

Ενότητα 7

ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

Τεχνητή Νοημοσύνη

Εκπαιδεύοντας ένα νευρωνικό
δίκτυο

Ενότητα 7. Τεχνητή Νοημοσύνη

7.1 Εισαγωγή

Τα τελευταία χρόνια, η Τεχνητή Νοημοσύνη (ΤΝ) έχει σημειώσει ραγδαία εξέλιξη, επηρεάζοντας πολλούς τομείς της επιστήμης, της τεχνολογίας και της κοινωνίας. Εφαρμογές, όπως οι έξυπνοι βοηθοί (π.χ. Siri, Alexa), τα συστήματα αυτόματης εξυπηρέτησης πελατών και τα εργαλεία ανάλυσης δεδομένων, βασίζονται σε αυτές τις τεχνολογίες για την παροχή πιο εξελιγμένων υπηρεσιών. Ο κλάδος της Τεχνητής Νοημοσύνης που έχει γνωρίσει αυτήν την ανάπτυξη είναι αυτός της Μηχανικής Μάθησης (machine learning). Η ραγδαία ανάπτυξη της Μηχανικής Μάθησης, τα τελευταία χρόνια, συνδέεται άμεσα με την εκρηκτική αύξηση των μεγάλων δεδομένων (big data), την ανάπτυξη του Διαδικτύου και τη συνεχή βελτίωση της υπολογιστικής ισχύος. Η τεράστια παραγωγή και διαθεσιμότητα δεδομένων από το Διαδίκτυο, τα Μέσα Κοινωνικής Δικτύωσης, τις συσκευές του Διαδικτύου των Πραγμάτων και άλλες ψηφιακές πηγές έχει δημιουργήσει μια πλούσια πηγή για την εκπαίδευση των μοντέλων Μηχανικής Μάθησης. Η αυξημένη υπολογιστική ισχύς, με την εξέλιξη των εξειδικευμένων επεξεργαστών για τεχνητή νοημοσύνη, έχει επιτρέψει την επεξεργασία αυτών των τεράστιων όγκων δεδομένων σε χρόνο ρεκόρ, καθιστώντας δυνατή την εκπαίδευση πολύπλοκων αλγορίθμων βαθιάς μάθησης. Όλα αυτά τα στοιχεία έχουν οδηγήσει σε σημαντικές προόδους σε διάφορους τομείς, όπως η αναγνώριση εικόνας, η επεξεργασία φυσικής γλώσσας και η πρόβλεψη τάσεων, ανοίγοντας νέες προοπτικές για το μέλλον. Στην ενότητα αυτή θα μάθουμε τη βασική διαφορά μεταξύ Τεχνητής Νοημοσύνης και Μηχανικής Μάθησης και θα έχουμε την ευκαιρία να εκπαιδεύσουμε το δικό μας νευρωνικό δίκτυο στην αναγνώριση φωτογραφιών που έχουν ληφθεί από τον Διεθνή Διαστημικό Σταθμό.



Η βασική διαφορά μεταξύ τεχνητής νοημοσύνης (Artificial Intelligence - AI) και μηχανικής μάθησης (Machine Learning - ML) είναι η εξής:

- **Τεχνητή Νοημοσύνη (AI):** Η AI είναι ένας ευρύτερος τομέας που περιλαμβάνει την ανάπτυξη συστημάτων και αλγορίθμων που μπορούν να εκτελούν καθήκοντα που απαιτούν νοημοσύνη. Αυτά τα καθήκοντα μπορεί να περιλαμβάνουν την επίλυση προβλημάτων, την κατανόηση φυσικής γλώσσας, την όραση υπολογιστών και άλλα. Στόχος της AI είναι η δημιουργία συστημάτων που μπορούν να μιμούνται ή να αναπαράγουν ανθρώπινη νοημοσύνη.
- **Μηχανική Μάθηση (ML):** Η ML είναι ένας υποτομέας της AI που επικεντρώνεται στην ανάπτυξη αλγορίθμων που επιτρέπουν στα συστήματα να "μαθαίνουν" από δεδομένα και να βελτιώνουν την απόδοσή τους με την πάροδο του χρόνου χωρίς να είναι ρητά προγραμματισμένα γι' αυτό. Στη ML, τα μοντέλα "εκπαιδεύονται" χρησιμοποιώντας μεγάλα σύνολα δεδομένων για να προβλέπουν ή να παίρνουν αποφάσεις με βάση νέα δεδομένα.

Με λίγα λόγια, η τεχνητή νοημοσύνη είναι η ομπρέλα που περιλαμβάνει διάφορες τεχνικές και προσεγγίσεις, με τη μηχανική μάθηση να είναι μία από αυτές.

