων που ευθυγραμμίζονται πιο στενά με τις επιθυμητές οδηγίες, βελτιώνοντας την απόδοσή της σε εργασίες που απαιτούν ρητή καθοδήγηση. Επιπλέον, η ενισχυτική μάθηση από την ανθρώπινη ανάδραση (RLHF - Reinforcement Learning from Human Feedback) είναι μια διαδικασία που απαιτεί τη συμμετοχή των ανθρώπων στην κατάταξη της αποτελεσματικότητας της παραγωγής (ανθρώπινη ανάδραση). Η διαδικασία RLHF ήταν θεμελιώδης για την επιτυχία του ChatGPT. Είναι αξιοσημείωτο ότι κάθε ένα από τα βήματα μικρορύθμισης ή RLHF μπορεί να εκτελεστεί είτε ανεξάρτητα είτε διαδοχικά. Αυτή η ευελιξία είναι ιδιαίτερα σημαντική, δεδομένου ότι πολλά μοντέλα συνομιλίας LLM

υποβάλλονται σε πολλαπλά στάδια εκπαίδευσης, οδηγώντας τελικά σε βελτιωμένη απόδοση και αποτελεσματικότητα σε διάφορες εργασίες κατανόησης φυσικής γλώσσας και δημιουργίας. Τέλος η προσέγγιση βελτιστοποίησης άμεσης προτίμησης (DPO) στοχεύει στην άμεση βελτιστοποίηση του μοντέλου για τις προτιμήσεις των χρηστών ή τα επιθυμητά αποτελέσματα. Προτείνεται να είναι εναλλακτική του RLHF. Όσον αφορά τη λειτουργία του, το DPO παρακάμπτει τα ενδιάμεσα βήματα και στοχεύει άμεσα στη βελτιστοποίηση των κριτηρίων που καθορίζονται από τον χρήστη. Ως εκ τούτου, το μοντέλο μπορεί να επικεντρωθεί σε πιο σχετικά και ικανοποιητικά αποτελέσματα για τους χρήστες που επιτυγχάνουν εξατομικευμένα και συναφή αποτελέσματα (Casella et al., 2024).



Εικόνα 4. Ενδεικτική χρονοσειρά ανάπτυξης διάφορων MML. (Πηγή: Casella et al., 2024)




Κεφάλαιο 7. Τεχνητή Νοημοσύνη σε διάφορους τομείς

7.1 Αρχιτεκτονική

Στον τομέα της αρχιτεκτονικής, της μηχανικής, των κατασκευών και των έργων διαχείρισης εγκαταστάσεων, οι έννοιες και οι απαιτήσεις σχεδιασμού παραδοσιακά μεταφέρονται μέσω ενός συνδυασμού οπτικών αναπαραστάσεων και γραπτών και λεκτικών περιγραφών. Αυτές οι οπτικές εικόνες, μαζί με τις συνοδευτικές επεξηγήσεις κειμένου, είναι ζωτικής σημασίας για την αποτελεσματική επικοινωνία των προδιαγραφών σχεδιασμού στους ενδιαφερόμενους φορείς που εμπλέκονται στη διαδικασία. Με την εξέλιξη των υπολογιστών και των ψηφιακών τεχνολογιών, η αρχιτεκτονική οπτικοποίηση που διευκολύνεται από τεχνικές απόδοσης υψηλής απόδοσης που δημιουργούν ρεαλιστικές απεικονίσεις παρόμοιες με τη φωτογραφία, συνεχίζει να διαδραματίζει κεντρικό ρόλο στην επικοινωνία σχεδιασμού. Η πρόσφατη προσοχή έχει στραφεί προς τις τεχνικές γενετικής τεχνητής νοημοσύνης (AI) που βασίζονται σε μοντέλα κειμένου σε εικόνα για οπτικές αναπαραστάσεις, οι οποίες συγκαταλέγονται στις μεγάλες προόδους στην τεχνολογία υπολογιστών. Αυτή η τεχνολογία αναμένεται να παρέχει καινοτόμες μεθοδολογίες για την αυτοματοποίηση της οπτικοποίησης στην αρχιτεκτονική βιομηχανία. Στην αρχιτεκτονική οπτικοποίηση ο σχεδιασμός των προσώπων κτιρίων μπορεί να έχει σημαντική συμβολική αναπαράσταση. Προέρχεται από τη γαλλική λέξη *facade* που σημαίνει «μέτωπο» ή «πρόσωπο», η πρόσοψη ενός κτιρίου παίζει καθοριστικό ρόλο στην ενοποίηση της οπτικής και αισθητικής πτυχής ενός ολόκληρου κτιρίου και χρησιμεύει ως ενδιάμεσος μεταξύ εξωτερικών και εσωτερικών συνθηκών. Η πρόσοψη του κτιρίου, από την άποψη της οπτικής της πτυχής, μπορεί να γίνει φόντο στους αστικούς χώρους και την περιφερειακή ζωή και να επηρεάσει τον τρόπο με τον οποίο οι άνθρωποι αντιλαμβάνονται και προσαρμόζονται στους αστικούς χώρους και την περιφερειακή ζωή, ακόμα και στο φυσικό πεδίο.

Στο παρελθόν, η διαδικασία του αρχιτεκτονικού σχεδιασμού ξεκινούσε με το χειροκίνητο σχέδιο. Ωστόσο, με την πρόοδο της τεχνολογίας των υπολογιστών, διάφορες καινοτομίες και εξελίξεις, όπως το λογισμικό σχεδιασμού με τη βοήθεια υπολογιστή (CAD) και η εισαγωγή των εννοιών της μοντελοποίησης πληροφοριών κτιρίων (BIM), έχουν οδηγήσει σε απλοποίηση και αυτοματοποίηση της σύνθετης διαδικασίας αρχιτεκτονικού σχεδιασμού. Αυτές οι εξελίξεις έχουν βελτιώσει σημαντικά την όλη διαδικασία σχεδιασμού. Αυτή η εξέλιξη είναι συνυφασμένη με καινοτομίες στο λογισμικό, τους υπολογιστές και την οπτικοποίηση, οι οποίες έχουν αλλάξει θεμελιωδώς τον τρόπο με τον οποίο συλλαμβάνονται, επικοινωνούν και παρουσιάζονται τα αρχιτεκτονικά σχέδια.

Ιδιαίτερα αξιοσημείωτη ήταν η εμφάνιση των Παραγωγικών Αντιπαλικών Δικτύων (GAN⁸s) το 2014, που σηματοδότησε μια αλλαγή παραδείγματος στην έρευνα δημιουργίας εικόνων. Τα GAN έχουν επιδείξει την ικανότητα να παράγουν ρεαλιστικές εικόνες, χρησιμοποιώντας μια ανταγωνιστική διαδικασία εκπαίδευσης που περιλαμβάνει έναν δημιουργό και έναν διακριτικό. Με την ωρίμανση των GAN οι μελετητές έχουν ξεκινήσει την εξερεύνηση τεχνικών για την ενίσχυση της σταθερότητας της εκπαίδευσης GAN. Συγκεκριμένα, το Google DeepMind αποκάλυψε ευρήματα που δείχνουν ότι τα μεγαλύτερα μεγέθη παρτίδας οδηγούν σε βελτιωμένη ποιότητα εικόνας, αν και τα μεγαλύτερα μεγέθη παρτίδας απαιτούν ισχυρούς υπολογιστικούς πόρους. Καθώς οι στρατηγικές για την εκπαίδευση GAN σταθεροποιήθηκαν, το εστιακό σημείο μετατοπίστηκε προς τη δημιουργία εικόνων προικισμένων με συγκεκριμένα χαρακτηριστικά και τον χειρισμό των παραγόμενων εικόνων (Jo et al., 2024).

