



ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

**Μελέτη εφαρμογών Τεχνητής Νοημοσύνης που
χρησιμοποιούνται σε συστήματα κρίσιμων υποδομών**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Σαββόπουλου Σάββα

Επιβλέπουσα: Βασιλική Διαμαντοπούλου

Μέλη εξεταστικής επιτροπής: Κωνσταντίνος Μαλιάτσος, Σπυρίδων Κοκολάκης

Σάμος, Οκτώβριος 2024

Περιεχόμενα

Περίληψη	3
Abstract.....	4
Κεφάλαιο 1: Εισαγωγή	5
Κεφάλαιο 2: Εφαρμογές Τεχνητής Νοημοσύνης σε Κρίσιμες Υποδομές	13
2.1. Κατηγορίες κρίσιμων υποδομών που επωφελούνται από την TN	13
2.2. Παραδείγματα εφαρμογών TN σε κρίσιμες υποδομές.....	20
2.3. Εφαρμογή ενός ML αλγορίθμου σε ένα σύστημα κρίσιμων υποδομών, επιλεγμένο με βάση κατάλληλα κριτήρια	26
2.4. Ανάλυση των υπαρχόντων εργαλείων TN που χρησιμοποιούνται σε αυτόν τον τομέα.....	32
2.5. Αξιολόγηση εργαλείων TN που χρησιμοποιούν αλγορίθμους Μηχανικής Μάθησης (Machine Learning- ML) για τη διαχείριση κρίσιμων υποδομών	35
Κεφάλαιο 3: Θέματα ασφάλειας σχετικά με τη TN σε κρίσιμες υποδομές.....	38
3.1. TN στην επιθετική ασφάλεια.....	38
3.2. TN στην αμυντική ασφάλεια	41
3.3.Τρωτά σημεία στην TN	44
3.4. Απειλές που οφείλονται στην TN	47
Κεφάλαιο 4: Πλεονεκτήματα και Μειονεκτήματα της χρήσης TN σε κρίσιμες υποδομές.....	51
4.1. Αντιμετώπιση πιθανών μειονεκτημάτων και προκλήσεων	58
4.2. Ανάλυση Πλεονεκτημάτων και Προκλήσεων στη χρήση ML σε συστήματα κρίσιμων υποδομών	62
Συμπεράσματα	66
Επίλογος	69
Βιβλιογραφικές Αναφορές.....	73

Περίληψη

Η παρούσα ερευνητική εργασία εξετάζει τις πολύπλευρες εφαρμογές της Τεχνητής Νοημοσύνης (TN) στις κρίσιμες υποδομές, εστιάζοντας στον τρόπο με τον οποίο οι τεχνολογίες TN υποστηρίζουν την εποπτεία, τη λειτουργία και τη διαχείριση των πόρων που είναι απαραίτητοι για την αδιάλειπτη λειτουργία των συστημάτων που στηρίζουν τις κρίσιμες υπηρεσίες. Δεδομένης της ύψιστης σημασίας υποδομών όπως τα ενεργειακά δίκτυα, τα συστήματα ύδρευσης, τα δίκτυα μεταφορών και οι δίαυλοι επικοινωνίας για την εθνική ασφάλεια και την οικονομική σταθερότητα, η μελέτη διερευνά τον καθοριστικό ρόλο της TN στην ενίσχυση της ανθεκτικότητας, της αποδοτικότητας και της προσαρμοστικότητάς τους. Μέσα από μελέτες περιπτώσεων και εμπειρικά δεδομένα, διερευνάται η ενσωμάτωση της μηχανικής μάθησης, της βαθιάς μάθησης και άλλων τεχνολογιών TN στην παρακολούθηση της υγείας των υποδομών, τη διαχείριση των ισορροπιών ζήτησης-προσφοράς και τη διασφάλιση έναντι κυβερνο-φυσικών απειλών. Προκλήσεις όπως η ασφάλεια των δεδομένων, οι ηθικοί προβληματισμοί στα αυτόνομα συστήματα και οι κοινωνικοοικονομικές επιπτώσεις της ενσωμάτωσης της TN αναλύονται κριτικά. Στην παρούσα μελέτη παρουσιάζεται μια ολοκληρωμένη αξιολόγηση των εργαλείων Τεχνητής Νοημοσύνης (TN), ειδικά εκείνων που χρησιμοποιούν αλγόριθμους Μηχανικής Μάθησης (ML), για τη διαχείριση συστημάτων υποδομών ζωτικής σημασίας. Η έμφαση δίνεται στην αξιολόγηση του τρόπου με τον οποίο αυτές οι τεχνολογίες μπορούν να βελτιστοποιήσουν τη λειτουργία, τη συντήρηση και την ανθεκτικότητα βασικών υπηρεσιών, όπως η ενέργεια, οι μεταφορές, η διαχείριση υδάτων και τα δίκτυα τηλεπικοινωνιών. Δεδομένου του κρίσιμου ρόλου που διαδραματίζουν αυτές οι υποδομές για την κοινωνική και οικονομική σταθερότητα, η έρευνα αποσκοπεί στην παροχή μιας εμπειριστατωμένης ανάλυσης των εργαλείων TN που καθοδηγούνται από ML για την ενίσχυση της προγνωστικής συντήρησης, της επιχειρησιακής απόδοσης και των δυνατοτήτων ανίχνευσης απειλών.

Λέξεις Κλειδιά: *αλγόριθμοι Μηχανικής Μάθησης (ML), εργαλεία Τεχνητής Νοημοσύνης (TN), κρίσιμες υποδομές*

Abstract

This research paper examines the multifaceted applications of Artificial Intelligence (AI) in critical infrastructure, focusing on how AI technologies support the monitoring, operation and management of resources necessary for the uninterrupted operation of the systems that support them. critical services. Given the paramount importance of infrastructure such as energy grids, water systems, transportation networks and communication channels to national security and economic stability, the study explores the key role of AI in enhancing their resilience, efficiency and adaptability. Through case studies and empirical data, it explores the integration of machine learning, deep learning, and other AI technologies in monitoring infrastructure health, managing supply-demand balances, and securing against cyber-physical threats. Challenges such as data security, ethical concerns in autonomous systems and socio-economic impacts of AI integration are critically analyzed. This study presents a comprehensive evaluation of Artificial Intelligence (AI) tools, especially those using Machine Learning (ML) algorithms, for managing critical infrastructure systems. The emphasis is on assessing how these technologies can optimize the operation, maintenance and resilience of essential services such as energy, transport, water management and telecommunications networks. Given the critical role these infrastructures play in social and economic stability, the research aims to provide an in-depth analysis of ML-driven AI tools to enhance predictive maintenance, operational performance and threat detection capabilities.

Κεφάλαιο 1: Εισαγωγή

Παρά τις συναρπαστικές προοπτικές και τις εκτεταμένες δυνατότητες που προσφέρει η Τεχνητή Νοημοσύνη (TN), είναι απαραίτητο για τις επιχειρήσεις και τις κυβερνητικές αρχές να προβούν σε ενδελεχή αξιολόγηση των αναγκών τους, προκειμένου να μεγιστοποιήσουν τα δυνητικά οφέλη. Είναι κρίσιμο να ληφθούν άμεσα προληπτικά μέτρα για τη δημιουργία των απαραίτητων υποδομών. Στον πυρήνα της TN βρίσκεται η σημασία των δεδομένων. Η επεξεργασία μεγάλων συνόλων δεδομένων απαιτεί σημαντικούς υπολογιστικούς πόρους. Με την αυξανόμενη εφαρμογή των τεχνολογιών TN, υπάρχει επιτακτική ανάγκη για αύξηση του όγκου των δεδομένων, γεγονός που απαιτεί μεγαλύτερη υπολογιστική ισχύ τόσο σε διανεμημένες όσο και σε κεντρικές εγκαταστάσεις. Αυτή η ανάγκη υποστηρίζεται από υποδομές και αρχιτεκτονικές που έχουν σχεδιαστεί για τη δρομολόγηση, μεταφορά και αποθήκευση δεδομένων, οι οποίες πρέπει να λειτουργούν με άριστη αποδοτικότητα και αποτελεσματικότητα.

Η εμπορική αξιοποίηση της TN έχει σημειώσει εντυπωσιακή πρόοδο τα τελευταία χρόνια, με προβλέψεις που τοποθετούν το 2024 ως ένα ακόμη έτος σημαντικών εξελίξεων. Επιπλέον, οι καινοτομίες στον τομέα της κβαντικής πληροφορικής και των ψηφιακών διδύμων για τα βιομηχανικά περιβάλλοντα είναι στενά συνδεδεμένες με τις εφαρμογές της TN. Ο συνδυασμός αυτών των τεχνολογιών αναμένεται να επιφέρει πρωτοφανή επίπεδα αύξησης της παραγωγικότητας στις βιομηχανίες. Η TN από μόνη της προβλέπεται να δημιουργήσει 97 εκατομμύρια νέες θέσεις εργασίας μέχρι το 2025 και να συνεισφέρει ετησίως με περισσότερα από 4,4 τρισεκατομμύρια δολάρια στην παγκόσμια οικονομία. Σύμφωνα με εκτιμήσεις της Goldman Sachs, η TN έχει τη δυνατότητα να αυξήσει το παγκόσμιο Ακαθάριστο Εγχώριο Προϊόν (ΑΕΠ) κατά 7% την επόμενη δεκαετία.

Ένα σημαντικό μέρος των ωφελειών από την Τεχνητή Νοημοσύνη (TN) αναμένεται να προκύψει από την ενίσχυση της αποδοτικότητας μέσω της ευρείας χρήσης προγνωστικής ανάλυσης και διαφόρων άλλων τεχνικών βελτιστοποίησης των επιχειρησιακών διαδικασιών. Εκτιμάται ότι η TN έχει τη δυνατότητα να ενισχύσει ή να αυτοματοποιήσει έως και το 70% των υφιστάμενων εργασιακών διαδικασιών, με πιθανή αύξηση της συνολικής παραγωγικότητας κατά 3,3%. Αυτή η εξέλιξη θα επιτρέψει στους εργαζόμενους να αφιερώνουν περισσότερο χρόνο σε δραστηριότητες υψηλότερης προστιθέμενης αξίας, όπως τα στρατηγικά αναπτυξιακά έργα. Πέραν της αυτοματοποίησης, η TN μπορεί να ενισχύσει την κριτική σκέψη και την επίλυση σύνθετων

προβλημάτων, προσφέροντας σημαντικά πλεονεκτήματα σε διάφορους τομείς. Οι δυνατότητες αυτές επιτρέπουν στις επιχειρήσεις να επιτυγχάνουν υψηλότερα αποτελέσματα με μειωμένους πόρους, γεγονός που αποκτά ιδιαίτερη σημασία λόγω της δημογραφικής τάσης γήρανσης του εργατικού δυναμικού.