Εικόνα 7.1. Απάντηση του ChatGPT στην ερώτηση «Ποια είναι η βασική διαφορά Τεχνητής Νοημοσύνης και Μηχανικής Μάθησης»

7.2 Τεχνητή Νοημοσύνη

Τα τελευταία χρόνια ακούμε ή διαβάζουμε όλο και πιο συχνά για τον όρο **Τεχνητή Νοημοσύνη (ΤΝ)**, στα αγγλικά (**Artificial Intelligence** ή **AI**). Τελευταία, συναντούμε πολλές αναφορές για την Τεχνητή Νοημοσύνη στα ηλεκτρονικά μέσα επικοινωνίας, όπως στην τηλεόραση και το ραδιόφωνο, σε άρθρα εφημερίδων και περιοδικών ή στις ιστοσελίδες του Παγκόσμιου Ιστού. Μάλιστα, πολλές φορές το περιεχόμενό τους διεγείρει τη φαντασία μας σε μεγάλο βαθμό, γιατί συνδέεται με εντυπωσιακά σενάρια. Αυτοκίνητα, τρένα ή συρμοί του μετρό που κινούνται χωρίς οδηγό (αυτόνομα οχήματα), δημιουργία κειμένου, λήψη απόφασης, αναγνώριση ομιλίας και διάγνωση ασθενειών, είναι κάποια από αυτά.

Η ενότητα αυτή θα απαντήσει σε μια σειρά εύλογων ερωτημάτων που γεννώνται. Τι είναι και πώς λειτουργεί η Τεχνητή Νοημοσύνη; Ποια είναι τα πλεονεκτήματα και τα μειονεκτήματά της; Σε ποιους τομείς εφαρμόζεται και πώς θα εφαρμοστεί στο μέλλον; Πώς μπορούμε να την αξιοποιήσουμε στην εκπαίδευση; Σκοπός δεν είναι να αναλυθεί σε βάθος το θέμα, αλλά να μας εισαγάγει στην έννοια της Τεχνητής Νοημοσύνης, με τρόπο σύγνοτομο και κατανοητό.

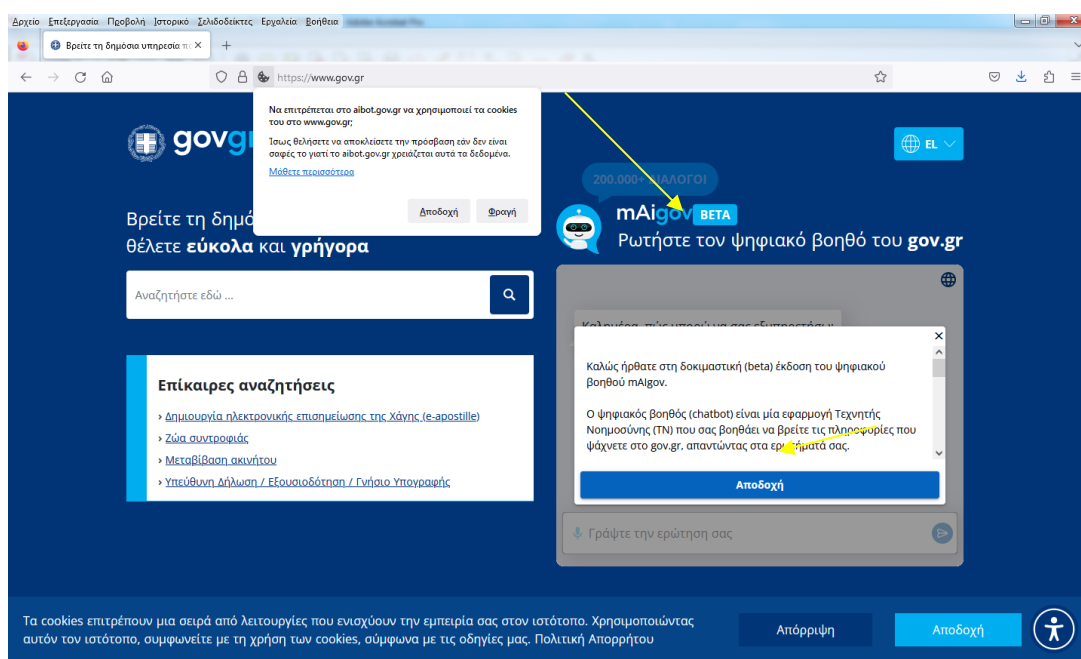
Η Τεχνητή Νοημοσύνη είναι κλάδος της επιστήμης των υπολογιστών, που σκοπό έχει τη δημιουργία συστημάτων τα οποία μπορούν να μιμηθούν την ανθρώπινη ευφυΐα. Συστήματα που μπορούν να προσεγγίζουν όλο και πιο πιστά την ανθρώπινη σκέψη και να παίρνουν αποφάσεις, να επιλύουν προβλήματα, να κάνουν προβλέψεις, να αναγνωρίζουν την ομιλία ή το πρόσωπο, να δημιουργούν περιεχόμενο, να δίνουν συμβουλές ή να παρέχουν βοήθεια και πολλά άλλα.