		A (OpenAI, n.d.)	B (Midjourney, n.d.)	C (Dreamstudio, n.d.)
INPUT	Case	Building facade		
	Trigger words	Realistic architecture, Building facade image		
OUTPUT	Generated Image			

Εικόνα 5. Παράδειγμα εξόδου εικόνων που παράχθηκαν από πλατφόρμες δημιουργίας εικόνων (Πηγή: Jo et al., 2024)

7.2 Τεχνητή Νοημοσύνη και Κυβερνοασφάλεια

Με την ολοένα και πιο βαθιά ενοποίηση του Διαδικτύου και της κοινωνικής ζωής το Διαδίκτυο αλλάζει τον τρόπο με τον οποίο μαθαίνουν και εργάζονται οι άνθρωποι αλλά μας εκθέτει επίσης σε όλο και πιο σοβαρές απειλές για την ασφάλεια. Ο τρόπος με τον οποίο αναγνωρίζονται διάφορες επιθέσεις δικτύου, ιδιαίτερα επιθέσεις που δεν έχουν παρατηρηθεί στο παρελθόν, είναι ένα βασικό ζήτημα που πρέπει να επιλυθεί επείγοντως.

Η κυβερνοασφάλεια είναι ένα σύνολο τεχνολογιών και διαδικασιών που έχουν σχεδιαστεί για την προστασία υπολογιστών, δικτύων, προγραμμάτων και δεδομένων από επιθέσεις και μη εξουσιοδοτημένη πρόσβαση, τροποποίηση ή καταστροφή. Ένα σύστημα ασφάλειας δικτύου αποτελείται από ένα σύστημα ασφάλειας δικτύου και ένα σύστημα

⁸ Τα Παραγωγικά Αντιπαλικά Δίκτυα (ΠΑΔ), γνωστά επίσης ως Αντιπαλικά Δίκτυα, Παραγωγικά Ανταγωνιστικά Δίκτυα, Παραγωγικά Δίκτυα Αντιπάλων και Αναγεννητικά Ανταγωνιστικά Δίκτυα (στα αγγλικά Generative Adversarial Networks - GAN) είναι μια κατηγορία συστημάτων μηχανικής μάθησης που εφευρέθηκε από τον Ian Goodfellow και τους συναδέλφους του το 2014.

ασφαλείας υπολογιστή. Καθένα από αυτά τα συστήματα περιλαμβάνει τείχη προστασίας, λογισμικό προστασίας από ιούς και συστήματα ανίχνευσης εισβολής (IDS). Τα IDS βοηθούν στην ανακάλυψη, τον προσδιορισμό και τον εντοπισμό μη εξουσιοδοτημένης συμπεριφοράς του συστήματος, όπως χρήση, αντιγραφή, τροποποίηση και καταστροφή. Οι παραβιάσεις ασφαλείας περιλαμβάνουν εξωτερικές και εσωτερικές εισβολές. Υπάρχουν τρεις κύριοι τύποι ανάλυσης δικτύου για τα IDS: βασισμένη σε κακή χρήση, γνωστή και ως βάση υπογραφής, βάσει ανωμαλίας και υβριδική. Οι τεχνικές ανίχνευσης που βασίζονται σε κακή χρήση στοχεύουν στον εντοπισμό γνωστών επιθέσεων χρησιμοποιώντας τις υπογραφές αυτών των επιθέσεων. Χρησιμοποιούνται για γνωστούς τύπους επιθέσεων χωρίς να δημιουργούν μεγάλο αριθμό ψευδών συναγεμίων. Ωστόσο οι διαχειριστές πρέπει συχνά να ενημερώνουν μη αυτόματα τους κανόνες και τις υπογραφές της βάσης δεδομένων. Δεν είναι δυνατή η ανίχνευση νέων επιθέσεων (zero-day) βάσει τεχνολογιών που χρησιμοποιούνται κατά λάθος. Οι τεχνικές που βασίζονται σε ανωμαλίες μελετούν την κανονική συμπεριφορά του δικτύου και του συστήματος και εντοπίζουν τις ανωμαλίες ως αποκλίσεις από την κανονική συμπεριφορά. Είναι ελκυστικά λόγω της ικανότητάς τους να εντοπίζουν επιθέσεις μηδενικής ημέρας. Ένα άλλο πλεονέκτημα είναι ότι τα προφίλ της κανονικής δραστηριότητας προσαρμόζονται για κάθε σύστημα, εφαρμογή ή δίκτυο, επομένως καθιστά δύσκολο για τους εισβολείς να γνωρίζουν ποιες δραστηριότητες μπορούν να πραγματοποιήσουν χωρίς να εντοπιστούν. Επιπλέον τα δεδομένα για τα οποία προειδοποιούν τεχνικές που βασίζονται σε ανωμαλίες (νέες επιθέσεις) μπορούν να χρησιμοποιηθούν για τον καθορισμό των υπογραφών για ανιχνευτές κακής χρήσης. Το κύριο μειονέκτημα των τεχνικών που βασίζονται σε ανωμαλίες είναι η πιθανότητα για υψηλά ποσοστά ψευδών συναγεμίων επειδή συμπεριφορές συστήματος που δεν είχαν παρατηρηθεί προηγουμένως μπορούν να κατηγοριοποιηθούν ως ανωμαλίες. Η υβριδική ανίχνευση συνδυάζει κακή χρήση και ανίχνευση ανωμαλιών. Χρησιμοποιείται για την αύξηση του ποσοστού ανίχνευσης γνωστών εισβολών και για τη μείωση του ψευδώς θετικού ποσοστού άγνωστων επιθέσεων. Οι περισσότερες μέθοδοι ML/DL είναι υβριδικές.