Η ενσωμάτωση της TN στα συστήματα κρίσιμων υποδομών αντιπροσωπεύει μια μετασχηματιστική στροφή προς την ενίσχυση της αποδοτικότητας, της ανθεκτικότητας και της ασφάλειας σε βασικές υπηρεσίες όπως η ενέργεια, η παροχή νερού, οι μεταφορές και τα δίκτυα επικοινωνιών. Μέσω της αξιοποίησης των δυνατοτήτων της TN για προγνωστική ανάλυση, παρακολούθηση σε πραγματικό χρόνο και αυτόνομη λήψη αποφάσεων, τα συστήματα αυτά μπορούν να προβλέπουν και να μετριάζουν τους κινδύνους, να βελτιστοποιούν τις λειτουργίες και να ανταποκρίνονται γρήγορα σε αναδυόμενες απειλές ή λειτουργικές ανωμαλίες. Για παράδειγμα, οι αλγόριθμοι TN μπορούν να προβλέψουν τις ανάγκες συντήρησης στα ενεργειακά δίκτυα, μειώνοντας τον χρόνο διακοπής λειτουργίας και αποτρέποντας βλάβες που θα μπορούσαν να οδηγήσουν σε εκτεταμένες διακοπές. Αντίστοιχα, στη διαχείριση του νερού, τη βιώσιμη χρήση των πόρων. Ωστόσο, η υιοθέτηση της TN στις υποδομές ζωτικής σημασίας απαιτεί επίσης αυστηρή προσοχή στα μέτρα κυβερνοασφάλειας, καθώς τα συστήματα αυτά καθίστανται όλο και περισσότερο διασυνδεδεμένα και, ως εκ τούτου, δυνητικά ευάλωτα σε εξελιγμένες κυβερνοεπιθέσεις. Η προσεκτική ανάπτυξη της TN σε αυτά τα πλαίσια υπόσχεται όχι μόνο να βελτιώσει την αποδοτικότητα και την αξιοπιστία των κρίσιμων υπηρεσιών, αλλά και να προωθήσει την καινοτομία στον τρόπο διαχείρισης και προστασίας αυτών των βασικών λειτουργιών απέναντι στις εξελισσόμενες προκλήσεις.

Τα Κυβερνο- φυσικά Συστήματα (Cyber-Physical Systems-CPS) είναι ενσωματωμένα συστήματα που συνδυάζουν υπολογιστικά στοιχεία, όπως ενσωματωμένους υπολογιστές και λογισμικό, με φυσικές διεργασίες, υποβοηθούμενα από αισθητήρες και ενεργοποιητές. Αυτά τα συστήματα έχουν σχεδιαστεί για να παρακολουθούν και να ελέγχουν φυσικές διεργασίες σε πραγματικό χρόνο, μέσω πολύπλοκων βρόχων ανάδρασης, όπου το φυσικό περιβάλλον επηρεάζει τους υπολογισμούς και τις αποφάσεις, οι οποίες με τη σειρά τους επηρεάζουν το φυσικό περιβάλλον. Τα βασικά συστατικά των CPS περιλαμβάνουν αισθητήρες για τη συλλογή δεδομένων, υπολογιστικές μονάδες για την επεξεργασία αυτών των δεδομένων και ενεργοποιητές για την εκτέλεση των απαραίτητων ενεργειών βάσει των υπολογισμών. Δίκτυα επικοινωνίας, τόσο

ενσύρματα όσο και ασύρματα, επιτρέπουν την απρόσκοπτη ανταλλαγή δεδομένων μεταξύ αυτών των στοιχείων, εξασφαλίζοντας συνεκτική λειτουργία σε όλο το σύστημα (Sheikh et al., 2022).

Τα CPS είναι κρίσιμα σε τομείς κρίσιμων υποδομών, όπως η ενέργεια, οι μεταφορές, η υγειονομική περίθαλψη και η βιομηχανία. Στον τομέα της ενέργειας, τα CPS είναι θεμελιώδη για τη λειτουργία των έξυπνων δικτύων, τα οποία χρησιμοποιούν δεδομένα σε πραγματικό χρόνο από αισθητήρες και προηγμένους αλγορίθμους για τη διαχείριση της διανομής ηλεκτρικής ενέργειας με αποδοτικότητα, εξασφαλίζοντας αξιόπιστη και σταθερή παροχή ενέργειας. Στις μεταφορές, τα CPS επιτρέπουν τη λειτουργία αυτόνομων οχημάτων, ενσωματώνοντας δεδομένα από πολυάριθμους αισθητήρες για την ασφαλή πλοήγηση και οδήγηση χωρίς ανθρώπινη παρέμβαση, ενισχύοντας έτσι την οδική ασφάλεια και τη διαχείριση της κυκλοφορίας. Τα συστήματα υγειονομικής περίθαλψης επίσης ωφελούνται από τα CPS, όπου η παρακολούθηση σε πραγματικό χρόνο και ο αυτόματος έλεγχος μπορούν να βελτιώσουν τα αποτελέσματα για τους ασθενείς, όπως με προηγμένες ιατρικές συσκευές, όπως οι βηματοδότες και οι αντλίες ινσουλίνης που ανταποκρίνονται στις ανάγκες του ασθενούς αυτόνομα.

Η ενσωμάτωση τους σε κρίσιμες υποδομές φέρνει πολλαπλά οφέλη, όπως βελτιωμένη αποδοτικότητα, ενισχυμένη ασφάλεια και μεγαλύτερη αξιοπιστία. Για παράδειγμα, στα έξυπνα κτίρια, τα CPS μπορούν να βελτιστοποιήσουν την κατανάλωση ενέργειας προσαρμόζοντας το φωτισμό, τη θέρμανση και τον κλιματισμό βάσει της κατοίκησης και των περιβαλλοντικών συνθηκών. Στη βιομηχανική αυτοματοποίηση, τα CPS επιτρέπουν τον ακριβή έλεγχο και τον συντονισμό των μηχανημάτων, οδηγώντας σε αυξημένη παραγωγικότητα και μειωμένο χρόνο διακοπής λειτουργίας (Sheikh et al., 2022). Ωστόσο, αυτή η ενσωμάτωση παρουσιάζει σημαντικές προκλήσεις, ιδιαίτερα σε ό,τι αφορά την ασφάλεια και την αξιοπιστία. Η εξασφάλιση ότι τα CPS είναι ασφαλή από κυβερνοεπιθέσεις είναι ζωτικής σημασίας, καθώς οποιαδήποτε παραβίαση θα μπορούσε να έχει σοβαρές φυσικές συνέπειες, ενδεχομένως διακόπτοντας τις βασικές υπηρεσίες. Επιπλέον, τα CPS πρέπει να είναι εξαιρετικά αξιόπιστα, καθώς οι αποτυχίες σε αυτά τα συστήματα θα μπορούσαν να οδηγήσουν σε καταστροφικά αποτελέσματα, όπως διακοπές ρεύματος στα δίκτυα ηλεκτρικής ενέργειας ή αποτυχίες σε κρίσιμες ιατρικές συσκευές.

Τα Κυβερνο-φυσικά Συστήματα αποτελούν μετασχηματιστικές τεχνολογίες που γεφυρώνουν το χάσμα μεταξύ υπολογιστικών και φυσικών κόσμων, παίζοντας καθοριστικό ρόλο στη βελτίωση της αποδοτικότητας και της αξιοπιστίας των κρίσιμων υποδομών. Παρά τα πολυάριθμα οφέλη που

προσφέρουν, η διασφάλιση της ασφάλειας και της αξιοπιστίας τους παραμένει πρωταρχικής σημασίας για την αποτροπή οποιωνδήποτε δυσμενών επιπτώσεων στις βασικές υπηρεσίες της κοινωνίας. Καθώς η τεχνολογία εξελίσσεται, ο ρόλος τους στις κρίσιμες υποδομές είναι πιθανό να αυξηθεί, υποκινούμενος από εξελίξεις σε τομείς όπως το Διαδίκτυο των Πραγμάτων (IoT), η τεχνητή νοημοσύνη και η υπολογιστική άκρη, που θα ενισχύσουν περαιτέρω τις δυνατότητες και την ενσωμάτωσή τους.

Η Μηχανική Μάθηση (ML) είναι ένας κλάδος της τεχνητής νοημοσύνης που επικεντρώνεται στην ανάπτυξη αλγορίθμων και στατιστικών μοντέλων που επιτρέπουν στους υπολογιστές να εκτελούν συγκεκριμένες εργασίες χωρίς ρητές οδηγίες, βασιζόμενοι σε μοτίβα και αναγωγή. Οι αλγόριθμοι ML δημιουργούν μοντέλα βασισμένα σε δείγματα δεδομένων, γνωστά ως δεδομένα εκπαίδευσης, για να κάνουν προβλέψεις ή να λαμβάνουν αποφάσεις χωρίς να προγραμματίζονται ρητά για την εκτέλεση της εργασίας (Alkhaleel, 2023). Τα κύρια συστατικά της ML περιλαμβάνουν την προεπεξεργασία δεδομένων, την εξαγωγή χαρακτηριστικών, την εκπαίδευση μοντέλων και την αξιολόγηση. Μαθαίνοντας από ιστορικά δεδομένα, τα συστήματα ML μπορούν να προσαρμοστούν σε νέα δεδομένα, να βελτιώνονται με την πάροδο του χρόνου και να κάνουν πιο ακριβείς προβλέψεις ή αποφάσεις.

Η Μηχανική Μάθηση μπορεί να ενισχύσει σημαντικά τις κρίσιμες υποδομές σε διάφορους τομείς όπως η ενέργεια, οι μεταφορές, η υγειονομική περίθαλψη και η βιομηχανία. Στον τομέα της ενέργειας, οι αλγόριθμοι ML μπορούν να βελτιστοποιήσουν τη λειτουργία των έξυπνων δικτύων προβλέποντας τα πρότυπα κατανάλωσης ενέργειας, εντοπίζοντας βλάβες στο σύστημα και βελτιστοποιώντας τη διανομή της ηλεκτρικής ενέργειας. Για παράδειγμα, η ML μπορεί να προβλέψει τη ζήτηση ενέργειας και να προσαρμόσει την προσφορά σε πραγματικό χρόνο για να αποτρέψει διακοπές ρεύματος και να μειώσει τη σπατάλη. Στις μεταφορές, η ML είναι καθοριστική για την ανάπτυξη αυτόνομων οχημάτων και έξυπνων συστημάτων διαχείρισης κυκλοφορίας. Αναλύοντας δεδομένα από διάφορους αισθητήρες και κάμερες κυκλοφορίας, η ML μπορεί να βελτιστοποιήσει τη ροή της κυκλοφορίας, να μειώσει τη συμφόρηση και να ενισχύσει την οδική ασφάλεια.

Στον τομέα της υγειονομικής περίθαλψης, η ML μπορεί να επαναστατήσει στην παροχή φροντίδας στους ασθενείς επιτρέποντας την πρόωμη διάγνωση και την εξατομικευμένη θεραπεία. Οι αλγόριθμοι ML μπορούν να αναλύσουν τεράστιες ποσότητες ιατρικών δεδομένων,

συμπεριλαμβανομένων των ηλεκτρονικών ιατρικών φακέλων, των αποτελεσμάτων εργαστηριακών εξετάσεων και των ιατρικών απεικονίσεων, για να εντοπίσουν μοτίβα που μπορεί να παραβλεφθούν από τους ανθρώπινους ιατρούς. Αυτό μπορεί να οδηγήσει σε έγκαιρη ανίχνευση ασθενειών όπως ο καρκίνος, πιο ακριβείς διαγνώσεις και προσαρμοσμένες θεραπευτικές συστάσεις με βάση τα δεδομένα του ατόμου. Επιπλέον, στη βιομηχανία, η ML μπορεί να βελτιώσει τη προληπτική συντήρηση, μειώνοντας τον χρόνο διακοπής λειτουργίας και επεκτείνοντας τη διάρκεια ζωής του εξοπλισμού (Alkhaleel, 2023). Αναλύοντας δεδομένα από αισθητήρες ενσωματωμένους σε μηχανήματα, τα μοντέλα ML μπορούν να προβλέψουν πότε είναι πιθανό να αποτύχει μια μηχανή και να προγραμματίσουν τη συντήρηση πριν από την εμφάνιση μιας βλάβης, αυξάνοντας έτσι την αποδοτικότητα και μειώνοντας το κόστος.