Δε θα ήταν υπερβολή να υποστηρίξουμε ότι στοιχεία Τεχνητής Νοημοσύνης συναντάμε καθημερινά στη ζωή μας. Ας δούμε μερικά χαρακτηριστικά παραδείγματα:

- 1) Η ανίχνευση και η αναγνώριση του προσώπου μας ή του δακτυλικού μας αποτυπώματος για να ξεκλειδώσει το κινητό μας τηλέφωνο.
- 2) Η πρόβλεψη του καιρού.
- 3) Η συμπεριφορά μας στα μέσα κοινωνικής δικτύωσης ή η πλοήγησή μας στο Διαδίκτυο από το κινητό μας τηλέφωνο γίνεται «κατανοητή» λόγω της ενσωματωμένης Τεχνητής Νοημοσύνης και εμφανίζεται περιεχόμενο, αλλά και διαφημίσεις, σχετικό με τα ενδιαφέροντά μας.
- 4) Η αναγνώριση της συντομότερης διαδρομής ή της κυκλοφοριακής συμφόρησης ή η πλοήγηση ενός οχήματος προς έναν προορισμό, μέσω κινητού τηλεφώνου, είναι εφαρμογές που χρησιμοποιούν Τεχνητή Νοημοσύνη.
- 5) Οι εφαρμογές που υποστηρίζουν την αλληλεπίδραση χρήστη και μηχανής μέσω μιας τεχνικής με ερωτήσεις και απαντήσεις. Με τις εφαρμογές αυτές, αναπτύσσεται ένας διάλογος (συχνά αναφέρεται ως **chatbot** από τη σύντμηση της λέξης **chat** και της κατάληξης της λέξης **robot**), μια συνομιλία, όπου ο άνθρωπος αναζητά πληροφορίες από τη μηχανή, όπως για παράδειγμα με την πληκτρολόγηση ενός κειμένου στο σχετικό πεδίο. Με άλλα λόγια, πληκτρολογούμε μια **προτροπή (prompt)** για το θέμα αναζήτησης που μας ενδιαφέρει και η μηχανή εμφανίζει την απάντησή της, πάλι με κείμενο στη φυσική γλώσσα του ανθρώπου ή άλλης μορφής περιεχόμενο (για παράδειγμα μια εικόνα). Λόγω της παραγωγής περιεχομένου με δημιουργικό τρόπο, όπως θα έκανε και ένας άνθρωπος, χαρακτηρίζεται ως **παραγωγική Τεχνητή Νοημοσύνη (generative AI)**. Η λειτουργία της είναι η εξής: στη μηχανή έχει εγκατασταθεί εφαρμογή (λογισμικό), η οποία διαβάσει την προτροπή (ερώτηση ή οδηγία) και, αφού την επεξεργαστεί, εμφανίζει την απάντηση στην

οθόνη. Παραδείγματα τέτοιων εφαρμογών, όπου χρησιμοποιούμε διαλόγους για να αντλήσουμε τις πληροφορίες που μας ενδιαφέρουν, είναι ο ψηφιακός βοηθός που έχει ενσωματωθεί στην ιστοσελίδα **gov.gr**, καθώς και αυτές που λέγονται **Gemini**, **Copilot** και **ChatGPT**. Στη συνέχεια, θα περιγράψουμε συνοπτικά πώς μπορούμε να χρησιμοποιήσουμε αυτές τις εφαρμογές, αναφέροντας και χρήσιμα παραδείγματα στην καθημερινή μας ζωή.

Η εφαρμογή Gemini αναπτύχθηκε από την εταιρεία Google. Πληκτρολογούμε την οδηγία/προτροπή (prompt) που επιθυμούμε και σε λίγα δευτερόλεπτα θα εμφανιστεί η απάντηση της εφαρμογής. Τονίζεται, ιδιαίτερα, ότι η απάντηση αυτή θα πρέπει να ελέγχεται ως προς την αξιοπιστία της και να μη χρησιμοποιείται άκριτα. Με ανάλογο τρόπο λειτουργούν και οι εφαρμογές **Copilot** και **ChatGPT**, οι οποίες δημιουργήθηκαν από τις εταιρείες Microsoft και OpenAI αντίστοιχα.



Εικόνα 7.2. Ο ψηφιακός βοηθός στην ιστοσελίδα gov.gr

Όπως φαίνεται, χρησιμοποιώντας την παραγωγική Τεχνητή Νοημοσύνη, μπορούμε να αντλήσουμε τις πληροφορίες που θέλουμε, εύκολα και γρήγορα, αλλά θα πρέπει πάντα να είμαστε επιφυλακτικοί για την αξιοπιστία τους. Στη συνέχεια, θα εξηγήσουμε πώς λειτουργούν οι εφαρμογές αυτές και θα καταλάβουμε γιατί οι επιφυλάξεις αυτές είναι βάσιμες.

Οι εφαρμογές Τεχνητής Νοημοσύνης «εκπαιδεύονται» (μαθαίνουν) σε πολύ μεγάλο όγκο δεδομένων (μεγάλα δεδομένα ή big data), με βάση κατάλληλα μοντέλα και αλγόριθμους και με σκοπό να παράγουν αποτελέσματα (πληροφορίες) που θα βοηθήσουν τον άνθρωπο που τις χρησιμοποιεί. Βασίζονται, δηλαδή, στην αρχιτεκτονική του παραγωγικού προ-εκπαιδευμένου μετασχηματιστή (**Generative Pre-trained Transformer** ή **GPT**). Τα δεδομένα μπορεί να είναι, για παράδειγμα, το περιεχόμενο ιστοσελίδων, βιβλίων ή άρθρων που έχουν αναρτηθεί στον Παγκόσμιο Ιστό. Περισσότερα δεδομένα και καλύτερα (με περισσότερη σαφήνεια και ακρίβεια) μοντέλα και αλγόριθμοι οδηγούν σε πιο αξιόπιστα αποτελέσματα.