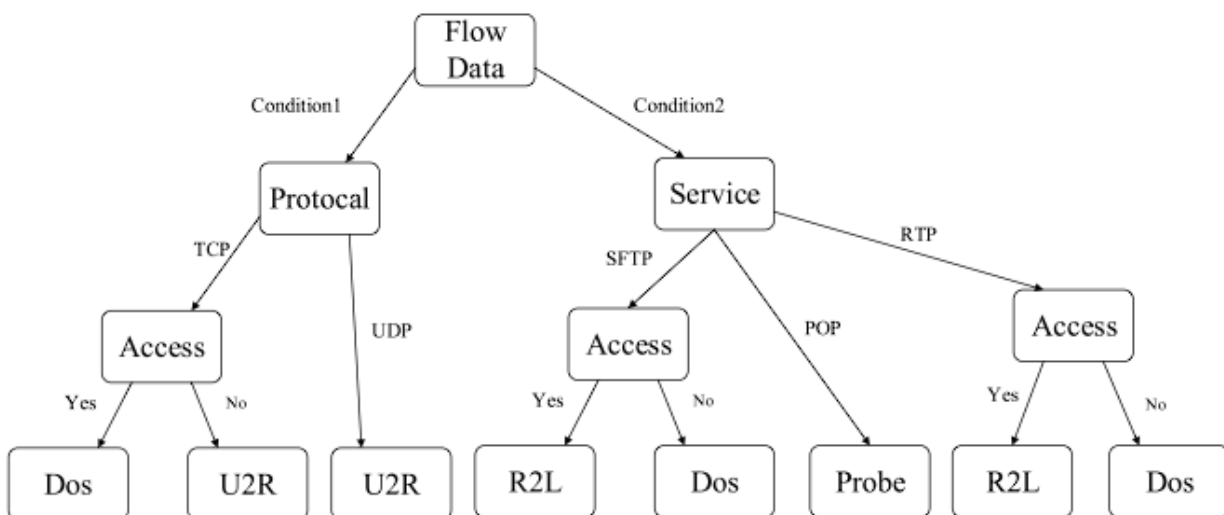
7.2.1 Support Vector Machine

Το Support Vector Machine (SVM) είναι μια από τις πιο ισχυρές και ακριβείς μεθόδους σε όλους τους αλγόριθμους μηχανικής μάθησης. Περιλαμβάνει κυρίως την ταξινόμηση διανυσμάτων υποστήριξης (SVC) και την παλινδρόμηση διανυσμάτων υποστήριξης (SVR). Το SVC βασίζεται στην έννοια των ορίων απόφασης. Ένα όριο απόφασης διαχωρίζει ένα σύνολο περιπτώσεων που έχουν διαφορετικές τιμές κλάσης μεταξύ δύο ομάδων. Το SVC υποστηρίζει τόσο δυαδικές όσο και πολλαπλές ταξινομήσεις. Το διάνυσμα υποστήριξης είναι το πλησιέστερο σημείο στο υπερεπίπεδο διαχωρισμού, το οποίο καθορίζει το βέλτιστο υπερεπίπεδο διαχωρισμού. Στη διαδικασία ταξινόμησης, τα διανύσματα εισόδου χαρτογράφησης που βρίσκονται στην πλευρά του υπερεπίπεδου διαχωρισμού του χώρου χαρακτηριστικών εμπίπτουν σε μια κατηγορία και οι θέσεις εμπίπτουν στην άλλη κατηγορία στην άλλη πλευρά του επιπέδου. Στην περίπτωση σημείων δεδομένων που δεν διαχωρίζονται γραμμικά, το SVM χρησιμοποιεί κατάλληλες συναρτήσεις

πυρήνα για να τα αντιστοιχίσει σε χώρους υψηλότερων διαστάσεων έτσι ώστε να διαχωρίζονται σε αυτούς τους χώρους.

7.2.2 Δέντρο Απόφασης (Decision Tree)

Ένα δέντρο απόφασης είναι μια δομή δέντρου στην οποία κάθε εσωτερικός κόμβος αντιπροσωπεύει μια δοκιμή σε μια ιδιότητα και κάθε κλάδος αντιπροσωπεύει μια έξοδο δοκιμής, με κάθε κόμβο φύλλου να αντιπροσωπεύει μια κατηγορία. Στη μηχανική μάθηση το δέντρο αποφάσεων είναι ένα μοντέλο πρόβλεψης, αντιπροσωπεύει μια αντιστοίχιση μεταξύ των χαρακτηριστικών αντικειμένων και των τιμών των αντικειμένων. Κάθε κόμβος στο δέντρο αντιπροσωπεύει ένα αντικείμενο, κάθε διαδρομή απόκλισης αντιπροσωπεύει μια πιθανή τιμή χαρακτηριστικού και κάθε κόμβος φύλλου αντιστοιχεί στην τιμή του αντικειμένου που αντιπροσωπεύεται από τη διαδρομή από τον κόμβο ρίζας προς τον κόμβο φύλλου. Το δέντρο



Εικόνα 6. Ένα παράδειγμα δέντρου αποφάσεων (Πηγή: Xin et al.)

απόφασης έχει μόνο μία έξοδο. Εάν θέλετε σύνθετη έξοδο μπορείτε να δημιουργήσετε ένα ανεξάρτητο δέντρο αποφάσεων για να χειριστείτε διαφορετικές εξόδους. Τα μοντέλα δέντρων αποφάσεων που χρησιμοποιούνται συνήθως είναι τα ID3, C4.5 και CART.

7.2.3 Δίκτυο Βαθιάς Πεποίθησης - Deep Belief Network

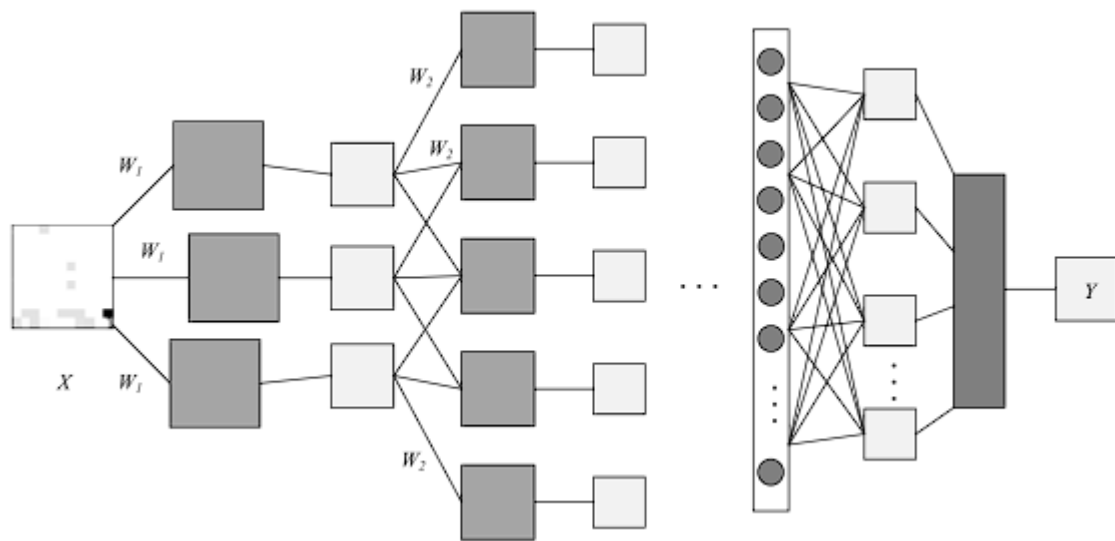
Το Δίκτυο Βαθιάς Πεποίθησης - Deep Belief Network (DBN) είναι ένα πιθανολογικό παραγωγικό μοντέλο που αποτελείται από πολλαπλά επίπεδα στοχαστικών και κρυφών μεταβλητών. Το Restricted Boltzmann Machine (RBM) και το DBN είναι αλληλένδετα, επειδή η σύνθεση και η στοίβαξη ενός αριθμού RBM επιτρέπει σε πολλά κρυφά επίπεδα να εκπαιδεύουν δεδομένα αποτελεσματικά μέσω ενεργοποιήσεων ενός RBM για περαιτέρω στάδια εκπαίδευσης. Το RBM είναι μια ειδική τοπολογική δομή μιας μηχανής Boltzmann

(BM). Η αρχή του BM προήλθε από τη στατιστική φυσική ως μέθοδος μοντελοποίησης που βασίζεται σε μια ενεργειακή συνάρτηση που μπορεί να περιγράψει τις αλληλεπιδράσεις υψηλής τάξης μεταξύ των μεταβλητών. Το BM είναι ένα συμμετρικό συζευγμένο νευρωνικό δίκτυο δυαδικής μονάδας τυχαίας ανάδρασης που αποτελείται από ένα ορατό στρώμα και ένα πλήθος κρυφών επιπέδων. Ο κόμβος δικτύου χωρίζεται σε μια ορατή μονάδα και μια κρυφή μονάδα και η ορατή μονάδα και η κρυφή μονάδα χρησιμοποιούνται για να εκφράσουν ένα τυχαίο δίκτυο και ένα τυχαίο περιβάλλον. Το μοντέλο μάθησης εκφράζει τη συσχέτιση μεταξύ των μονάδων με στάθμιση.