Ωστόσο, η ενσωμάτωση της ML σε κρίσιμες υποδομές παρουσιάζει επίσης σημαντικές προκλήσεις. Η εξασφάλιση της ασφάλειας και της ιδιωτικότητας των δεδομένων είναι ζωτικής σημασίας, καθώς οποιαδήποτε παραβίαση θα μπορούσε να έχει σοβαρές συνέπειες. Για παράδειγμα, στην υγειονομική περίθαλψη, η προστασία των δεδομένων των ασθενών από κυβερνοεπιθέσεις είναι κρίσιμη για τη διατήρηση της εμπιστοσύνης και τη συμμόρφωση με τους κανονισμούς. Ομοίως, στον τομέα της ενέργειας, η ασφάλεια των συστημάτων ML από κυβερνοαπειλές είναι ουσιώδης για την αποτροπή διαταραχών στην παροχή ενέργειας. Επιπλέον, η αξιοπιστία των μοντέλων ML είναι καθοριστική· πρέπει να λειτουργούν με ακρίβεια και συνέπεια σε πραγματικές συνθήκες. Η διασφάλιση ότι τα μοντέλα ML είναι διαφανή και οι διαδικασίες λήψης αποφάσεων τους κατανοητές είναι επίσης κρίσιμη για τη δημιουργία εμπιστοσύνης και την εξασφάλιση λογοδοσίας.

Η Μηχανική Μάθηση προσφέρει μετασχηματιστικό δυναμικό για την ενίσχυση της αποδοτικότητας, της αξιοπιστίας και της ασφάλειας των κρίσιμων υποδομών. Αξιοποιώντας δεδομένα και προηγμένους αλγορίθμους, η ML μπορεί να βελτιστοποιήσει τις λειτουργίες, να προβλέψει και να αποτρέψει αποτυχίες και να παρέχει εξατομικευμένες υπηρεσίες σε διάφορους τομείς. Παρά τα πολυάριθμα οφέλη που προσφέρει, η αντιμετώπιση των προκλήσεων της ασφάλειας, της ιδιωτικότητας, της αξιοπιστίας και της διαφάνειας είναι απαραίτητη για την πλήρη αξιοποίηση των πλεονεκτημάτων της ML στις κρίσιμες υποδομές.

Η χρήση εργαλείων Τεχνητής Νοημοσύνης (TN), και ειδικότερα εκείνων που βασίζονται σε αλγορίθμους μηχανικής μάθησης (ML), μετασχηματίζει ριζικά τη διαχείριση κρίσιμων υποδομών,

όπως τα ενεργειακά δίκτυα, τα δίκτυα μεταφορών, οι εγκαταστάσεις επεξεργασίας νερού και οι τηλεπικοινωνιακές υπηρεσίες. Τα εργαλεία αυτά διαδραματίζουν καθοριστικό ρόλο στην αξιοποίηση των δεδομένων για την πρόβλεψη βλαβών συστημάτων, τη βελτιστοποίηση της κατανομής πόρων και την ενίσχυση των μέτρων ασφαλείας έναντι τόσο φυσικών όσο και κυβερνοαπειλών. Μέσω της ανάλυσης προτύπων και ανωμαλιών σε εκτεταμένα σύνολα δεδομένων, οι αλγόριθμοι ML έχουν τη δυνατότητα να προβλέπουν πιθανές διαταραχές και να αυτοματοποιούν διορθωτικές ενέργειες, διασφαλίζοντας τη σταθερότητα και την αποτελεσματικότητα των υποδομών που είναι ζωτικής σημασίας για την κοινωνική λειτουργία και την οικονομική σταθερότητα.

Ωστόσο, παρά τις σημαντικές βελτιώσεις που υπόσχονται αυτές οι τεχνολογίες στη διαχείριση των υποδομών, ανακύπτουν σημαντικά ζητήματα σχετικά με την προστασία της ιδιωτικότητας των δεδομένων, την αλγοριθμική προκατάληψη και την ανάγκη για ισχυρές άμυνες κυβερνοασφάλειας. Καθώς τα συστήματα υποδομών ζωτικής σημασίας εξαρτώνται όλο και περισσότερο από την ΤΝ, η εξισορρόπηση μεταξύ καινοτομίας και διαχείρισης κινδύνων καθίσταται κεντρικό ζήτημα για τους ερευνητές και τους επαγγελματίες του κλάδου.

Η διατήρηση της βιωσιμότητας στις παραδοσιακές τεχνολογίες που είναι ζωτικής σημασίας για τις υποδομές, αναδεικνύεται σε μία από τις πιο κρίσιμες προκλήσεις για τις σύγχρονες κοινωνίες και τον επιχειρηματικό κόσμο. Επιπλέον, η ανάγκη για προσαρμογή αυτών των τεχνολογιών ώστε να αντιμετωπίζουν τις επιπτώσεις της κλιματικής αλλαγής, ειδικά σε ευαίσθητα οικοσυστήματα, καθίσταται ένα από τα πλέον επιτακτικά ζητήματα στο πλαίσιο της διαχείρισης κινδύνων και της ασφάλειας. Σύμφωνα με μελέτη των Ermolieva & Ermoliev (2005), οι κοινωνικές επιπτώσεις από καταστροφικά γεγονότα βρίσκονται σε συνεχή άνοδο, κυρίως λόγω της συγκέντρωσης πληθυσμών και κρίσιμων υποδομών σε περιοχές εκτεθειμένες σε κινδύνους, με την κατάσταση να επιδεινώνεται από την έλλειψη επαρκούς ενημέρωσης σχετικά με αυτούς τους κινδύνους. Πέραν των κοινωνικών επιπτώσεων, ο βιομηχανικός τομέας, ως ένας από τους πλέον κρίσιμους πυλώνες της οικονομίας, αντιμετωπίζει σημαντικές προκλήσεις σε περιπτώσεις αστοχίας, κάτι που καθιστά την ασφάλεια στον βιομηχανικό τομέα ένα ζήτημα ύψιστης σημασίας για την κοινωνία, όπως επισημαίνεται στην έκθεση Groth & Cortekar (2015). Σε αυτό το πλαίσιο, οι βιομηχανικές οντότητες και οι ρυθμιστικές αρχές προχωρούν σε αξιολογήσεις κινδύνου, προσπαθώντας να

προσδιορίσουν με ακρίβεια τις πιθανότητες αποτυχίας και τις συνέπειες που συνδέονται με τέτοια γεγονότα.

Η αντιμετώπιση των αβέβαιων κινδύνων που συνδέονται με κρίσιμες υποδομές παρουσιάζει ένα σημαντικό εμπόδιο λόγω της εκτεταμένης πολυπλοκότητας, του αυξημένου κόστους και της έλλειψης πληροφοριών που προκύπτει από τον μεγάλο αριθμό των εμπλεκόμενων μεταβλητών. Στο πλαίσιο αυτό, το παρόν έγγραφο επιδιώκει να εξετάσει και να συγκρίνει αλγορίθμους τεχνητής νοημοσύνης (AI) με τη δυνατότητα αποτελεσματικής αντιμετώπισης της πολυπλοκότητας και της αβεβαιότητας στην αξιολόγηση κινδύνων. Τα πλεονεκτήματα της εφαρμογής μεθόδων τεχνητής νοημοσύνης στην αξιολόγηση κινδύνων είναι σημαντικά, καθώς έχουν αποδειχθεί ιδιαίτερα αποτελεσματικά στην επίλυση μη-γραμμικών προβλημάτων εντός περιβαλλόντων υψηλής αβεβαιότητας. Ένα ενδεικτικό παράδειγμα, η συγκριτική μελέτη που παρουσιάζεται στην Bahrammirzaee (2010), δείχνει ότι η ακρίβεια των μεθόδων AI υπερέχει σε σχέση με τις παραδοσιακές στατιστικές τεχνικές, ιδιαίτερα σε συνθήκες ανάλυσης μη γραμμικών δεδομένων κινδύνου. Η χρήση τεχνητής νοημοσύνης στον τομέα της ασφάλειας εστιάζει στον πρωταρχικό στόχο της μίμησης της ανθρώπινης εμπειρογνωμοσύνης στην εκτίμηση κινδύνων και την καθιέρωση επιπέδων ανοχής.

Σύμφωνα με τον Herbert Simon, έναν κορυφαίο ειδικό στον τομέα της Τεχνητής Νοημοσύνης, η Τεχνητή Νοημοσύνη έχει παρουσιάσει καινοτόμες τεχνικές που βελτιώνουν τον διαδικαστικό ορθολογισμό των ατόμων που λαμβάνουν αποφάσεις, βοηθώντας στη διαχείριση της πολυπλοκότητας και του μεγάλου όγκου πληροφοριών. Στην περιοχή της αξιολόγησης κινδύνων, οι πιο κοινά χρησιμοποιούμενες τεχνολογίες Τεχνητής Νοημοσύνης περιλαμβάνουν συστήματα εμπειρογνομόνων, τεχνητά νευρωνικά δίκτυα και υβριδικά ευφυή συστήματα. Αυτές οι μεθοδολογίες Τεχνητής Νοημοσύνης διαδραματίζουν καθοριστικό ρόλο στην ενίσχυση της ακρίβειας και της αποτελεσματικότητας της διαδικασίας αξιολόγησης κινδύνων, παρέχοντας πολύτιμα εργαλεία για την επεξεργασία και την ανάλυση περίπλοκων δεδομένων σε δύσκολα περιβάλλοντα.

Οι εξελίξεις στην Τεχνητή Νοημοσύνη και τον Αυτοματισμό φέρνουν επανάσταση στις υποδομές ζωτικής σημασίας, υποσχόμενες αύξηση της αποδοτικότητας και της ανθεκτικότητας. Αυτά τα συστήματα, ζωτικής σημασίας για τη δημόσια ευημερία, περιλαμβάνουν ένα ευρύ φάσμα τομέων, από τα δίκτυα ηλεκτρικής ενέργειας έως την υγειονομική περίθαλψη. Αντί να αντικαταστήσει τον

άνθρωπο, η TN ενισχύει τις ανθρώπινες ικανότητες, επιτρέποντας την εστίαση σε καθήκοντα υψηλότερου επιπέδου. Ωστόσο, εξακολουθούν να υφίστανται προκλήσεις, όπως η γήρανση των υποδομών και ηθικά ζητήματα. Η εξισορρόπηση της καινοτομίας με την ασφάλεια είναι υψίστης σημασίας, καθώς η TN ανοίγει νέα σύνορα στην κυβερνοασφάλεια, προσφέροντας τόσο τρωτά σημεία όσο και λύσεις. Παρά τους κινδύνους, το μέλλον των υποδομών ζωτικής σημασίας έγκειται στην αξιοποίηση των δυνατοτήτων της TN για την οικοδόμηση ενός ασφαλέστερου και αποτελεσματικότερου κόσμου.

Τα επόμενα χρόνια, αναμένεται να υπάρξει σημαντική αύξηση της χρήσης και της ενσωμάτωσης της αυτοματοποίησης και της ενσωμάτωσης της TN στα συστήματα κρίσιμων υποδομών. Τα συστήματα αυτά είναι απαραίτητα για τη δημόσια υγεία, την ασφάλεια και την ευημερία και καλύπτουν ένα ευρύ φάσμα, συμπεριλαμβανομένων των δικτύων ηλεκτρικής ενέργειας, των πετρελαιαγωγών, του ελέγχου εναέριας κυκλοφορίας, των τηλεπικοινωνιακών δικτύων, των μεταφορών, της υγειονομικής περίθαλψης και των χρηματοπιστωτικών ιδρυμάτων και υπηρεσιών.