Μπορείς να μου δώσεις έναν ορισμό για τον όρο "Τεχνητή Νοημοσύνη";



Βεβαίως! Η Τεχνητή Νοημοσύνη (ΤΝ) αναφέρεται στον τομέα της πληροφορικής που ασχολείται με τη δημιουργία συστημάτων ή μηχανών που μπορούν να εκτελέσουν εργασίες οι οποίες συνήθως απαιτούν ανθρώπινη νοημοσύνη. Αυτές οι εργασίες περιλαμβάνουν την εκμάθηση από εμπειρίες, την αναγνώριση προτύπων, την κατανόηση και επεξεργασία φυσικής γλώσσας, τη λήψη αποφάσεων και την επίλυση προβλημάτων.

Οι εφαρμογές της Τεχνητής Νοημοσύνης κυμαίνονται από απλούς αλγορίθμους μέχρι πολύπλοκα συστήματα που μπορούν να αυτο-βελτιώνονται με την πάροδο του χρόνου. Οι βασικές τεχνολογίες που χρησιμοποιούνται περιλαμβάνουν τη μηχανική μάθηση (machine learning), τη βαθιά μάθηση (deep learning), τη νευρωνική δικτύωση και τη φυσική γλώσσα (natural language processing, NLP).

Ο στόχος της Τεχνητής Νοημοσύνης είναι να δημιουργήσει μηχανές που μπορούν να σκέφτονται και να ενεργούν όπως οι άνθρωποι, ή ακόμη και να ξεπεράσουν τις ανθρώπινες ικανότητες σε συγκεκριμένους τομείς.

Εικόνα 7.3. Εμφάνιση αποτελέσματος εφαρμογής ChatGPT για το τι είναι η Τεχνητή Νοημοσύνη



Μπορείς να μου δώσεις έναν ορισμό για τον όρο "Τεχνητή Νοημοσύνη";

Εμφάνιση πρόχειρων ✓



Φυσικά! Μπορώ να σου δώσω έναν ορισμό για την Τεχνητή Νοημοσύνη (ΤΝ):

Τεχνητή Νοημοσύνη είναι ο κλάδος της πληροφορικής που ασχολείται με τη δημιουργία μηχανών που μπορούν να μιμούνται στοιχεία της ανθρώπινης νοημοσύνης, όπως:

- **Μάθηση:** Η ικανότητα να αποκτούν νέες πληροφορίες και να προσαρμόζουν τη συμπεριφορά τους βάσει αυτών.
- **Εξαγωγή συμπερασμάτων:** Η ικανότητα να συνδέουν πληροφορίες και να φτάνουν σε λογικά συμπεράσματα.
- **Ανάγνωση και κατανόηση γλωσσών:** Η ικανότητα να κατανοούν και να παράγουν ανθρώπινη γλώσσα.
- **Επίλυση προβλημάτων:** Η ικανότητα να αναλύουν προβλήματα και να βρίσκουν λύσεις.
- **Αναγνώριση προτύπων:** Η ικανότητα να αναγνωρίζουν μοτίβα σε δεδομένα.

Με απλούστερα λόγια, η ΤΝ αφορά στην ανάπτυξη μηχανών που μπορούν να σκέφτονται, να μαθαίνουν και να εκτελούν εργασίες που συνήθως απαιτούν ανθρώπινη νοημοσύνη.

Εικόνα 7.4. Εμφάνιση αποτελέσματος εφαρμογής Gemini για το τι είναι η ΤΝ

Για παράδειγμα, αν η «εκπαίδευση» γίνει με περισσότερα δεδομένα που αφορούν διάφορες ασθένειες, είναι πιθανότερο να οδηγήσει σε πιο αξιόπιστη διάγνωση για την ασθένεια ενός ατόμου που παρουσιάζει παρόμοια συμπτώματα με εκείνα στα οποία η εφαρμογή εκπαιδεύτηκε. Όμως, η διάγνωση αυτή δε σημαίνει ότι είναι η σωστή. Ο ιατρός θα πρέπει να την εξετάσει ως ένα ενδεχόμενο και να συνεκτιμήσει και άλλα στοιχεία, που θα κρίνει εκείνος ότι χρειάζονται, προκειμένου να αποφανθεί για το ποια είναι η ασθένεια. Επομένως, ο ιατρός θα μπορεί έγκαιρα να καταλήξει σε συμπέρασμα για την κατάσταση του ασθενούς του και, με την κατάλληλη θεραπευτική αγωγή, να επιδιώξει να τον οδηγήσει στη θεραπεία του, αποφεύγοντας τυχόν δυσάρεστα αποτελέσματα ή παρενέργειες.

Μπορούμε να καταλάβουμε ότι, στην περίπτωση της παραγωγικής Τεχνητής Νοημοσύνης, τα μεγάλα δεδομένα που χρησιμοποιούνται για την «εκπαίδευση» της εφαρμογής θα περιλαμβάνουν, μεταξύ άλλων, κείμενα και κώδικα, ενώ όσον αφορά την πλοήγηση οχημάτων σε διάφορους προορισμούς, τα δεδομένα θα περιέχουν οδούς και την κατεύθυνση διέλευσής τους.