7.2.4 Αναδρομικά Νευρωνικά Δίκτυα (Convolutional Neural Networks)

Το αναδρομικό νευρωνικό δίκτυο (RNN) χρησιμοποιείται για την επεξεργασία δεδομένων ακολουθίας. Στο παραδοσιακό μοντέλο νευρωνικών δικτύων δεδομένα από το επίπεδο εισόδου στο κρυφό στρώμα στο επίπεδο εξόδου. Τα επίπεδα είναι πλήρως συνδεδεμένα και δεν υπάρχει σύνδεση μεταξύ των κόμβων μεταξύ κάθε επιπέδου. Υπάρχουν πολλά προβλήματα που αυτό το συμβατικό νευρωνικό δίκτυο δεν μπορεί να λύσει. Τα συνελκτικά νευρωνικά δίκτυα (CNN) είναι ένας τύπος τεχνητού νευρωνικού δικτύου που έχει γίνει hotspot στον τομέα της ανάλυσης ομιλίας και της αναγνώρισης εικόνων. Η δομή του δικτύου κατανομής βάρους το καθιστά περισσότερο παρόμοιο με ένα βιολογικό νευρωνικό δίκτυο, μειώνοντας έτσι την πολυπλοκότητα του μοντέλου δικτύου και μειώνοντας τον αριθμό των βαρών. Αυτό το πλεονέκτημα είναι πιο εμφανές όταν η είσοδος δικτύου είναι μια πολυδιάστατη εικόνα και η εικόνα μπορεί να χρησιμοποιηθεί απευθείας ως είσοδος του δικτύου για να αποφευχθεί η σύνθετη εξαγωγή χαρακτηριστικών και η ανακατασκευή δεδομένων στον παραδοσιακό αλγόριθμο αναγνώρισης. Το Συνελκτικό Δίκτυο είναι ένας πολυεπίπεδος αισθητήρας, ειδικά σχεδιασμένος για να αναγνωρίζει δισδιάστατα σχήματα που είναι πολύ αμετάβλητα στη μετάφραση, την κλιμάκωση, την κλίση ή άλλες μορφές παραμόρφωσης. Το CNN είναι ο πρώτος πραγματικά επιτυχημένος αλγόριθμος εκμάθησης για την εκπαίδευση δομών δικτύων πολλαπλών επιπέδων. Μειώνει τον αριθμό των παραμέτρων που πρέπει να μάθουμε για τη βελτίωση της απόδοσης εκπαίδευσης του αλγορίθμου BP μέσω χωρικών σχέσεων. Ως αρχιτεκτονική βαθιάς μάθησης, το CNN προτείνεται να ελαχιστοποιεί τις απαιτήσεις προ επεξεργασίας δεδομένων. Υπάρχουν τρία κύρια μέσα για το CNN για τη μείωση των παραμέτρων εκπαίδευσης δικτύου: τοπική δεκτικότητα, κοινή χρήση βάρους και συγκέντρωση. Το πιο ισχυρό μέρος του CNN είναι οι ιεραρχίες χαρακτηριστικών εκμάθησης από μεγάλες ποσότητες δεδομένων χωρίς ετικέτα. Ως εκ τούτου, τα CNN είναι πολλά υποσχόμενα για εφαρμογή στο πεδίο ανίχνευσης εισβολής δικτύου.

Συγκριτική Μελέτη εργαλείων Γενετικής Τεχνητής Νοημοσύνης



Εικόνα 7. Ένα παράδειγμα συνελκτικού νευρωνικού δικτύου (Πηγή: Xin et al.)

7.3 Κατηγοριοποίηση οδικών ατυχημάτων

Σε μία δημοσίευση με τίτλο «A Distributed Big Data Analytics Model for Traffic Accidents Classification and Recognition based on SparkMLLib Cores» ο κ. El Mallahi και η ομάδα του, προτείνουν μια λύση που βασίζεται σε μεγάλα δεδομένα και μοντέλα μηχανικής μάθησης για την πρόβλεψη τροχαίων ατυχημάτων. Αυτή η επεξεργασία γίνεται καταγράφοντας σε πραγματικό χρόνο όλα τα εισερχόμενα σύνολα δεδομένων, τα οποία είναι αποθηκευμένα στο σύμπλεγμα κατακευματισμένου συστήματος αρχείων Hadoop. Στη συνέχεια μέσω κλήσης του πυρήνα SparkMLLib, ο οποίος χρησιμοποιεί όλες τις προ-υλοποιημένες λειτουργίες LamBig Data Analytic. Στη συνέχεια εστίασαν στην πρόβλεψη σοβαρότητας για τροχαία ατυχήματα, η οποία είναι ένα τεράστιο βήμα στη διαχείριση τροχαίων ατυχημάτων. Για να αξιολογήσουν τη σοβαρότητα των τροχαίων ατυχημάτων, αξιολόγησαν τον πιθανό αντίκτυπο του ατυχήματος και υλοποίησαν αποτελεσματικές διαδικασίες διαχείρισης ατυχημάτων. Σε αυτήν την προτεινόμενη μελέτη, εφάρμοσαν μερικούς αλγόριθμους για την ταξινόμηση της σοβαρότητας των τροχαίων ατυχημάτων και παρουσίασαν τη μήτρα σύγχυσης για να καθοριστούν οι έννοιες: Πεζός, Όχημα ή Συνεπιβάτης, ή Οδηγός ή Αναβάτης. Για την επικύρωση αυτού του πειραματισμού, χρησιμοποιήθηκε το σύνολο δεδομένων TRAFFIC_ACCIDENTS_2019_LEEDS για την ταξινόμηση της πρόβλεψης σοβαρότητας για τροχαία ατυχήματα σε τρεις κατηγορίες: Πεζός, Όχημα ή Συνεπιβάτης ή Οδηγός ή Αναβάτης. Επίσης μπορούμε να εφαρμόσουμε το κόστος για την πρόβλεψη της βαρύτητας των τροχαίων ατυχημάτων. Συμπερασματικά στο τέλος αναφέρουν ότι το πολύ σημαντικό όφελος από τη χρήση του παραδείγματος μεγάλων δεδομένων είναι ότι βελτιώνει την επεξεργασία των δεδομένων και καθορίζει ένα καλό ποσοστό στην οδική ασφάλεια με βάση την ταξινόμηση και την αναγνώριση των τροχαίων ατυχημάτων.