Με την εμφάνιση αυτών των τεχνολογιών, υπάρχει μια αλλαγή παραδείγματος στον τρόπο με τον οποίο αντιλαμβανόμαστε και διαχειριζόμαστε αυτά τα κρίσιμα συστήματα. Σε αντίθεση με τη δημοφιλή πεποίθηση, ο ρόλος της τεχνητής νοημοσύνης και της αυτοματοποίησης στις κρίσιμες υποδομές δεν αφορά την αντικατάσταση των ανθρώπων, αλλά μάλλον την επαύξηση των δυνατοτήτων τους.

Η παρούσα εργασία έχει ως στόχο να ερευνήσει τις εφαρμογές τεχνητής νοημοσύνης σε κρίσιμες υποδομές καθώς και τις κατηγορίες κρίσιμων υποδομών που επωφελούνται από την TN. Θα αναφερθούν παραδείγματα εφαρμογών tn σε κρίσιμες υποδομές και θα αναληθούν τα υπάρχοντα εργαλεία TN που χρησιμοποιούνται σε αυτόν τον τομέα. Τέλος, θα αξιολογηθούν τα εργαλεία TN που χρησιμοποιούν αλγορίθμους Μηχανικής Μάθησης (Machine Learning- ML) για τη διαχείριση κρίσιμων υποδομών. Τα ερευνητικά ερωτήματα που προκύπτουν μέσα από τη βιβλιογραφική έρευνα είναι τα εξής: Πώς η TN υποστηρίζει την εποπτεία, λειτουργία, διαχείριση απαιτούμενων πόρων των συστημάτων που υποστηρίζουν κρίσιμες υποδομές. Ποια είναι τα υπάρχοντα εργαλεία; Ποια είναι τα πλεονεκτήματα και ποι μειονεκτήματα χρήσης TN. Η παρούσα ανασκόπηση δεν επιδιώκει μόνο να εξακριβώσει ποιες TN τεχνικές χρησιμοποιούνται ευρέως σε συστήματα υποδομής, αλλά και να συγκρίνει την ωριμότητα και το βάθος της έρευνας σε όλα τα συστήματα, με την ελπίδα ότι θα εντοπιστούν πιθανά ερευνητικά κενά.

Κεφάλαιο 2: Εφαρμογές Τεχνητής Νοημοσύνης σε Κρίσιμες Υποδομές

2.1. Κατηγορίες κρίσιμων υποδομών που επωφελούνται από την ΤΝ

Ο όρος «κρίσιμη υποδομή» αναφέρεται στα φυσικά και εικονικά συστήματα, δίκτυα και περιουσιακά στοιχεία που είναι απαραίτητα για τη λειτουργία της κοινωνίας και της οικονομίας. Η Τεχνολογία Πληροφοριών (ΤΠ) και η Επιχειρησιακή Τεχνολογία (ΟΤ) είναι ζωτικής σημασίας για την κρίσιμη υποδομή. Ο όρος "ΤΠ" συνήθως αναφέρεται στα τυποποιημένα υπολογιστικά συστήματα που χρησιμοποιούνται για την επεξεργασία, αποθήκευση και επικοινωνία δεδομένων εντός των οργανισμών. Παραδείγματα περιλαμβάνουν διακομιστές, επιτραπέζιους υπολογιστές, φορητούς υπολογιστές, δικτυακές συσκευές και εφαρμογές λογισμικού. Αντίθετα, ο όρος "ΟΤ" αναφέρεται στα συστήματα υλικού και λογισμικού που χρησιμοποιούνται για την παρακολούθηση και τον έλεγχο φυσικών λειτουργιών σε διάφορους τομείς της κρίσιμης υποδομής, όπως η ενέργεια, η μεταποίηση, οι μεταφορές και η παροχή νερού. Παραδείγματα συστημάτων ΟΤ περιλαμβάνουν τα Συστήματα Εποπτικού Ελέγχου και Απόκτησης Δεδομένων (SCADA), τα Κατανεμημένα Συστήματα Ελέγχου (DCS), τα Βιομηχανικά Συστήματα Ελέγχου (ICS) και τους Προγραμματιζόμενους Λογικούς Ελεγκτές (PLC) (Kayan et al., 2022). Αυτές οι υποδομές παρέχουν υπηρεσίες ζωτικής σημασίας και υποστήριξη, διασφαλίζοντας την ομαλή λειτουργία διαφόρων τομέων και την ευημερία του πληθυσμού μιας χώρας. Οι τομείς των κρίσιμων υποδομών είναι αλληλένδετοι και εξαρτώνται ο ένας από τον άλλο για τη σωστή λειτουργία τους.

Ωστόσο, η ΚτΠ είναι επίσης ευάλωτη σε καταστροφές ή απειλές, καθώς μπορεί να έχει άμεσο αντίκτυπο στην κοινωνία (Ampratwum, et al., 2023.). Είναι κρίσιμα επειδή η διατάραξη ή η καταστροφή τους μπορεί να επηρεάσει σοβαρά τη δημόσια ασφάλεια, την εθνική ασφάλεια, την οικονομική σταθερότητα και τη δημόσια υγεία. Ενώ τα ΤΠ συστήματα είναι υπεύθυνα για τη διαχείριση δεδομένων και την επικοινωνία, τα συστήματα ΟΤ ελέγχουν τις φυσικές διεργασίες που διαχειρίζονται βασικές υπηρεσίες. Συνεπώς, η προστασία και η διασφάλιση αυτών των υποδομών από φυσικές και κυβερνοαπειλές είναι ζωτικής σημασίας για τη διατήρηση της εθνικής συνολικής ευημερία και σταθερότητα ενός έθνους.

Οι έξυπνες τεχνολογίες αποτελούν μέρος πολλών πτυχών της καθημερινής μας ζωής, συμβάλλοντας σημαντικά στην πρόοδο της κοινωνίας μέσω της χρήσης διασυνδεδεμένων συσκευών που στοχεύουν στη βελτίωση του καθημερινού βίου. Από τα smartphones και τα tablets

μέχρι τις έξυπνες συσκευές και τα φώτα που ελέγχονται μέσω του διαδικτύου, οι έξυπνες τεχνολογίες διαδραματίζουν ρόλο σε πολλά στοιχεία της καθημερινής ζωής. Πέρα από τις καθημερινές εφαρμογές, οι τεχνολογίες πληροφόρησης και επικοινωνίας (ICTs) έχουν διαδραματίσει βασικό ρόλο στον σχηματισμό των οικονομικών δραστηριοτήτων και της αστικής υποδομής (Hollands, 2020). Η εκθετική τεχνολογική ανάπτυξη έχει προκαλέσει μεγάλο ενδιαφέρον και συζητήσεις σχετικά με την ενσωμάτωση των ICTs σε αστικά αναπτυξιακά έργα όπως τα έξυπνα δίκτυα (smart grids) και οι έξυπνες πόλεις (smart cities).

Η έξυπνη τεχνολογία, η οποία περιλαμβάνει την ενσωμάτωση αισθητήρων και δικτυωμένων υποδομών, βρίσκει εφαρμογή σε πολλούς τομείς σήμερα. Χάρη στην έξυπνη τεχνολογία, έχουμε τη δυνατότητα να βελτιώσουμε κάθε υπηρεσία ή προϊόν, καθώς αυτή έχει ήδη εισαχθεί και εφαρμόζεται εδώ και δεκαετίες. Για παράδειγμα, κατά τη δεκαετία του 1980, φοιτητές του Πανεπιστημίου Carnegie Mellon χρησιμοποίησαν αισθητήρες που ήταν συνδεδεμένοι στο διαδίκτυο σε έναν αυτόματο πωλητή, προκειμένου να παρακολουθούν πόσα αναψυκτικά διατίθενται. Αυτό επέτρεψε στους προμηθευτές να παρακολουθούν τις πωλήσεις των προϊόντων αποτελεσματικότερα και να βελτιώσουν την παράδοση και εξυπηρέτησή τους. Παρόμοια παραδείγματα της εφαρμογής έξυπνης τεχνολογίας είναι πολλά και συνεχώς αυξάνονται τόσο σε ποικιλία όσο και σε πολυπλοκότητα (Graham & Marvin, 1999).

Μία από τις σημαντικότερες εφαρμογές της έξυπνης τεχνολογίας και του Διαδικτύου των Πραγμάτων (IoT) αφορά τον τομέα της υγειονομικής περίθαλψης, ο οποίος παρουσιάζει τεράστιο δυναμικό και προοπτικές ανάπτυξης. Σε συνδυασμό με το Cloud Computing, το IoT έχει κερδίσει έμφαση και αναγνώριση για τη συμβολή του στην εξέλιξη των υγειονομικών υπηρεσιών (Sakhnini & Karimipour, 2023). Σύμφωνα με την εταιρεία αναλύσεων αγοράς Grand View Research, η αγορά του IoT στην υγειονομική περίθαλψη έφτασε τα 330 δισεκατομμύρια δολάρια το 2020. Η ενσωμάτωση διαδικτυακών συνδέσεων στις ιατρικές συσκευές έχει τη δυνατότητα να ενισχύσει τις δυνατότητες παρακολούθησης και διάγνωσης. Η ιατρική βιομηχανία προσανατολίζεται προς την υιοθέτηση του IoT και των δικτυακών υποδομών, με τις νέες τεχνολογίες να προτείνονται για την ανάλυση διαφόρων ιατρικών παραμέτρων, όπως ο καρδιακός ρυθμός και ο κορεσμός οξυγόνου στο αίμα. Παράλληλα, τα συστήματα παρακολούθησης υγείας με βάση το cloud προσφέρουν ενισχυμένη επιτήρηση από το ιατρικό προσωπικό, συμβάλλοντας στη βελτίωση της ποιότητας της περίθαλψης και στη μείωση του κόστους (Ruzbahani et al., 2016).

Η έννοια του έξυπνου δικτύου (smart grid) περιλαμβάνει την ενσωμάτωση έξυπνων μετρητών, αισθητήρων και προηγμένων πληροφοριακών τεχνολογιών στα συστήματα ηλεκτρικής ενέργειας. Αυτή η τεχνολογία βελτιώνει σημαντικά την αποδοτικότητα της παραγωγής ενέργειας και ενθαρρύνει την ενοποίηση διαφορετικών πηγών ενέργειας σε ένα ενιαίο σύστημα (Karimipour & Dinavahi, 2017). Οι επικοινωνίες και τα κέντρα ελέγχου του δικτύου υποστηρίζουν τη συνδυασμένη χρήση πηγών πράσινης ενέργειας. Η σύνδεση των έξυπνων μετρητών και αισθητήρων σε όλο το ηλεκτρικό δίκτυο επιτρέπει στα κέντρα παραγωγής να έχουν πρόσβαση σε δεδομένα ζήτησης ενέργειας σε πραγματικό χρόνο, γεγονός που μπορεί να χρησιμοποιηθεί για την εφαρμογή αποδοτικών σχεδίων παραγωγής και διανομής (Yang et al., 2016). Ως εκ τούτου, η ενσωμάτωση αυτών των τεχνολογιών στην υποδομή του ηλεκτρικού δικτύου έχει αυξήσει σημαντικά την ενεργειακή αποδοτικότητα, ενώ ταυτόχρονα έχει μειώσει το κόστος της ηλεκτρικής ενέργειας.