Η επιτυχία της εφαρμογής αυτής στηρίζεται σε δύο άξονες. Από τη μία πλευρά, είναι σημαντική η κατανόηση της ερώτησης του ανθρώπου από τη μηχανή και από την άλλη, η δημιουργία απάντησης από τη μηχανή με φυσικό τρόπο που θα κατανοεί ο άνθρωπος. Υπάρχει, όμως, και μία άλλη παράμετρος που πρέπει να έχουμε υπόψη μας. Η απάντηση της μηχανής είναι αξιόπιστη; Αναγνωρίστηκε σωστά η έννοια των λέξεων και το νόημα της ερώτησής μας; Μήπως η ερώτηση είναι γενική και δεν είναι επί της ουσίας; Τα δεδομένα στα οποία στηρίχθηκε η απάντηση είναι επικαιροποιημένα;

Η απάντηση στα ερωτήματα αυτά μας οδηγεί στο συμπέρασμα ότι πρέπει να είμαστε επιφυλακτικοί με τις πληροφορίες που μας δίνονται και να τις αξιολογούμε με κριτική προσέγγιση. Για να περιορίσουμε τις επιφυλάξεις μας, θα μπορούσαμε να συνεχίσουμε τα ερωτήματά μας προς την εφαρμογή, ζητώντας να μας παρουσιάσει επιπλέον στοιχεία στα οποία στηρίζεται η απάντησή της, όπως, για παράδειγμα, βιβλιογραφικές αναφορές που στη συνέχεια μπορούμε να εξετάσουμε ως προς την αξιοπιστία τους. Πρέπει να υπογραμμίσουμε ότι η ποιότητα των απαντήσεων εξαρτάται και από την ποιότητα των ερωτήσεων-οδηγιών που απευθύνουμε προς τη μηχανή. Οι απαντήσεις της μηχανής μπορεί να γίνονται πιο εμπεριστατωμένες, όσο ο διάλογος συνεχίζεται, οι ερωτήσεις μας γίνονται πιο σαφείς και εξειδικεύονται σε μεγαλύτερο βαθμό.

Η εφαρμογή της Τεχνητής Νοημοσύνης, θα πρέπει να χρησιμοποιείται ως ένα βοηθητικό εργαλείο στη μαθησιακή διαδικασία και χρειάζεται πάντα ο ανθρώπινος παράγοντας για την τελική επεξεργασία των αποτελεσμάτων της. Επιπλέον, όπως θα μπορούσε να γίνει αντιληπτό, χρειάζεται μια σχετική εξοικείωση ως προς την τεχνική ανάπτυξης μιας ερώτησης ή ενός νήματος διαλόγου (πολλών διαδοχικών ερωτήσεων) με τη μηχανή, προκειμένου να καταλήξουμε σε αξιόπιστα αποτελέσματα.

Η Τεχνητή Νοημοσύνη εμπλέκεται σε πολλές δραστηριότητες των ανθρώπων στην καθημερινή τους ζωή. Η τεχνολογία των συστημάτων αυτών βελτιώνεται διαρκώς. Τα αποτελέσματα που μας παρουσιάζουν σήμερα είναι συχνά αναξιόπιστα. Για αυτό και ο άνθρωπος εκφράζει ανησυχίες. Από τη μια πλευρά, αναγνωρίζει τη θετική πλευρά της τεχνολογίας που περιέχει λειτουργίες Τεχνητής Νοημοσύνης με εφαρμογές στην καθημερινότητά του και από την άλλη αντιλαμβάνεται ότι είναι πιθανό η τεχνολογία αυτή να του δημιουργήσει προβλήματα. Θα πρέπει να την υιοθετήσει ή όχι; Υπάρχει τρόπος ώστε η εμπλοκή του με αυτή να αποβεί εποικοδομητική στη ζωή του; Πώς θα μπορούσε να εκμεταλλευτεί τα οφέλη της, αποφεύγοντας ταυτόχρονα τα προβλήματα που μπορεί να ανακύψουν; Οι απαντήσεις σε τέτοιου είδους εύλογα ερωτήματα, καθώς και σε θέματα ηθικής και δεοντολογίας ως προς τη χρήση συστημάτων Τεχνητής Νοημοσύνης, θα εξεταστούν στη συνέχεια.

Όπως αναφέρθηκε παραπάνω, οι εφαρμογές Τεχνητής Νοημοσύνης βασίζονται σε μεγάλο όγκο δεδομένων προκειμένου να καταλήξουν σε αποτελέσματα. Τι θα γίνει, όμως, αν τα δεδομένα αυτά δεν είναι αξιόπιστα ή αν περιέχουν στοιχεία που δεν ανταποκρίνονται στην πραγματικότητα; Ας σκεφτούμε, για παράδειγμα, την περίπτωση της παραγωγικής Τεχνητής Νοημοσύνης. Αν τα δεδομένα περιέχουν ανακριβείς ή παραπλανητικές πληροφορίες, η εφαρμογή θα μας εμφανίσει ψευδείς ειδήσεις (fake news) ή θα καταλήξει σε λάθος συμπεράσματα, προβλέψεις (για παράδειγμα, η πρόβλεψη του καιρού) ή αποφάσεις. Επίσης, είναι πιθανό η εφαρμογή Τεχνητής Νοημοσύνης να έχει προβλήματα στην ορθή επεξεργασία των δεδομένων και την παραγωγή αποτελέσματος.