Κεφάλαιο 8: Chat-GPT & DALL-E

8.1 Εισαγωγή

Η ανάπτυξη προηγμένων μοντέλων όπως το GPT-3 και το DALL-E έφερε επανάσταση στον τρόπο με τον οποίο προσεγγίζουμε αυτά τα δύο πεδία. Το GPT-3 είναι ένα ευέλικτο μοντέλο επεξεργασίας γλώσσας που μπορεί να εκτελέσει μια ποικιλία εργασιών, συμπεριλαμβανομένης της επεξεργασίας φυσικής γλώσσας, των γλωσσικών μοντέλων chatbot, των συστημάτων απάντησης σε ερωτήσεις και της δημιουργίας κειμένου, ενώ το DALL-E, ένα μοντέλο επεξεργασίας εικόνας, είναι ικανό να δημιουργεί εικόνες από κείμενο. Αυτά τα μοντέλα έχουν δείξει πολλά υποσχόμενα αποτελέσματα σε διάφορες εφαρμογές και έχουν χρησιμοποιηθεί σε ένα ευρύ φάσμα εφαρμογών λογισμικού. Όντας ένα μοντέλο γλώσσας μεγάλης κλίμακας το GPT-3 μπορεί να κατανοήσει και να δημιουργήσει φυσική γλώσσα με επάρκεια. Εκπαιδευμένο σε ένα τεράστιο σύνολο δεδομένων κειμένου, έχει τη δυνατότητα να εκτελεί διάφορες εργασίες επεξεργασίας γλώσσας, όπως μετάφραση γλώσσας, περίληψη και ανάλυση συναισθημάτων. Επιπλέον το GPT-3 είναι ικανό να δημιουργεί κείμενο που μοιάζει με άνθρωπο, το οποίο μπορεί να χρησιμοποιηθεί σε εφαρμογές όπως τα chatbots και η αυτοματοποιημένη παραγωγή περιεχομένου. Το DALL-E, από την άλλη πλευρά, είναι ένα μοντέλο παραγωγής εικόνας που βασίζεται σε νευρωνικά δίκτυα. Μπορεί να δημιουργήσει εικόνες υψηλής ποιότητας από μηνύματα κειμένου, επιτρέποντας στους χρήστες να δημιουργήσουν ένα ευρύ φάσμα εικόνων με λίγες μόνο λέξεις. Αυτή η τεχνολογία έχει σημαντικές εφαρμογές στη δημιουργική βιομηχανία όπως στον σχεδιασμό γραφικών, τη διαφήμιση και την ψυχαγωγία.

8.2 Στα ενδότερα του Chat-GPT

Αναπτύχθηκε από την OpenAI, το Generative Pre-trained Transformer (GPT), είναι ένας τύπος μοντέλου επεξεργασίας γλώσσας. Με προ-εκπαίδευση σε τεράστιες ποσότητες συνόλων δεδομένων κειμένου, είναι ένα νευρωνικό δίκτυο μεγάλης κλίμακας, ικανό να εκτελέσει ένα ευρύ φάσμα εργασιών επεξεργασίας γλώσσας οι οποίες περιλαμβάνουν μετάφραση γλώσσας, περίληψη και ανάλυση συναισθήματος. Επιπλέον το GPT μπορεί να παράγει κείμενο που μοιάζει με ανθρώπινη γραφή, καθιστώντας το πολύτιμο εργαλείο σε εφαρμογές όπως τα chatbot και η αυτοματοποιημένη παραγωγή περιεχομένου. Η ανάπτυξη του GPT είναι σημαντική γιατί αντιπροσωπεύει μια σημαντική πρόοδο στον τομέα του NLP (Natural Language Processing). Τα προηγούμενα μοντέλα επεξεργασίας γλώσσας ήταν περιορισμένα στην ικανότητά τους να κατανοούν και να δημιουργούν φυσική γλώσσα και απαιτούσαν σημαντικές ποσότητες ανθρώπινης συμβολής και εκπαίδευσης για την εκτέλεση των καθηκόντων τους. Αντίθετα το GPT έχει προ εκπαιδευτεί σε τεράστιους όγκους δεδομένων κειμένου, επιτρέποντάς του να κατανοεί και να δημιουργεί γλώσσα με ελάχιστη ανθρώπινη συμμετοχή. Το GPT έχει μια χαρακτηριστική ικανότητα δημιουργίας κειμένου από μια δεδομένη προτροπή. Αυτό επιτυγχάνεται με μια τεχνική γνωστή ως μοντελοποίηση

γλώσσας. Το μοντέλο εκπαιδεύεται να προβλέπει την πιθανότητα μιας συγκεκριμένης λέξης ή φράσης με βάση τις προηγούμενες λέξεις του κειμένου. Αυτό επιτρέπει στο GPT να δημιουργεί κείμενο υψηλής ποιότητας που μοιάζει με άνθρωπο το οποίο δεν διακρίνεται από το κείμενο που γράφτηκε από ανθρώπους. Το GPT έχει επίσης σχεδιαστεί για να είναι ένα ευέλικτο και προσαρμόσιμο μοντέλο επεξεργασίας γλώσσας. Η προσαρμογή του GPT για συγκεκριμένες εργασίες του επιτρέπει να εκτελεί ένα ποικίλο σύνολο εργασιών επεξεργασίας γλώσσας με εξαιρετική ακρίβεια. Για παράδειγμα το GPT μπορεί να ρυθμιστεί με ακρίβεια για να εκτελεί ανάλυση συναισθήματος σε δεδομένα μέσω κοινωνικής δικτύωσης ή για τη δημιουργία περιγραφών προϊόντων υψηλής ποιότητας για ιστότοπους ηλεκτρονικού εμπορίου. Αυτήν τη στιγμή υπάρχουν πολλές εκδόσεις του μοντέλου GPT, καθεμία από τις οποίες βασίζεται στην επιτυχία του προκατόχου της και εισάγει νέες δυνατότητες και βελτιώσεις.