Η Τεχνητή Νοημοσύνη (TN) έχει τη δυνατότητα να βελτιώσει ουσιαστικά την αποτελεσματικότητα, την ασφάλεια και τη λειτουργικότητα σε διάφορους τομείς κρίσιμων υποδομών. Ορισμένοι βασικοί τομείς στους οποίους η TN έχει σημαντικό αντίκτυπο είναι η ενέργεια και οι υπηρεσίες κοινής ωφέλειας. Η TN μπορεί να προβλέψει τις διακυμάνσεις της ζήτησης και της προσφοράς, να βελτιστοποιήσει τις λειτουργίες του δικτύου και να διαχειριστεί πιο αποτελεσματικά τις ανανεώσιμες πηγές ενέργειας. Τα έξυπνα δίκτυα, ενισχυμένα με TN, βοηθούν στην παρακολούθηση και διαχείριση της ροής ενέργειας σε πραγματικό χρόνο, αυξάνοντας την ανθεκτικότητα και την αποτελεσματικότητα της διανομής ενέργειας.

Στον τομέα των μεταφορών, η TN βελτιώνει τη διαχείριση της κυκλοφορίας, μειώνει τη συμφόρηση και ενισχύει την ασφάλεια και την αξιοπιστία των μέσων μαζικής μεταφοράς. Τα αυτόνομα οχήματα, συμπεριλαμβανομένων των μη επανδρωμένων αεροσκαφών για εφοδιαστική και παράδοση, αξιοποιούν την TN για την πλοήγηση και τη βελτιστοποίηση των διαδρομών. Επιπλέον, οι εφαρμογές TN στη διαχείριση υδάτων περιλαμβάνουν την προγνωστική συντήρηση των υποδομών, την ανίχνευση διαρροών και την παρακολούθηση της ποιότητας του νερού. Η TN μπορεί να βελτιστοποιήσει τις διαδικασίες επεξεργασίας νερού και τη χρήση ενέργειας, οδηγώντας σε πιο βιώσιμες πρακτικές διαχείρισης υδάτων.

Ένας άλλος σημαντικός τομέας είναι η υγεία. Η TN στην υγειονομική περίθαλψη διευκολύνει την πρόβλεψη ασθενειών, τη διάγνωση και την εξατομίκευση της θεραπείας μέσω ανάλυσης

δεδομένων και μοντέλων μηχανικής μάθησης. Τα εργαλεία TN μπορούν να παρακολουθούν την υγεία των ασθενών σε πραγματικό χρόνο και να ειδοποιούν το ιατρικό προσωπικό για αλλαγές που απαιτούν άμεση προσοχή, βελτιώνοντας τους χρόνους απόκρισης και τα θεραπευτικά αποτελέσματα. Επιπλέον, η TN ενισχύει τη βελτιστοποίηση των τηλεπικοινωνιακών δικτύων, προβλέπει δυσλειτουργίες υλικού και βελτιώνει την εξυπηρέτηση πελατών μέσω αυτοματοποιημένων συστημάτων και chatbots. Η TN μπορεί να διαχειριστεί μαζικές ροές δεδομένων, βελτιστοποιώντας την κατανομή του εύρους ζώνης και ενισχύοντας την ασφάλεια των επικοινωνιών.

Η TN μπορεί επίσης να ενισχύσει τη δημόσια ασφάλεια και την προστασία μέσω αναγνώρισης προσώπου και ανίχνευσης μη φυσιολογικής συμπεριφοράς. Συμβάλλει στην αντιμετώπιση καταστροφών αναλύοντας δεδομένα από διάφορες πηγές για την πρόβλεψη των επιπτώσεων των καταστροφών και τη βελτιστοποίηση της ανταπόκρισης σε καταστάσεις έκτακτης ανάγκης. Κάθε ένας από αυτούς τους τομείς επωφελείται από την TN, βελτιώνοντας τη λειτουργική αποδοτικότητα, μειώνοντας το κόστος, ενισχύοντας την ασφάλεια και την προστασία και διευκολύνοντας καλύτερες διαδικασίες λήψης αποφάσεων. Καθώς η τεχνολογία της TN συνεχίζει να εξελίσσεται, η ενσωμάτωσή της στις κρίσιμες υποδομές είναι πιθανό να επιφέρει περαιτέρω μετασχηματιστικές αλλαγές σε αυτούς τους σημαντικούς τομείς.

CPS στην υγειονομική περίθαλψη

Οι πρόσφατες εξελίξεις στον τομέα των ιατρικών συσκευών έχουν επικεντρωθεί στην ενσωμάτωση πολύπλοκων υπολογιστικών δικτύων, των CPS (Lee & Sokolsky, 2010). Αυτή η αλλαγή στην τεχνολογία επιτρέπει τη συνεχή παρακολούθηση και έλεγχο των φυσιολογικών διεργασιών και λειτουργιών των ασθενών. Η ενσωμάτωση των CPS στις ιατρικές συσκευές παρέχει επίσης τη δυνατότητα παρακολούθησης της κατάστασης των ασθενών από απόσταση (Milenković et al., 2006). Επιπλέον, η ενσωματωμένη ανίχνευση, υπολογιστών, επικοινωνιών και η συνδυασμένη χρήση με φυσικά στοιχεία και διεργασίες επιτρέπουν σε αυτά τα CPS να επιτύχουν ένα επίπεδο λειτουργικότητας που υπερβαίνει τα απλούστερα παθητικά συστήματα. Τα CPS έχουν επίσης τη δυνατότητα να προσεγγίσουν το σώμα χρησιμοποιώντας ελάχιστα επεμβατικές τεχνικές. Επιπλέον, μειώνουν το κόστος, ενισχύουν την κινητικότητα, την ανεξαρτησία και βελτιώνουν την ποιότητα της ζωής. Η ευαίσθητη φύση της αντιμετώπισης ιατρικών θεμάτων καθιστά τον σχεδιασμό των CPS κρίσιμο στην υγειονομική περίθαλψη. Η συνεχής εκθετική ανάπτυξη των

ιατρικών συσκευών απαιτεί ολοκληρωμένες και αποτελεσματικές δοκιμές. Επιπλέον, τα πρότυπα της υγειονομικής περίθαλψης είναι από τα υψηλότερα σε όλους τους τομείς της τεχνολογίας. Η ασφάλεια και η ιδιωτικότητα είναι επίσης μεταξύ των βασικών ανησυχιών για τα CPS στην υγειονομική περίθαλψη. Η ροή ευαίσθητων δεδομένων και η πρόσβαση των CPS στο σώμα σημαίνουν ότι οι επιθέσεις μπορούν να έχουν καταστροφικές και μερικές φορές απειλητικές για τη ζωή επιπτώσεις.

CPS στις μεταφορές

Για να ανταποκριθούν στις αυξανόμενες απαιτήσεις της κοινωνίας, τα συστήματα μεταφορών αναπτύσσονται και εξελίσσονται προς ολοένα και πιο πολύπλοκα συστήματα. Το Ευρωπαϊκό Σύστημα Διαχείρισης Σιδηροδρομικής Κυκλοφορίας (ERTMS) αποτελεί ένα εξαιρετικό παράδειγμα ενσωμάτωσης CPS στις μεταφορές. Το ERTMS χρησιμοποιεί την επικοινωνία GSM, ή αλλιώς παγκόσμιο σύστημα κινητών επικοινωνιών, για τη σύνδεση των τρένων και του συστήματος υποδομής για καλύτερη διαχείριση. Πολλές άλλες εταιρείες και οργανισμοί ενσωματώνουν CPS στις μεταφορές. Cloud Computing και διαδικτυακές βάσεις δεδομένων χρησιμοποιούνται για την παρακολούθηση και τη βελτιστοποίηση των λεωφορείων ή των δρομολογίων των τρένων. Επιπλέον, τα έξυπνα τηλέφωνα διαδραματίζουν ολοένα και μεγαλύτερο ρόλο στις μεταφορές, ένα έξυπνο τηλέφωνο μπορεί να λειτουργήσει ως εισιτήριο τρένου, ανιχνευτής GPS και, σε ορισμένες περιπτώσεις, να συνδεθεί και να λειτουργήσει ορισμένα μέρη ενός οχήματος. Οι σύγχρονες λύσεις για την κινητικότητα δίνουν έμφαση στις αυτοματοποιημένες μορφές μεταφορών, με διασυνδεδεμένα οχήματα, για την αντιμετώπιση των κοινωνικών στόχων της αύξησης της ασφάλειας, της προστασίας, της ευκολίας και της αποδοτικότητας.

CPS στη μεταποίηση

Η μεταποίηση ήταν πρωτοπόρος στην αυτοματοποίηση. Εργαλεία αυτοματισμού, όπως βιομηχανικά ρομπότ και αυτόνομες μηχανές, εισήλθαν στον τομέα της μεταποίησης προτού εφαρμοστούν σε άλλους τομείς. Στη μεταποίηση, υπάρχουν CPS σε διάφορες μορφές, από τρισδιάστατους εκτυπωτές έως την κατασκευή στο νέφος. Επομένως, η ενσωμάτωση των CPS στην κατασκευή δημιουργεί ευκαιρίες για προσαρμοσμένα συστήματα κατασκευής που μπορούν να είναι επωφελείς και προσιτές για τις μικρές επιχειρήσεις. Επιπλέον, ο κύκλος ζωής ενσωμάτωσης εισάγεται από τις βιομηχανικές εταιρείες, η οποία περιλαμβάνει την ανίχνευση των

δεδομένων πίσω στην ανάπτυξη και την κατασκευή για μελλοντική βελτίωση. Τα CPS επίσης παρέχουν λύσεις για αυξημένη βιωσιμότητα μέσω ακριβούς και συνεχούς παρακολούθησης των φυσικών συστημάτων.

CPS σε συστήματα ισχύος: Το έξυπνο δίκτυο

Το έξυπνο δίκτυο αποτελεί μία από τις πιο σημαντικές και μελετημένες εφαρμογές των CPS στον τομέα των συστημάτων ισχύος, στα οποία περιλαμβάνονται τα συστήματα ηλεκτρικής ενέργειας της επόμενης γενιάς. Η ανάπτυξη των σημερινών συστημάτων ηλεκτρικής ενέργειας στοχεύει στην ενσωμάτωση έξυπνων μετρητών και αισθητήρων, καθώς και προηγμένων υπολογιστικών τεχνολογιών για τη βελτίωση της αποδοτικότητας της ηλεκτροπαραγωγής. Η σύνδεση των έξυπνων μετρητών και αισθητήρων κατά μήκος του δικτύου ηλεκτρικής ενέργειας επιτρέπει τη συλλογή δεδομένων ζήτησης ισχύος σε πραγματικό χρόνο, τα οποία μπορούν να χρησιμοποιηθούν για την αποτελεσματικότερη διανομή ενέργειας (Fang et al., 2011). Η ενσωμάτωση αυτών των τεχνολογιών στην υποδομή του συστήματος ηλεκτρικής ενέργειας έχει αυξήσει σημαντικά την ενεργειακή απόδοση και μειώσει την τιμή της ηλεκτρικής ενέργειας. Επιπλέον, το έξυπνο δίκτυο αποτελείται από διάφορους πόρους και τεχνολογίες, όπως οι έξυπνοι μετρητές για τη συλλογή δεδομένων κατανάλωσης ενέργειας και η διασύνδεση του εποπτικού ελέγχου και της απόκτησης δεδομένων (SCADA) για την πιο διευρυμένη κεντρική διανομή κατά μήκος μεγάλων γεωγραφικών περιοχών (Karimipour & Dinavahi, 2015). Το έξυπνο δίκτυο επιτρέπει επίσης την αλληλεπίδραση μεταξύ του δικτύου μεταφοράς και του δικτύου διανομής, των ελεγκτών κτιρίων, καθώς και διαφόρων πηγών παραγωγής ενέργειας.