Ποια, λοιπόν, είναι η λύση για την εύρυθμη λειτουργία των συστημάτων αυτών; Η απάντηση στο ερώτημα είναι η κατάλληλη συμπεριφορά όλων των εμπλεκόμενων. Η συμπεριφορά αυτή προκύπτει, όταν όλοι αναγνωρίζουν την ευθύνη που έχουν απέναντι στο κοινωνικό σύνολο. Ο διαχειριστής θα φροντίσει για την αξιοπιστία των δεδομένων, ο δημιουργός για την ενδεδειγμένη λειτουργία της εφαρμογής και ο χρήστης για τον κριτικό έλεγχο των αποτελεσμάτων της μηχανής, έχοντας πάντα επιφυλάξεις ως προς την ορθότητά τους. Και βέβαια, είναι απαραίτητο να υπάρχει συμμόρφωση με τη νομοθεσία που ρυθμίζει τους κανόνες λειτουργίας των συστημάτων Τεχνητής Νοημοσύνης, η οποία θα πρέπει να εμπλουτίζεται με νέο περιεχόμενο, όσο η τεχνολογία εξελίσσεται.

Με τον τρόπο αυτό, η παραγωγική Τεχνητή Νοημοσύνη θα μπορέσει να αποφύγει αποτελέσματα που περιέχουν, για παράδειγμα, προκαταλήψεις, μεροληψία, διακρίσεις, ακατάλληλο περιεχόμενο ή ψευδείς ειδήσεις και να φροντίζει για την προστασία της ιδιωτικότητας, τον σεβασμό των δικαιωμάτων του ανθρώπου και την ασφάλεια των δεδομένων που διαχειρίζεται. Τέλος, να έχουμε υπόψη μας ότι θα πρέπει να αναφέρουμε την πηγή (τον δημιουργό του έργου) για το υλικό που χρησιμοποιούμε, ενσωματώνοντας στις εργασίες μας τη σχετική αναφορά. Διαφορετικά, η εργασία μας θα θεωρηθεί ως προϊόν λογοκλοπής, η οποία, πρέπει να σημειωθεί, μπορεί εύκολα να ανιχνευτεί με το κατάλληλο λογισμικό. Ο δημιουργός κάθε έργου (π.χ. βιβλίου, άρθρου ή λογισμικού) έχει τα πνευματικά δικαιώματα (copyright) για το δημιούργημά του. Η λογοκλοπή παραβιάζει τα δικαιώματα αυτά και για τον λόγο αυτό έχει ποινικές συνέπειες, σύμφωνα με τη νομοθεσία.

7.2.1 Ερωτήσεις

E.1: Έχετε χρησιμοποιήσει την εφαρμογή Gemini για να ζητήσετε πληροφορίες για ένα θέμα εργασίας που σας ανέθεσε ο καθηγητής ή η καθηγήτριά σας; Εξηγήστε πώς θα ελέγξετε την αξιοπιστία των αποτελεσμάτων που παρήγαγε η εφαρμογή.

E.2: Περιγράψτε δύο παραδείγματα από την καθημερινή ζωή, που δεν αναφέρονται στην ενότητα αυτή, στα οποία εμπλέκεται η Τεχνητή Νοημοσύνη.

E.3: Εξηγήστε τι είναι η Τεχνητή Νοημοσύνη και περιγράψτε δύο παραδείγματα όπου η χρήση της θα μπορούσε να εγκυμονεί κινδύνους για τον άνθρωπο.

7.3 Τεχνητή Νοημοσύνη: Ορισμοί



Η **Τεχνητή Νοημοσύνη (TN, Artificial Intelligence, AI)** είναι ένας κλάδος της επιστήμης της Πληροφορικής που ασχολείται με τη δημιουργία συστημάτων ή μηχανών ικανών να εκτελούν εργασίες που απαιτούν ανθρώπινη νοημοσύνη. Αυτές οι εργασίες περιλαμβάνουν την αναγνώριση ομιλίας, την οπτική αναγνώριση αντικειμένων, τη λήψη αποφάσεων, την επίλυση προβλημάτων, την κατανόηση της φυσικής γλώσσας και τη μάθηση από εμπειρία.