8.2.1 Εφαρμογές βασισμένες στο Chat-GPT

8.2.2 Copy A.I.

Το Copy.ai είναι ένα εργαλείο επεξεργασίας γλώσσας που χρησιμοποιεί το GPT-3 για τη δημιουργία κειμένου φυσικής γλώσσας για διάφορους σκοπούς, όπως τη σύνταξη αντιγράφων διαφημίσεων, αναρτήσεις μέσω κοινωνικής δικτύωσης και περιγραφές προϊόντων. Είναι σχεδιασμένο ώστε να είναι φιλικό προς το χρήστη, να απαιτεί ελάχιστη έως καθόλου τεχνική εξειδίκευση και παρέχει στους χρήστες μια απλή διεπαφή για τη δημιουργία κειμένου που ταιριάζει με το επιθυμητό στυλ, τόνο και περιεχόμενο. Το Copy.ai λειτουργεί παρέχοντας στους χρήστες μια προτροπή κειμένου, η οποία μπορεί να είναι οτιδήποτε, από μια λέξη έως μια μεγαλύτερη περιγραφή του τι θέλουν να δημιουργήσουν. Στη συνέχεια το μοντέλο χρησιμοποιεί το GPT-3 για να δημιουργήσει μια σειρά επιλογών κειμένου, οι οποίες μπορούν να επεξεργαστούν και να βελτιωθούν από τον χρήστη για να διασφαλιστεί ότι ταιριάζουν με την προβλεπόμενη έξοδο. Τα πλεονεκτήματα της χρήσης του copy.ai περιλαμβάνουν την ευκολία χρήσης του, την ικανότητά του να δημιουργεί υψηλής ποιότητας συνεκτικό κείμενο που ταιριάζει με τον επιθυμητό τόνο και στυλ και την ικανότητά του να εξοικονομεί χρόνο και πόρους αυτοματοποιώντας τη διαδικασία σύνταξης κειμένου. Επιπλέον η χρήση του GPT-3 διασφαλίζει ότι το κείμενο που δημιουργείται είναι υψηλής ποιότητας, με υψηλό βαθμό συνοχής και ευαισθησίας περιβάλλοντος. Ωστόσο υπάρχουν επίσης ορισμένα πιθανά μειονεκτήματα στη χρήση του copy.ai. Μία από τις κύριες ανησυχίες είναι η πιθανότητα μεροληψίας στο κείμενο που δημιουργείται από το μοντέλο, καθώς το GPT-3 έχει αποδειχθεί ότι παρουσιάζει προκατάληψη με βάση τα δεδομένα στα οποία εκπαιδεύτηκε. Μια άλλη ανησυχία είναι η πιθανότητα λογοκλοπής, καθώς το κείμενο που δημιουργείται από το copy.ai θα μπορούσε να θεωρηθεί ως αντιγραφή ή επαναχρησιμοποίηση υπάρχοντος περιεχομένου. Επιπλέον, όπως συμβαίνει με κάθε αυτοματοποιημένο εργαλείο δημιουργίας κειμένου, υπάρχει κίνδυνος η έξοδος να μην ταιριάζει πάντα με τις προσδοκίες ή τις απαιτήσεις του χρήστη και μπορεί να απαιτεί περαιτέρω επεξεργασία και βελτίωση.

8.2.3 OpenAI GPT-3 Playground

Το GPT-3 Playground είναι ένα διαδικτυακό εργαλείο που επιτρέπει στους χρήστες να αλληλεπιδρούν με το μοντέλο γλώσσας GPT-3 σε μια φιλική προς το χρήστη διεπαφή. Παρέχει μια σειρά λειτουργιών, συμπεριλαμβανομένης της δυνατότητας εισαγωγής μηνυμάτων κειμένου και λήψης κειμένου που δημιουργείται ως έξοδος, καθώς και δυνατότητα αναπαραγωγής παιχνιδιών και δημιουργίας chatbot χρησιμοποιώντας το GPT-3 API. Το Playground χρησιμοποιεί το μοντέλο γλώσσας GPT-3, ένα μοντέλο αιχμής σε βάθος εκμάθησης που χρησιμοποιείται για την επεξεργασία φυσικής γλώσσας. Το GPT-3 εκπαιδεύεται σε ένα τεράστιο σύνολο δεδομένων διαφορετικών πηγών κειμένου επιτρέποντάς του να παράγει ανθρώπινες αποκρίσεις σε ένα ευρύ φάσμα εισόδων. Το Playground παρέχει μια φιλική προς το χρήστη διεπαφή για αλληλεπίδραση με το GPT-3, επιτρέποντας στους χρήστες να δοκιμάσουν εύκολα τις δυνατότητες του μοντέλου και να δημιουργήσουν εξόδους κειμένου. Ένα πλεονέκτημα του GPT-3 Playground είναι η προσβασιμότητά του. Το εργαλείο είναι διαθέσιμο online, πράγμα που σημαίνει ότι οι χρήστες δεν χρειάζεται να εγκαταστήσουν κάποιο λογισμικό ή να διαθέτουν εξειδικευμένο υλικό για να το χρησιμοποιήσουν. Αυτό διευκολύνει τους μη τεχνικούς χρήστες να πειραματιστούν με το μοντέλο GPT-3 και να εξερευνήσουν τις δυνατότητές του. Ένα άλλο πλεονέκτημα της Παιδικής Χαράς είναι η ευελιξία της. Παρέχει μια σειρά λειτουργιών και ρυθμίσεων που οι χρήστες μπορούν να προσαρμόσουν για να βελτιστοποιήσουν την έξοδο του μοντέλου GPT-3. Για παράδειγμα οι χρήστες μπορούν να προσαρμόσουν τη ρύθμιση "θερμοκρασία" για να ελέγξουν το επίπεδο τυχαίας στο κείμενο που δημιουργείται ή μπορούν να εισάγουν προσαρμοσμένες προτροπές για να προσαρμόσουν την έξοδο στις συγκεκριμένες ανάγκες τους. Ωστόσο υπάρχουν επίσης ορισμένα μειονεκτήματα στην παιδική χαρά GPT-3. Ένα πιθανό ζήτημα είναι το περιορισμένο εύρος του εργαλείου. Παρόλο που παρέχει μια σειρά λειτουργιών για αλληλεπίδραση με το GPT-3 εξακολουθεί να είναι μια απλοποιημένη διεπαφή και ενδέχεται να μην παρέχει το ίδιο επίπεδο προσαρμογής και ελέγχου με άλλα εργαλεία που έχουν σχεδιαστεί ειδικά για εφαρμογές βαθιάς εκμάθησης. Ένα άλλο πιθανό μειονέκτημα είναι το κόστος. Ενώ η ίδια η Playground είναι δωρεάν για χρήση, η πρόσβαση στο GPT-3 API για πιο προηγμένες εφαρμογές απαιτεί συνδρομή επί πληρωμή. Αυτό μπορεί να το κάνει λιγότερο προσβάσιμο για ορισμένους χρήστες που δεν έχουν τους οικονομικούς πόρους να πληρώσουν για το API.

8.2.4 AI Dungeon

Το AI Dungeon είναι ένα διαδικτυακό παιχνίδι όπου οι παίκτες συμμετέχουν σε μια περιπέτεια βασισμένη σε κείμενο και το μοντέλο γλώσσας GPT-3 δημιουργεί απαντήσεις για να παρέχει μια διαδραστική εμπειρία αφήγησης. Πληκτρολογώντας εντολές και περιγραφές οι παίκτες μπορούν να εξερευνήσουν και να αλληλοεπιδράσουν με έναν εικονικό κόσμο, με το μοντέλο GPT-3 να ερμηνεύει την είσοδο για να δημιουργήσει μια απάντηση. Ο κόσμος

και η ιστορία του παιχνιδιού εξελίσσονται σε πραγματικό χρόνο με βάση τις επιλογές και τις ενέργειες του παίκτη, παρέχοντας μια άκρως καθηλωτική και συναρπαστική εμπειρία. Το AI Dungeon χρησιμοποιεί το μοντέλο γλώσσας GPT-3 για να δημιουργήσει απαντήσεις στην είσοδο του παίκτη. Το μοντέλο υποβάλλεται εκ των προτέρων σε μια τεράστια ποσότητα δεδομένων κειμένου και μπορεί να δημιουργήσει κείμενο εξαιρετικά συνεκτικό και σχετικό με τα συμφραζόμενα. Όταν ένας παίκτης εισάγει μια εντολή ή μια περιγραφή το μοντέλο GPT-3 δημιουργεί μια απόκριση με βάση την είσοδο και το περιβάλλον του κόσμου του παιχνιδιού. Η απάντηση μπορεί να είναι οτιδήποτε, από μια απλή περιγραφή του περιβάλλοντος έως μια περίπλοκη αφήγηση που οδηγεί την ιστορία προς τα εμπρός. Τα πλεονεκτήματα της χρήσης του AI Dungeon με GPT-3 είναι πολλά. Το μοντέλο GPT-3 επιτρέπει άκρως καθηλωτικό και συναρπαστικό παιχνίδι, καθώς είναι ικανό να παράγει εξαιρετικά συνεκτικές και σχετικές με τα συμφραζόμενα απαντήσεις. Το παιχνίδι μπορεί να δημιουργήσει έναν πρακτικά άπειρο αριθμό ιστοριών και σεναρίων, παρέχοντας στους παίκτες μια εμπειρία με μεγάλη δυνατότητα αναπαραγωγής. Επιπλέον, το AI Dungeon μπορεί να παιχτεί σε διάφορα είδη από φαντασία έως sci-fi, και μπορεί να καλύψει ένα ευρύ φάσμα προτιμήσεων των παικτών. Ωστόσο, υπάρχουν επίσης ορισμένα μειονεκτήματα στη χρήση του AI Dungeon με το GPT-3.