Εφαρμογές της τεχνητής νοημοσύνης στην ασφάλεια στον κυβερνοχώρο

Η Τεχνητή Νοημοσύνη (TN) είναι ένας ταχέως αναπτυσσόμενος τομέας στη βιομηχανία της τεχνολογίας. Πολλοί ερευνητές και προγραμματιστές εφαρμόζουν μεθόδους και τεχνικές που προσομοιώνουν την ανθρώπινη νοημοσύνη σε αλγοριθμικές λειτουργίες. Η τεχνητή νοημοσύνη και η εμφάνιση της βαθιάς μάθησης έχουν αυξήσει τις δυνατότητες πολλών εφαρμογών. Η τεχνητή νοημοσύνη, σε αντίθεση με τα συνήθη εργαλεία προγραμματισμού, δεν απαιτεί μια σαφώς μοντελοποιημένη διαδικασία- σε πολλές περιπτώσεις, η TN μπορεί να χρησιμοποιήσει δεδομένα εισόδου-εξόδου για να μάθει και να δημιουργήσει παραμέτρους συστήματος που μπορούν να ταξινομήσουν και να προβλέψουν μη μαθημένα δεδομένα. Ένα τέτοιο εργαλείο

μπορεί να έχει ευρύ φάσμα εφαρμογών, όπως η αναγνώριση ομιλίας, η αναγνώριση προσώπου και η ρομποτική. Η ασφάλεια στον κυβερνοχώρο είναι επίσης μία από τις εφαρμογές της TN που έχουν μελετηθεί σε μεγάλο βαθμό (Kiwia et al., 2018) στην πραγματικότητα, οι περισσότεροι οργανισμοί που ασχολούνται με την TN εστιάζουν στην ανίχνευση και την αποτροπή εισβολών ασφαλείας.

Τα ευφυή συστήματα χρησιμοποιούνται επίσης για την παρακολούθηση κακόβουλου λογισμικού και την ανίχνευση εισβολών. Επιπλέον, οι αλγόριθμοι μηχανικής μάθησης (ML) χρησιμοποιούνται συνήθως για την ταξινόμηση και ανίχνευση απειλών στον κυβερνοχώρο στα CPS. Η άφθονη χρήση της τεχνητής νοημοσύνης στην ασφάλεια στον κυβερνοχώρο οφείλεται στην υψηλή δυνητική αποδοτικότητα και την επεκτασιμότητά της. Σε αντίθεση με τις λύσεις που βασίζονται σε μοντέλα, τα έξυπνα συστήματα έχουν τη δυνατότητα να μαθαίνουν και έτσι να προσαρμόζονται σε μεγαλύτερα συστήματα.

Το ζήτημα με τη χρήση της μηχανικής μάθησης (ML) στην ασφάλεια στον κυβερνοχώρο προέρχεται κυρίως από την εξάρτηση αυτών των αλγορίθμων από ένα σύνολο μεταβλητών ή χαρακτηριστικών. Κατά το σχεδιασμό ενός αλγορίθμου ML για την ανίχνευση ενός κακόβουλου λογισμικού, για παράδειγμα, τα βασικά χαρακτηριστικά ενός κακόβουλου λογισμικού πρέπει να προσδιοριστούν. Και αν ορισμένες μεταβλητές δεν χρησιμοποιούνται ως χαρακτηριστικά, ο αλγόριθμος ML θα τις αγνοήσει. Ως εκ τούτου, η επιλογή χαρακτηριστικών ή η εξαγωγή χαρακτηριστικών είναι μια βασική έννοια στην ανάπτυξη αλγορίθμων ML. Ως εκ τούτου, για να διατηρηθεί υψηλή ακρίβεια ταξινόμησης, οι ML αλγόριθμοι πρέπει να χρησιμοποιούν μια ολοκληρωμένη και αποτελεσματική τεχνική εξαγωγής χαρακτηριστικών (Karimipour et al., 2019).

Για να ξεπεραστεί αυτό το ελάττωμα, οι ερευνητές άρχισαν να αλλάζουν από παραδοσιακούς ML αλγορίθμους σε βαθιά νευρωνικά δίκτυα ή βαθιά μάθηση (DL). Το DL είναι ένας υποτομέας του ML με τη δυνατότητα άμεσης εκπαίδευσης σε πρωτότυπα δεδομένα χωρίς την εξαγωγή των χαρακτηριστικών τους. Κατά την τελευταία δεκαετία, η DL έχει βελτιώσει σημαντικά τις επιδόσεις στους τομείς της όρασης υπολογιστών, της αναγνώρισης ομιλίας και του κειμένου επιτυγχάνοντας ένα ιστορικό άλμα στον τομέα της τεχνητής νοημοσύνης (Deng, & Yu, 2014). Τα κύρια πλεονεκτήματα της DL είναι η ικανότητά του να ανιχνεύει μη γραμμικές σχέσεις και να υποστηρίζει νέους τύπους αρχείων και ως εκ τούτου, μπορεί να ανιχνεύσει άγνωστες επιθέσεις σε συστήματα, γεγονός που αποτελεί ελκυστικό χαρακτηριστικό για ασφάλεια στον κυβερνοχώρο. Η

DL, καθώς και οι γενικοί αλγόριθμοι ML, μπορούν να κατηγοριοποιηθούν στις ακόλουθες τρεις κατηγορίες:

Μάθηση με επίβλεψη: Απαιτεί τη χρήση δεδομένων εισόδου με ετικέτες για την εκπαίδευση μιας ταξινόμησης, ή αλγορίθμου παλινδρόμησης. Η ταξινόμηση χρησιμοποιείται για την ανάλυση δεδομένων και την κατανομή τους σε μια συγκεκριμένη κλάση, ενώ η παλινδρόμηση χρησιμοποιείται για την έξοδο μιας τιμής πρόβλεψης συνεχούς χαρακτήρα.

Μάθηση χωρίς επίβλεψη: Χρησιμοποιεί μη επισημασμένα δεδομένα εισόδου. Αυτή η κατηγορία αλγορίθμων χρησιμοποιείται συχνά για την ομαδοποίηση δεδομένων ή τη μείωση των διαστάσεων τους.

Ενισχυμένη Μάθηση: Βασίζεται σε ένα σύστημα ανταμοιβής που ανταμείβει ικανοποιητικές ενέργειες. Αυτού του είδους οι αλγόριθμοι μπορούν να θεωρηθούν ως συγχώνευση της επιβλεπόμενης και της μη επιβλεπόμενης μάθησης και είναι κυρίως κατάλληλοι για εργασίες που έχουν μακροπρόθεσμες ανατροφοδοτήσεις (Arulkumaran et al., 2017).

2.2. Παραδείγματα εφαρμογών TN σε κρίσιμες υποδομές

Η Τεχνητή Νοημοσύνη (TN) διαδραματίζει σημαντικό ρόλο στην ενίσχυση και τη διαφύλαξη κρίσιμων υποδομών σε διάφορους τομείς. Ακολουθούν μερικά παραδείγματα του τρόπου με τον οποίο εφαρμόζεται η Τεχνητή Νοημοσύνη σε διάφορους τομείς υποδομών κρίσιμης σημασίας:

Ενέργεια

Η Τεχνητή Νοημοσύνη (TN) συμβάλλει σημαντικά στη διαχείριση των δικτύων ηλεκτρικής ενέργειας, προβλέποντας τη ζήτηση, βελτιστοποιώντας τη διανομή ενέργειας και ενσωματώνοντας αποτελεσματικά τις ανανεώσιμες πηγές ενέργειας. Η χρήση της TN στον ενεργειακό τομέα είναι ιδιαίτερα διαδεδομένη, κυρίως για την προστασία των σταθμών παραγωγής ενέργειας, των ηλεκτρικών δικτύων και άλλων ενεργειακών υποδομών από κυβερνοαπειλές. Για παράδειγμα, οι Yu et al. (2021) παρουσιάζουν μια προληπτική προσέγγιση για την αντιμετώπιση προηγμένων επίμονων απειλών (APT) στο Βιομηχανικό Διαδίκτυο των Πραγμάτων (IIoT), χρησιμοποιώντας επισημασμένα δεδομένα από ένα ιδιωτικό δίκτυο ηλεκτρικής ενέργειας. Οι Yang et al. (2019) αναπτύσσουν ένα σύστημα ανίχνευσης εισβολών βασισμένο στη βαθιά μάθηση για δίκτυα

SCADA, χρησιμοποιώντας δεδομένα κίνησης που συλλέγονται από δύο δοκιμαστικά περιβάλλοντα κυβερνοασφάλειας SCADA. Οι Khaw et al. (2020) προτείνουν ένα σύστημα ανίχνευσης κυβερνοεπιθέσεων βασισμένο στη βαθιά μάθηση για τα προστατευτικά συστήματα μεταφοράς relays. Παράλληλα, μια ολοκληρωμένη μελέτη για τις μεθόδους και τις λύσεις της μηχανικής μάθησης στην κυβερνοασφάλεια των έξυπνων δικτύων παρουσιάζεται από τους Haghighi et al. (2020). Επιπλέον, οι Das et al. (2020) προτείνουν ένα μοντέλο ανίχνευσης ανωμαλιών βασισμένο σε κανόνες, ενώ οι Haghighi et al. (2020) αναπτύσσουν ένα τείχος προστασίας βασισμένο σε μηχανική μάθηση και προληπτικούς κανόνες για την ασφάλεια των δικτύων ηλεκτρικής ενέργειας. Η μοντελοποίηση που βασίζεται σε κανόνες μπορεί επίσης να χρησιμοποιηθεί για την παρακολούθηση δραστηριοτήτων χρηστών και προτύπων πρόσβασης, επιτρέποντας τον εντοπισμό εσωτερικών απειλών ή άλλων ύποπτων δραστηριοτήτων και αποτρέποντας τη μη εξουσιοδοτημένη πρόσβαση ή την κλοπή δεδομένων. Συνολικά, οι οργανισμοί μπορούν να ενισχύσουν την ικανότητά τους στον εντοπισμό και τον μετριασμό των κυβερνοαπειλών μέσω της ανάλυσης και της εξόρυξης κανόνων με βάση τη διαφανή μοντελοποίηση.

Μεταφορές

Οι αλγόριθμοι Τεχνητής Νοημοσύνης (TN) επεξεργάζονται δεδομένα από κάμερες, αισθητήρες και συστήματα GPS για τη διαχείριση της κυκλοφοριακής ροής, τη μείωση της συμφόρησης και την ενίσχυση της οδικής ασφάλειας. Επιπλέον, η TN επιτρέπει τη λειτουργία αυτόνομων οχημάτων, τα οποία στο μέλλον θα μπορούσαν να μειώσουν τα ατυχήματα και να βελτιώσουν την αποδοτικότητα της κυκλοφορίας. Ο τομέας των μεταφορών αποτελεί ζωτικό στοιχείο της κρίσιμης υποδομής μιας χώρας και είναι καθοριστικός για την ομαλή λειτουργία της κοινωνίας και της οικονομίας. Ωστόσο, οι συσκευές του Διαδικτύου των Πραγμάτων (IoT) και τα συνδεδεμένα οχήματα στη βιομηχανία μεταφορών μπορεί να είναι ευάλωτα σε κυβερνοεπιθέσεις, θέτοντας σε κίνδυνο τη δημόσια ασφάλεια και προκαλώντας ακόμη και ατυχήματα. Για παράδειγμα, οι επιθέσεις σε συστήματα ελέγχου της κυκλοφορίας μπορεί να αλλοιώσουν τα πρότυπα κυκλοφορίας, προκαλώντας συμφόρηση και δημιουργώντας δυνητικούς κινδύνους για την ασφάλεια.