Υπάρχουν διάφοροι τύποι τεχνητής νοημοσύνης:

- **Ασθενής Τεχνητή Νοημοσύνη (Narrow AI ή Weak AI):** Αυτή η μορφή Τεχνητής Νοημοσύνης είναι σχεδιασμένη για να εκτελεί συγκεκριμένες εργασίες όπως, για παράδειγμα, η αναγνώριση μοτίβων, αλλά δεν διαθέτει γενική νοημοσύνη. Ωστόσο, αυτές τις εργασίες τις εκτελεί καλύτερα από τον άνθρωπο. Παραδείγματα αυτής της κατηγορίας είναι τα συστήματα αναγνώρισης ομιλίας, όπως το Siri της Apple και τα μεγάλα γλωσσικά μοντέλα όπως το ChatGPT και το Gemini.
- **Γενική Τεχνητή Νοημοσύνη (General AI):** Αυτή η μορφή TN έχει την ικανότητα να κατανοεί, να μαθαίνει και να εφαρμόζει τη γνώση σε ένα ευρύ φάσμα δραστηριοτήτων, προσομοιάζοντας στην ανθρώπινη νοημοσύνη. Αυτή η μορφή Τεχνητής Νοημοσύνης είναι ακόμα υπό ανάπτυξη και αποτελεί αντικείμενο έρευνας.
- **Υπερνοημοσύνη (Superintelligence):** Αυτή η θεωρητική μορφή Τεχνητής Νοημοσύνης υπερβαίνει κατά πολύ την ανθρώπινη νοημοσύνη σε όλους τους τομείς. Αποτελεί θέμα φιλοσοφικής και επιστημονικής συζήτησης, αλλά δεν έχει ακόμη υλοποιηθεί.



Η **Μηχανική Μάθηση (machine learning)** είναι ένας κλάδος της Τεχνητής Νοημοσύνης, ο οποίος μελετά την ικανότητα των υπολογιστών να μαθαίνουν, χωρίς να προγραμματίζονται ρητά για τον σκοπό αυτό.

Σήμερα έχει εμφανιστεί μια νέα κατηγορία συστημάτων Τεχνητής Νοημοσύνης, τα **Μεγάλα Γλωσσικά Μοντέλα (Large Language Models, LLMs)**, τα οποία χρησιμοποιούν αρκετά πολύπλοκα μοντέλα νευρωνικών δικτύων, για να επεξεργάζονται και να παράγουν κείμενο. Τέτοια γνωστά μοντέλα είναι το ChatGPT και το Gemini, στα οποία μπορεί όλοι να έχουν πρόσβαση μέσω του Διαδικτύου. Τα Μεγάλα Γλωσσικά Μοντέλα εκπαιδεύονται σε τεράστια σύνολα κειμένου που προέρχονται από διάφορες πηγές, όπως βιβλία, άρθρα, ιστοσελίδες και άλλα κείμενα. Μέσω αυτής της διαδικασίας, εξάγουν κάποια μοντέλα, με βάση τα οποία μπορούν να προβλέπουν την επόμενη λέξη σε μια πρόταση ή ένα κείμενο. Τα αποτελέσματα αυτών των μοντέλων είναι εντυπωσιακά και οφείλονται στον τεράστιο όγκο δεδομένων με τον οποίο έχουν εκπαιδευτεί.

Τα Μεγάλα Γλωσσικά Μοντέλα σημειώνουν εντυπωσιακές επιδόσεις στην παραγωγή κειμένου, περιλήψεων ή μεταφράσεων. Ωστόσο, επειδή δε σκέφτονται με τον τρόπο που σκέφτεται ο άνθρωπος, δεν τα πάνε τόσο καλά σε πρωτότυπα προβλήματα, όσο απλά και να είναι αυτά, όπως φαίνεται στα παραδείγματα 1 και 2. Και στις δύο περιπτώσεις είναι φανερό ότι το ChatGPT αδυνατεί να καταλάβει ένα τόσο απλό πρόβλημα, διότι δεν έχει εκπαιδευτεί σε μια τέτοιου είδους διατύπωση ή τέτοιου είδους πρόβλημα. Ειδικά στη δεύτερη περίπτωση, παρόλο που «ξέρει» να λύσει το πρόβλημα με τον λύκο, το πρόβατο και το λάχανο που είναι πιο δύσκολο, αδυνατεί να λύσει το ίδιο πρόβλημα χωρίς το λάχανο, που είναι προφανές ακόμα και για ένα μικρό παιδί. Ωστόσο, τα μοντέλα αυτά εκπαιδεύονται σε όλο και μεγαλύτερο όγκο δεδομένων με αποτέλεσμα να βελτιώνονται συνεχώς.

Παράδειγμα 1 - Πόσα μήλα σου έμειναν;**You**

Έχω 999 μήλα και σου έδωσα το 1, Έφαγα 9. Πόσα σου έμειναν;

**ChatGPT**

Αν είχες αρχικά 999 μήλα και μου έδωσες ένα, τότε έμειναν 998 μήλα. Αν έφαγες 9 από αυτά, τότε μείνανε 989.

**You**

Δεν ρώτησα πόσα έμειναν σε μένα αλλά σε σένα

**ChatGPT**

Απολογούμαι για την παρερμηνεία. Αν εσύ έφαγες 9 μήλα από τα 999 που είχες αρχικά και στη συνέχεια μου έδωσες ένα, τότε εγώ έχω πλέον 989 μήλα.