8.3 Στα ενδότερα του DALL-E

Το OpenAI δημιούργησε ένα μοντέλο δημιουργίας εικόνων που ονομάζεται DALL-E, το οποίο μπορεί να παράγει εικόνες υψηλής ποιότητας με βάση περιγραφές κειμένου. Λειτουργεί κωδικοποιώντας πρώτα την είσοδο κειμένου σε μια αριθμητική αναπαράσταση χρησιμοποιώντας ένα μοντέλο που βασίζεται σε μετασχηματιστή, παρόμοιο με το GPT. Αυτή η αριθμητική αναπαράσταση χρησιμοποιείται στη συνέχεια για τη δημιουργία ενός 2D πλέγματος από "tokens", το καθένα αντιπροσωπεύει μια μικρή ενημέρωση κώδικα εικόνας. Στη συνέχεια, το DALL-E δημιουργεί αυτές τις ενημερώσεις κώδικα εικόνας, μία κάθε φορά, τροφοδοτώντας την αριθμητική αναπαράσταση και τις προηγούμενες ενημερώσεις κώδικα μέσω ενός δικτύου γεννήτριας. Αυτό το δίκτυο γεννήτριας χρησιμοποιεί μια σειρά από συνελκτικά στρώματα για τη δημιουργία των εικόνων, δημιουργώντας σταδιακά τις λεπτομέρειες και την πολυπλοκότητα της εικόνας καθώς προστίθενται περισσότερες ενημερώσεις κώδικα. Η γεννήτρια εκπαιδεύεται χρησιμοποιώντας μια παραλλαγή της αρχιτεκτονικής Generative Adversarial Network (GAN) η οποία φέρνει αντιμέτωπους ένα δίκτυο γεννήτριας ενάντια σε ένα δίκτυο διακριτικού που είναι εκπαιδευμένο να διακρίνει μεταξύ πραγματικών και ψεύτικων εικόνων. Το DALL-E δημιουργεί εικόνες ακολουθώντας μια διαδικασία βήμα προς βήμα αφού τροφοδοτηθεί με μια περιγραφή κειμένου της επιθυμητής εικόνας, όπως "μια πολυθρόνα σε σχήμα αβοκάντο". Το μοντέλο δημιουργεί κάθε patch εικόνας με τη σειρά η οποία σταδιακά δημιουργείται μέχρι την τελική εικόνα. Η τελική εικόνα είναι μια σύνθεση αυτών των επιδιορθώσεων εικόνας και έχει σχεδιαστεί για να ταιριάζει με την περιγραφή κειμένου που παρέχεται στο μοντέλο. Το χαρακτηριστικό γνώρισμα του DALL-E είναι η ικανότητά του να δημιουργεί ευφάνταστες και καινοτόμες εικόνες που εκτείνονται πέρα από την κυριολεκτική σημασία της εισαγωγής κειμένου. Αυτό

γίνεται εφικτό με τη χρήση ενός «λανθάνοντος χώρου» ο οποίος είναι ένας συνεχής διανυσματικός χώρος που περιέχει όλες τις πιθανές αναπαραστάσεις εικόνων. Προσαρμόζοντας τις τιμές σε αυτόν τον λανθάνοντα χώρο το DALL-E μπορεί να δημιουργήσει μια τεράστια γκάμα διαφορετικών και καινοτόμων εικόνων που υπερβαίνουν τη σαφή εισαγωγή κειμένου.

Κεφάλαιο 9. Συμπεράσματα

Οι μελλοντικές κατευθύνσεις του GPT και του DALL-E είναι συναρπαστικές και έχουν τη δυνατότητα για μεταμορφωτικές προόδους στην κατανόηση της γλώσσας και τη δημιουργία εικόνων. Καθώς ο τομέας της τεχνητής νοημοσύνης συνεχίζει να εξελίσσεται, αυτά τα μοντέλα αναμένεται να διαδραματίζουν ολοένα και πιο σημαντικό ρόλο σε ένα ευρύ φάσμα εφαρμογών από chatbot και προσωπικούς βοηθούς έως την ανάπτυξη βιντεοπαιχνιδιών και τη δημιουργία περιεχομένου. Ένας βασικός τομέας εστίασης για το GPT είναι πιθανό να είναι η ανάπτυξη πιο εξελιγμένων μοντέλων γλώσσας. Η πιο πρόσφατη επανάληψη του GPT, το GPT-3, περιέχει εκπληκτικά 175 δισεκατομμύρια παραμέτρους, επιτρέποντάς του να δημιουργεί εξαιρετικά ρεαλιστικό κείμενο που μοιάζει με άνθρωπο. Ωστόσο υπάρχουν ακόμη πολλά περιθώρια βελτίωσης ιδιαίτερα όσον αφορά την κατανόηση του πλαισίου και των αποχρώσεων. Στο μέλλον μπορούμε να περιμένουμε να δούμε προσπάθειες για τη βελτίωση του GPT για συγκεκριμένες περιπτώσεις χρήσης, όπως η ιατρική διάγνωση ή η εξυπηρέτηση πελατών, καθώς και η ανάπτυξη μεγαλύτερων και πιο περίπλοκων μοντέλων που μπορούν να επιτύχουν ακόμη υψηλότερα επίπεδα κατανόησης της γλώσσας. Ένας άλλος σημαντικός τομέας ανάπτυξης για το GPT είναι η ικανότητα συμμετοχής σε φυσική και άπταιστη συνομιλία. Ενώ το GPT είναι ικανό να δημιουργεί εξαιρετικά ρεαλιστικό κείμενο, εξακολουθεί να αγωνίζεται να διατηρήσει τη συνοχή και τη συνοχή σε εκτεταμένες συνομιλίες. Στο μέλλον οι ερευνητές και οι προγραμματιστές πιθανότατα θα επικεντρωθούν στην αντιμετώπιση αυτού του περιορισμού, ενδεχομένως μέσω της χρήσης ενισχυτικής μάθησης ή άλλων τεχνικών. Αυτό θα μπορούσε να οδηγήσει στη δημιουργία chatbots AI και εικονικών βοηθών που δεν διακρίνονται από τους αντίστοιχους ανθρώπους, με ικανότητα κατανόησης και απάντησης σε σύνθετα ερωτήματα και εκτέλεσης εργασιών με ελάχιστη συνεισφορά χρήστη. Όσο για το DALL-E το μέλλον είναι πιθανό να περιλαμβάνει την επέκταση του εύρους των εξόδων του πέρα από τις εικόνες για να συμπεριλάβει άλλες μορφές μέσων, όπως βίντεο ή ήχο. Αυτό θα μπορούσε να έχει σημαντικές επιπτώσεις σε τομείς όπως η παραγωγή ταινιών και τηλεόρασης, όπου η ικανότητα δημιουργίας εξαιρετικά ρεαλιστικών και πολύπλοκων οπτικών και ήχων είναι ζωτικής σημασίας. Επιπλέον μπορεί να γίνουν προσπάθειες για τη βελτίωση της ταχύτητας και της αποτελεσματικότητας του DALL-E, ενδεχομένως μέσω της χρήσης ισχυρότερου υλικού ή νέων τεχνικών εκπαίδευσης. Ένας άλλος τομέας ανάπτυξης για το DALL-E είναι η ικανότητα δημιουργίας πιο εξελιγμένων και κομψών εικόνων. Στο μέλλον μπορούμε να αναμένουμε να δούμε προσπάθειες για την αντιμετώπιση αυτών των περιορισμών, ενδεχομένως μέσω της χρήσης πιο προηγμένων αλγορίθμων ή της ενσωμάτωσης πρόσθετων πηγών δεδομένων. Συνολικά το μέλλον του GPT και του DALL-E είναι γεμάτο από δυνατότητες και δυνατότητες για πρωτοποριακές εξελίξεις στην τεχνητή νοημοσύνη. Αν και υπάρχουν σίγουρα προκλήσεις και περιορισμοί που πρέπει να αντιμετωπιστούν η συνεχιζόμενη έρευνα και ανάπτυξη σε αυτούς τους τομείς είναι βέβαιο ότι θα αποφέρει συναρπαστικά και μεταμορφωτικά αποτελέσματα τα επόμενα χρόνια.