Οι Lehto et al. (2020) συζητούν τα ζητήματα κυβερνοασφάλειας σε κρίσιμους τομείς των μεταφορών, συμπεριλαμβανομένων της αεροπορίας, της ναυτιλίας και της

αυτοκινητοβιομηχανίας. Οι Argyropoulos et al. (2021) υπογραμμίζουν τη σημασία των τεχνικών TN/ML στον εντοπισμό και τον μετριασμό των απειλών κυβερνοασφάλειας και ιδιωτικότητας σε διάφορες πτυχές του οικοσυστήματος κινητικότητας επόμενης γενιάς. Μια προληπτική προσέγγιση βασισμένη σε κανόνες μπορεί να συμβάλει στην προστασία των κρίσιμων υποδομών της βιομηχανίας μεταφορών από εντεινόμενες κυβερνοαπειλές. Για παράδειγμα, η εξαγωγή ενός συνόλου εξελιγμένων κανόνων και η αντίστοιχη ρύθμιση των συστημάτων ανίχνευσης εισβολών μπορεί να χρησιμοποιηθεί για την ανίχνευση και τον αποκλεισμό ύποπτων δραστηριοτήτων σε πραγματικό χρόνο. Οι τεχνικές μηχανικής μάθησης (ML) μπορούν να εξάγουν κανόνες μέσω της ανίχνευσης ανωμαλιών και της ανάλυσης συμπεριφοράς.

Υγειονομική περίθαλψη

Προηγμένα εργαλεία Τεχνητής Νοημοσύνης (TN) αναλύουν συστηματικά μοτίβα και προβλέπουν την εμφάνιση επιδημιών ασθενειών. Αυτή η δυνατότητα είναι κρίσιμη, καθώς επιτρέπει την έγκαιρη αντίδραση των υγειονομικών αρχών και την ορθή κατανομή ιατρικών και άλλων πόρων, προλαμβάνοντας την εξάπλωση της νόσου και περιορίζοντας τον αντίκτυπό της στη δημόσια υγεία. Η TN διαδραματίζει επίσης σημαντικό ρόλο στη βελτίωση της ανάλυσης ιατρικών εικόνων, όπως οι ακτινογραφίες και οι μαγνητικές τομογραφίες. Χρησιμοποιώντας προηγμένους αλγορίθμους, η TN βελτιώνει την ταχύτητα και την ακρίβεια των ιατρικών διαγνώσεων, προσφέροντας στους ιατρούς ένα ισχυρό εργαλείο για την ακριβέστερη εκτίμηση και αντιμετώπιση των παθήσεων.

Ο τομέας της υγείας αποτελεί ζωτικό μέρος της υποδομής κάθε χώρας, και η ασφάλεια στον κυβερνοχώρο είναι κρίσιμη για τη διασφάλιση της προσβασιμότητας των ιατρικών υπηρεσιών, την αποφυγή δυνητικά καταστροφικών διαταραχών και την προστασία της ιδιωτικής ζωής των ασθενών. Οι Radanliev et al. (2022) διερευνούν την προώθηση της κυβερνοασφάλειας στον τομέα της υγειονομικής περίθαλψης μέσω της χρήσης αυτοβελτιστοποιούμενης και αυτοπροσαρμοζόμενης TN, η οποία μπορεί να διαχειρίζεται συσκευές υγείας άκρων με δεδομένα πραγματικού χρόνου. Οι He et al. (2022) εστιάζουν στην ανακάλυψη και πρόληψη επιθέσεων σε ιατρικά συστήματα βασισμένα σε τεχνολογίες TN. Δεδομένου ότι ο τομέας της υγείας διαχειρίζεται τεράστιους όγκους ευαίσθητων δεδομένων ασθενών, όπως προσωπικές πληροφορίες και ιατρικά αρχεία, η προστασία αυτών των δεδομένων αποτελεί πρωταρχικό στόχο για την αποτροπή παραβιάσεων. Μοντέλα TN που ενσωματώνουν σύνθετους και σχετικούς κανόνες

μπορούν να ανιχνεύουν και να αντιδρούν σε κυβερνοαπειλές, προστατεύοντας ταυτόχρονα την ιδιωτική ζωή των ασθενών.

Διαχείριση νερού

Τα συστήματα Τεχνητής Νοημοσύνης (TN) αναλύουν την ποιότητα του νερού σε πραγματικό χρόνο, αξιοποιώντας δεδομένα από αισθητήρες για την ανίχνευση ρυπαντών και τη βελτιστοποίηση των διαδικασιών επεξεργασίας νερού. Οι αισθητήρες αυτοί συλλέγουν δεδομένα σε πραγματικό χρόνο, παρέχοντας τη δυνατότητα για άμεση ανάλυση και αντίδραση. Η TN μπορεί να αναγνωρίσει ρυπαντές, όπως τοξικά χημικά ή βιολογικούς παράγοντες, και να προσδιορίσει την ένταση και τη φύση της ρύπανσης. Αυτό επιτρέπει την έγκαιρη ενημέρωση και την κατάλληλη ανταπόκριση, οδηγώντας σε πιο αποτελεσματική διαχείριση των διαδικασιών επεξεργασίας νερού. Με αυτόν τον τρόπο, η TN συμβάλλει στη διασφάλιση της ποιότητας του πόσιμου νερού και στην προστασία της δημόσιας υγείας. Αυτή η εφαρμογή της TN είναι κρίσιμη για την αειφόρο διαχείριση των υδατικών πόρων, καθώς επιτρέπει τον αποτελεσματικό και οικονομικά αποδοτικό καθαρισμό και ανακύκλωση του νερού, προωθώντας την υγειονομική ασφάλεια και την περιβαλλοντική προστασία.

Τα έξυπνα συστήματα ύδρευσης αποτελούν πλέον υποδομές ζωτικής σημασίας για κάθε σύγχρονη πόλη. Ωστόσο, οι απειλές στον κυβερνοχώρο μπορούν να επηρεάσουν αυτά τα συστήματα, τα οποία είναι υπεύθυνα για τον έλεγχο της παροχής νερού και της επεξεργασίας λυμάτων. Οι Bello et al. (2020) επισημαίνουν τις προκλήσεις της κυβερνοασφάλειας στη διαχείριση των υποδομών νερού στην Αυστραλία. Οι Bakalos et al. (2019) παρουσιάζουν ένα πλαίσιο για την ανίχνευση επιθέσεων σε κρίσιμες υποδομές νερού, βασισμένο στην πολυτροπική συγχώνευση δεδομένων και την προσαρμοστική βαθιά μάθηση (μοντέλο CNN). Οι Sobien et al. (2023) διερευνούν τη χρήση της TN για την κυβερνοασφάλεια σε συστήματα ύδρευσης, με έμφαση σε διάφορα δοκιμαστικά περιβάλλοντα και ασφαλή επεξεργασία νερού. Η χρήση μοντέλων TN που βασίζονται σε πολυδιάστατους κανόνες μπορεί να αποδειχθεί εξαιρετικά επωφέλης για την ανίχνευση πιθανών παραβιάσεων, τη διασφάλιση της ποιότητας και της ακεραιότητας του νερού, καθώς και την αποτροπή μη εξουσιοδοτημένης πρόσβασης σε κρίσιμα στοιχεία της υποδομής. Για παράδειγμα, οι Das et al. (2020) αναπτύσσουν κανόνες για το σχεδιασμό ενός συστήματος ανίχνευσης ανωμαλιών, χρησιμοποιώντας μετρήσεις αισθητήρων από ένα σύστημα επεξεργασίας νερού. Επιπλέον, η εξόρυξη δεδομένων και οι τεχνικές μηχανικής μάθησης χρησιμοποιούνται για τη

δημιουργία σταθερών κανόνων ανίχνευσης ανωμαλιών σε Συστήματα Βιομηχανικού Ελέγχου (ICS).

Τηλεπικοινωνίες

Η Τεχνητή Νοημοσύνη (TN) εφαρμόζεται για τη βελτίωση της δρομολόγησης και της διαχείρισης της κυκλοφορίας στα τηλεπικοινωνιακά δίκτυα. Χρησιμοποιώντας σύνθετους αλγορίθμους, η TN αναλύει τα δεδομένα κυκλοφορίας και προσαρμόζει αυτόματα τις διαδρομές των δεδομένων, διασφαλίζοντας την ομαλότητα και την αποτελεσματικότητα της επικοινωνίας. Αυτό συμβάλλει στην ενίσχυση της αξιοπιστίας και στη μείωση των διακοπών στις υπηρεσίες, καθιστώντας τα δίκτυα πιο αποδοτικά και ανθεκτικά σε προβλήματα. Επιπλέον, οι αλγόριθμοι TN αναλύουν μεγάλους όγκους δεδομένων για να εντοπίσουν ασυνήθιστα μοτίβα ή συμπεριφορές που μπορεί να υποδηλώνουν δόλιες δραστηριότητες. Αυτή η δυνατότητα είναι ιδιαίτερα σημαντική για την προστασία των χρηστών και των παρόχων από απάτες και άλλες παραβιάσεις ασφαλείας. Η πρόληψη και η άμεση αντίδραση σε αυτές τις απειλές ενισχύουν την ασφάλεια των δικτύων και την εμπιστοσύνη των χρηστών προς τους παρόχους υπηρεσιών.

Χρηματοοικονομικές υπηρεσίες

Τα μοντέλα TN χρησιμοποιούνται για να αξιολογήσουν τα επίπεδα κινδύνου σε δάνεια και επενδύσεις. Αυτά τα μοντέλα αναλύουν μεγάλους όγκους δεδομένων για να εντοπίσουν παράγοντες κινδύνου, όπως την πιστωτική ικανότητα του δανειολήπτη, τις αγοραστικές τάσεις και τις οικονομικές συνθήκες. Η χρήση της TN επιτρέπει μια πιο ακριβή και τεκμηριωμένη αξιολόγηση κινδύνου, η οποία με τη σειρά της βοηθά τους αναλυτές και τους επενδυτές να προβαίνουν σε πιο ενημερωμένες αποφάσεις και να διαχειρίζονται καλύτερα τους πόρους τους. Επιπλέον, τα συστήματα διαπραγμάτευσης που βασίζονται στην TN αξιοποιούνται για την ανάλυση των δεδομένων της αγοράς και για τη λήψη αποφάσεων συναλλαγών υψηλής ταχύτητας. Αυτοί οι αλγόριθμοι είναι σε θέση να ανταποκρίνονται σε αλλαγές της αγοράς σε κλάσματα δευτερολέπτου, πράγμα που ξεπερνά τις δυνατότητες των ανθρώπινων εμπόρων. Η αλγοριθμική διαπραγμάτευση βελτιώνει την αποδοτικότητα και την απόδοση των συναλλαγών, καθώς και μειώνει το κόστος των συναλλαγών, δημιουργώντας σημαντικό ανταγωνιστικό πλεονέκτημα για εκείνους που χρησιμοποιούν την τεχνολογία.