Ωστόσο, αν διατυπώσουμε λίγο διαφορετικά τις παραπάνω ερωτήσεις, δηλαδή δοκιμάσουμε λίγο διαφορετική προτροπή (prompt), πάλι δε θα λύσει το πρόβλημα, αλλά μετά τη δεύτερη προτροπή θα φτάσει στη σωστή απάντηση:

Έχω 999 μήλα και σου έδωσα το 1. Έφαγα 9. Πόσα σου έμειναν;



Αρχικά, είχες 999 μήλα. Έδωσες σε εμένα 1 μήλο, άρα σου μένουν 998 μήλα. Έπειτα, έφαγες 9 μήλα. Άρα τα μήλα που σου έμειναν είναι:

$998 - 9 = 989$ μήλα.

Εσένα πόσα σου έμειναν όχι σε μένα.



Σωστά. Ας το υπολογίσουμε από την αρχή:

Είχες 999 μήλα και μου έδωσες 1 μήλο. Άρα εγώ έχω 1 μήλο. Δεν έφαγα κάποιο μήλο, οπότε μου μένει 1 μήλο.

Προσοχή!! Το γλωσσικό μοντέλο μπορεί να αξιοποιήσει και τις προηγούμενες προτροπές/ερωτήσεις (prompts) που έγιναν στην ίδια συζήτηση, όμως δε μαθαίνει από αυτήν τη συζήτηση. Τα μοντέλα αυτά είναι προ-εκπαιδευμένα σε σύνολα δεδομένων που φτάνουν μέχρι μια συγκεκριμένη χρονική στιγμή. Παρακάτω φαίνεται η απάντηση του ChatGPT στην ερώτηση «Πότε ήταν η τελευταία φορά που εκπαιδευτήκες/έμαθες από νέα δεδομένα;»



Η τελευταία φορά που ενημερώθηκα και εκπαιδεύτηκα με νέα δεδομένα ήταν τον Αύγουστο του 2023. Από τότε, δεν έχω λάβει νέες ενημερώσεις ή δεδομένα. Έτσι, οι πληροφορίες μου είναι ακριβείς μέχρι τότε, αλλά δεν περιλαμβάνουν αλλαγές ή γεγονότα που συνέβησαν μετά τον Αύγουστο του 2023.

Η ερώτηση αυτή έγινε στις 8 Αυγούστου του 2024.

Για το ίδιο πρόβλημα το Gemini της Google δίνει και αυτό λανθασμένη απάντηση:



Έχω 999 μήλα και σου έδωσα το 1. Έφαγα 9. Πόσα σου έμειναν;

Εμφάνιση προχείρων ▾



Αν σας έδωσα 1 μήλο και έφαγα 9, τότε σας έχουν μείνει **990 μήλα**.

Ας δούμε τα βήματα για την επίλυση του προβλήματος:

- Αρχικά, ξεκινάμε με 999 μήλα.
- Σας δίνω 1 μήλο, οπότε σας μένουν $999 - 1 = 998$ μήλα.
- Τρώω 9 μήλα, άρα αφαιρώ 9 από τα 998 μήλα που σας έμειναν, $998 - 9 = 989$ μήλα.

Συνεπώς, σας έχουν μείνει **990 μήλα**.

Παράδειγμα 2 - Ο λύκος, το πρόβατο και το λάχανο που λείπει



Ένας άνθρωπος, ένα πρόβατο και ένας λύκος βρίσκονται στη μια όχθη ενός ποταμού και έχουν μια βάρκα στην οποία χωράνε μόνο δύο, εκ των οποίων ο ένας πρέπει να είναι πάντα ο άνθρωπος/βαρκάρης. Πώς θα περάσουν απέναντι;



Ο άνθρωπος πηγαίνει πρώτος στην άλλη όχθη με τη βάρκα, αφήνοντας πίσω τον λύκο. Στη συνέχεια, ο άνθρωπος επιστρέφει με τη βάρκα στην αρχική όχθη, παίρνει το πρόβατο και το περνάει στην άλλη πλευρά. Όταν το πρόβατο φτάσει στην άλλη όχθη, ο άνθρωπος αφήνει το πρόβατο εκεί και παίρνει τον λύκο μαζί του πίσω στην αρχική όχθη. Τέλος, ο άνθρωπος περνάει ξανά τον ποταμό με τη βάρκα, αφήνοντας τον λύκο στην άλλη όχθη, και έτσι ολοκληρώνεται η μεταφορά.

όταν ο άνθρωπος αφήσει το πρόβατο στην άλλη όχθη που είναι και ο λύκος γιατί να μην βγει από τη βάρκα και αυτός αφού έχουν περάσει και οι τρεις? Γιατί πρέπει να γυρίσει πίσω;



Ο λόγος που ο άνθρωπος πρέπει να επιστρέψει πίσω είναι για να αποτρέψει τον λύκο να φάει το πρόβατο. Αν ο λύκος έμενε μόνος του με το πρόβατο στην άλλη όχθη, θα το έτρωγε, καθώς έχουν εχθρική σχέση. Έτσι, ο άνθρωπος πρέπει να επιστρέψει για να είναι εκεί να εμποδίσει τον λύκο να κάνει κάτι τέτοιο.

Σε αυτήν την περίπτωση, το γλωσσικό μοντέλο δεν μπορεί να λύσει μια πολύ πιο εύκολη εκδοχή του αρχικού προβλήματος, το οποίο έχει εκπαιδευτεί να λύνει. Το ερώτημα εδώ είναι αν τα μοντέλα αυτά «σκέφτονται».