Βιβλιογραφία

- Alfano, M., Abedin, E., Reimann, R., Ferreira, M., & Cheong, M. (2024). Now you see me, now you don't: An exploration of religious exnomination in DALL-E. *Ethics and Information Technology*, 26(2), 27. <https://doi.org/10.1007/s10676-024-09760-y>
- Ayodele, T. O. (n.d.). Machine Learning Overview. *New Advances in Machine Learning*.
- Becker, B. A., Denny, P., Finnie-Ansley, J., Luxton-Reilly, A., Prather, J., & Santos, E. A. (2023). Programming Is Hard - Or at Least It Used to Be: Educational Opportunities and Challenges of AI Code Generation. *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*, 500–506. <https://doi.org/10.1145/3545945.3569759>
- Bengio, Y. (2009). Learning Deep Architectures for AI. *Foundations and Trends® in Machine Learning*, 2(1), 1–127. <https://doi.org/10.1561/22000000006>
- Bengio, Y., Lecun, Y., & Hinton, G. (2021). Deep learning for AI. *Communications of the ACM*, 64(7), 58–65. <https://doi.org/10.1145/3448250>
- Boonstra, L. (2021). *The Definitive Guide to Conversational AI with Dialogflow and Google Cloud: Build Advanced Enterprise Chatbots, Voice, and Telephony Agents on Google Cloud* (p. 408). Scopus. <https://doi.org/10.1007/978-1-4842-7014-1>
- Boza, P., & Evgeniou, T. (2021). Implementing Ai Principles: Frameworks, Processes, and Tools. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3783124>
- Cascella, M., Semeraro, F., Montomoli, J., Bellini, V., Piazza, O., & Bignami, E. (2024). The Breakthrough of Large Language Models Release for Medical Applications: 1-Year Timeline and Perspectives. *Journal of Medical Systems*, 48(1). Scopus. <https://doi.org/10.1007/s10916-024-02045-3>
- Clune, J. (2020). AI-GAs: AI-generating algorithms, an alternate paradigm for producing general artificial intelligence (arXiv:1905.10985). arXiv. <http://arxiv.org/abs/1905.10985>
- Dunder, N., Lundborg, S., Wong, J., & Viberg, O. (2024). Kattis vs ChatGPT: Assessment and Evaluation of Programming Tasks in the Age of Artificial Intelligence. *Proceedings of the 14th Learning Analytics and Knowledge Conference*, 821–827. <https://doi.org/10.1145/3636555.3636882>
- Helm, J. M., Swiergosz, A. M., Haeberle, H. S., Karnuta, J. M., Schaffer, J. L., Krebs, V. E., Spitzer, A. I., & Ramkumar, P. N. (2020). Machine Learning and Artificial Intelligence: Definitions, Applications, and Future Directions. *Current Reviews in Musculoskeletal Medicine*, 13(1), 69–76. <https://doi.org/10.1007/s12178-020-09600-8>
- Hutter, F., Kotthoff, L., & Vanschoren, J. (Eds.). (2019). *Automated Machine Learning: Methods, Systems, Challenges*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-05318-5>
- Jo, H., Lee, J.-K., Lee, Y.-C., & Choo, S. (2024). Generative artificial intelligence and building design: Early photorealistic render visualization of façades using local identity-trained models. *Journal of Computational Design and Engineering*, 11(2), 85–105. Scopus. <https://doi.org/10.1093/jcde/qwae017>
- Li, C., Luo, Y., Xie, Y., Zhang, Z., Liu, Y., Zou, L., & Xiao, F. (2024). Structural and functional prediction, evaluation, and validation in the post-sequencing era. *Computational and Structural Biotechnology Journal*, 23, 446–451. <https://doi.org/10.1016/j.csbj.2023.12.031>

- Nasteski, V. (2017). An overview of the supervised machine learning methods. *HORIZONS.B*, 4, 51–62. <https://doi.org/10.20544/HORIZONS.B.04.1.17.P05>
- Prasad, P., & Sane, A. (2024). *A Self-Regulated Learning Framework using Generative AI and its Application in CS Educational Intervention Design*. 1, 1070–1076. Scopus. <https://doi.org/10.1145/3626252.3630828>
- Rao, N., Tsay, J., Kate, K., Hellendoorn, V., & Hirzel, M. (2024). AI for Low-Code for AI. *Proceedings of the 29th International Conference on Intelligent User Interfaces*, 837–852. <https://doi.org/10.1145/3640543.3645203>
- Salvaris, M., Dean, D., & Tok, W. H. (2018). *Deep learning with azure: Building and deploying artificial intelligence solutions on the microsoft AI platform* (p. 277). Scopus. <https://doi.org/10.1007/978-1-4842-3679-6>
- Sheard, J., Denny, P., Hellas, A., Leinonen, J., Malmi, L., & Simon. (2024). Instructor Perceptions of AI Code Generation Tools—A Multi-Institutional Interview Study. *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, 1223–1229. <https://doi.org/10.1145/3626252.3630880>
- Wang, M., Tai, C., Zhang, Q., Yang, Z., Li, J., Shen, K., & Wang, K. (2021). Application of BigML in the Classification Evaluation of Top Coal Caving. *Shock and Vibration*, 2021(1), 8552247. <https://doi.org/10.1155/2021/8552247>
- Xin, Y., Kong, L., Liu, Z., Chen, Y., Li, Y., Zhu, H., Gao, M., Hou, H., & Wang, C. (2018). Machine Learning and Deep Learning Methods for Cybersecurity. *IEEE Access*, 6, 35365–35381. <https://doi.org/10.1109/ACCESS.2018.2836950>