Υπηρεσίες έκτακτης ανάγκης

Τα εργαλεία TN εφαρμόζονται για την ανάλυση δορυφορικών εικόνων και δεδομένων από κοινωνικά μέσα για να χαρτογραφήσουν περιοχές που έχουν πληγεί από φυσικές καταστροφές όπως πλημμύρες, πυρκαγιές, σεισμούς ή τυφόνες. Η άμεση και ακριβής χαρτογράφηση είναι κρίσιμη, καθώς επιτρέπει στις αρχές και στους φορείς αντίδρασης να οργανώσουν και να ιεραρχήσουν τις προσπάθειες διάσωσης και αντιμετώπισης, εστιάζοντας πόρους εκεί όπου είναι περισσότερο απαραίτητοι.

Οι αλγόριθμοι TN χρησιμοποιούνται για την πρόβλεψη της ζήτησης για υπηρεσίες έκτακτης ανάγκης και τη βελτιστοποίηση της κατανομής των πόρων και του προσωπικού. Αυτό περιλαμβάνει τη διαχείριση των διαθέσιμων ιατρικών αποθεμάτων, των τροφίμων, του νερού και άλλων σημαντικών προμηθειών, καθώς και την κατανομή του διασωστικού προσωπικού και του εξοπλισμού. Η χρήση της TN σε αυτό τον τομέα στοχεύει στην πιο αποδοτική και αποτελεσματική ανταπόκριση στις καταστάσεις κρίσης, βελτιώνοντας την γενική ετοιμότητα και ανταπόκριση σε έκτακτες ανάγκες. Οι εφαρμογές TN ενδυναμώνουν τις υπηρεσίες έκτακτης ανάγκης, βελτιώνοντας τόσο την ικανότητα ως προς την πρόβλεψη και την αντίδραση σε κρίσεις όσο και την αποδοτικότητα στη διαχείριση και κατανομή πόρων κατά τη διάρκεια κρίσεων.

Αμυντική βιομηχανία

Οι επιθέσεις στον κυβερνοχώρο εναντίον αμυντικών συστημάτων μπορεί να έχουν καταστροφικές συνέπειες, κυμαίνονται από τη διακοπή των στρατιωτικών επιχειρήσεων έως την υπονόμευση της εθνικής ασφάλειας. Συνεπώς, η ανάγκη για ανθεκτικότητα στον κυβερνοχώρο πρέπει να ενσωματωθεί σε κάθε στάδιο του κύκλου ζωής και της ανάπτυξης των αμυντικών δυνατοτήτων, όπου η τεχνητή νοημοσύνη μπορεί να συμβάλει σημαντικά. Για παράδειγμα, οι Shin et al. (2018), αναπτύσσουν μια μέθοδο για την αποτελεσματική λειτουργία του κυβερνοχώρου ISR (νοημοσύνη, επιτήρηση και αναγνώριση) μέσα σε ένα κλειστό στρατιωτικό δίκτυο, χρησιμοποιώντας τεχνικές σταδιακής μάθησης. Οι Shin et al. (2018), εξετάζουν επίσης τη διαδικασία του κυβερνο-ISR με έμφαση στην αποτελεσματική λήψη αποφάσεων βάσει επιλογών χαρακτηριστικών. Επιπλέον, οι Eom et al. (2012), παρουσιάζουν μια ισχυρή στρατηγική για την επιχειρησιακή υπεροχή στον κυβερνοπόλεμο. Η εξαγωγή κανόνων ασφαλείας και η υλοποίηση αντίστοιχων μοντέλων τεχνητής νοημοσύνης μπορεί να διαδραματίσει κρίσιμο ρόλο στην υπεράσπιση της κρίσιμης αμυντικής βιομηχανίας. Για παράδειγμα, η εφαρμογή κανόνων φιλτραρίσματος ηλεκτρονικού ταχυδρομείου μπορεί να αποκλείσει γνωστούς τομείς phishing και

να προλάβει την είσοδο κακόβουλων συνημμένων στα συστήματα, μειώνοντας τον κίνδυνο επιθέσεων phishing.

2.3. Εφαρμογή ενός ML αλγορίθμου σε ένα σύστημα κρίσιμων υποδομών, επιλεγμένο με βάση κατάλληλα κριτήρια

Η μηχανική μάθηση (ML) είναι ένας ταχέως αναπτυσσόμενος τομέας της μηχανικής νοημοσύνης που επιτρέπει στους υπολογιστές να μαθαίνουν από δεδομένα χωρίς να προγραμματίζονται ρητά. Τεχνικές ML, όπως η βαθιά μάθηση, έχουν αποδειχθεί αποτελεσματικές στη μοντελοποίηση πολύπλοκων συστημάτων και στην αποκάλυψη κρυφών μοτίβων που ο άνθρωπος συχνά αγνοεί. (Xu & Saleh, 2021) ορίζει τη μηχανική μάθηση ως "ένα σύνολο μεθόδων που μπορούν να ανιχνεύουν αυτόματα μοτίβα σε δεδομένα και στη συνέχεια να χρησιμοποιούν τα αποκαλυφθέντα μοτίβα για να προβλέπουν μελλοντικά δεδομένα ή να εκτελούν άλλους τύπους λήψης αποφάσεων υπό αβεβαιότητα". Τα τελευταία χρόνια, η μηχανική μάθηση έχει ανθίσει σε ένα ευρύ φάσμα τομέων, από την υγειονομική περίθαλψη έως την εκπαίδευση, τις μεταφορές, το ηλεκτρονικό εμπόριο, τη μηχανική ποιότητας και αξιοπιστίας και πολλούς άλλους. Ο τομέας της μηχανικής ανθεκτικότητας των συστημάτων κρίσιμων υποδομών (CIS) αναμένεται να ακολουθήσει αυτή την τάση. Πρόσφατες φυσικές καταστροφές έχουν αναδείξει τη σημασία των ανθεκτικών υποδομών στη σύγχρονη κοινωνία (Zio, 2016). Οι πόλεις και τα έθνη σε όλο τον κόσμο επενδύουν πλέον σημαντικά σε προηγμένα συστήματα υποδομών, όπως οι έξυπνες πόλεις, τα έξυπνα δίκτυα και τα έξυπνα κτίρια, για να βελτιώσουν την ανθεκτικότητά τους και να αυξήσουν τη βιωσιμότητά τους.

Ωστόσο, η ανάπτυξη και συντήρηση αυτών των συστημάτων υψηλής τεχνολογίας είναι δαπανηρή και υπάρχει επιτακτική ανάγκη για καινοτόμες λύσεις που μπορούν να βελτιώσουν την ανθεκτικότητα αυτών των πολύπλοκων συστημάτων υποδομής χωρίς να απαιτούν εκτεταμένες επενδύσεις κεφαλαίου. Η χρήση μεθόδων μηχανικής μάθησης μπορεί να συμβάλει στην κάλυψη αυτής της ανάγκης. Υπάρχουν διάφορες εφαρμογές της μηχανικής μάθησης στον τομέα της μηχανικής ανθεκτικότητας κρίσιμων υποδομών που μπορούν να βελτιώσουν την ανθεκτικότητα των υποδομών τόσο των νέων όσο και των υφιστάμενων συστημάτων υποδομής. Ορισμένα από τα βασικά οφέλη της χρήσης τεχνικών μηχανικής μάθησης για τη βελτίωση της ανθεκτικότητας περιλαμβάνουν βελτιωμένες δυνατότητες λήψης αποφάσεων και μειωμένες κοινωνικοοικονομικές απώλειες από διαταραχές.

Οι κρίσιμες υποδομές ή τα δίκτυα κρίσιμων υποδομών είναι τα συστήματα που στηρίζουν την παροχή ζωτικών δημόσιων υπηρεσιών, είτε αυτά ανήκουν στον ιδιωτικό είτε στον δημόσιο τομέα. Παραδείγματα αποτελούν τα συστήματα μεταφορών, τα δίκτυα ηλεκτρικής ενέργειας, τα δίκτυα τηλεπικοινωνιών και τα νοσοκομεία. Οποιαδήποτε διαταραχή αυτών των συστημάτων μπορεί να έχει σημαντικές επιπτώσεις στην οικονομία και την κοινωνία. Λόγω της κρίσιμης φύσης αυτών των συστημάτων υποδομής, οι κυβερνήσεις σε όλο τον κόσμο έχουν θεσπίσει ρυθμιστικά πλαίσια σχεδιασμένα να διασφαλίζουν ότι μπορούν να παραμείνουν σε αποδεκτά επίπεδα εξυπηρέτησης ακόμη και σε περιόδους διαταραχών (Zhang & Peeta, 2011). Αυτά τα πλαίσια περιλαμβάνουν νομοθεσία που διέπει το σχεδιασμό των κρίσιμων συστημάτων και την ανάγκη διατήρησης επαρκούς πλεονασμού για να είναι δυνατή η επισκευή και συντήρηση σε περιόδους βλάβης του συστήματος. Η μη διατήρηση αυτών των επιπέδων ανθεκτικότητας μπορεί να οδηγήσει σε αποτυχίες στην παροχή υπηρεσιών που θα μπορούσαν να έχουν σημαντικό αντίκτυπο στην οικονομία και, σε ορισμένες περιπτώσεις, στην υγεία και την ασφάλεια του κοινού.

Παρά τις σημαντικές προσπάθειες που καταβάλλουν οι κυβερνήσεις και η βιομηχανία για τη βελτίωση της ανθεκτικότητας των υποδομών ζωτικής σημασίας, οι διαταραχές εξακολουθούν να συμβαίνουν σε τακτική βάση. Αυτό οφείλεται σε διάφορους παράγοντες, όπως φυσικές καταστροφές, ακραία καιρικά φαινόμενα, επιθέσεις στον κυβερνοχώρο, τρομοκρατία και άλλα ανίκητα φυσικά ή ανθρωπογενή φαινόμενα.

Η ανάλυση της τρωτότητας και του κινδύνου των υποδομών ζωτικής σημασίας είναι μια σύνθετη πρόκληση που πρέπει να αντιμετωπιστεί για τη βελτίωση της ανθεκτικότητας των εν λόγω συστημάτων. Πολλές υποδομές ζωτικής σημασίας είναι πολύπλοκα συστήματα, όπου οι ανθρώπινες και τεχνικές αλληλεπιδράσεις είναι συχνά περίπλοκα συνυφασμένες. Αυτή η πολυπλοκότητα, σε συνδυασμό με τη δυναμική φύση του περιβάλλοντος, μπορεί να οδηγήσει σε σημαντικά κενά στην κατανόηση των τρωτών σημείων και των κινδύνων που ενέχουν αυτές οι υποδομές. Επιπλέον, λόγω της αλληλοεξαρτώμενης φύσης πολλών υποδομών ζωτικής σημασίας, μια μεμονωμένη αστοχία μπορεί να έχει αλυσιδωτές επιπτώσεις σε πολλαπλά συστήματα και τομείς. Ως εκ τούτου, η αποτελεσματική ανάλυση τρωτότητας και κινδύνων απαιτεί μια ολοκληρωμένη προσέγγιση που λαμβάνει υπόψη τις αλληλεπιδράσεις μεταξύ διαφορετικών στοιχείων και συστημάτων. Αυτές οι αλληλεπιδράσεις ή αλληλεξαρτήσεις περιλαμβάνουν λειτουργικές, χωρικές, κυβερνο-, κοινωνικές και λογικές αλληλεξαρτήσεις και άλλες μορφές

αλληλεξάρτησης που θα μπορούσαν να επηρεάσουν την προληπτική ανθεκτικότητα (π.χ. ασφάλεια) και την αντιδραστική ανθεκτικότητα (π.χ. σχεδιασμός αποκατάστασης και ανάκαμψης). Αν και αυτή η σύνθετη μορφή αλληλεξάρτησης των υποδομών αναγνωρίστηκε πριν από δύο δεκαετίες, η σχετική έρευνα για τα αλληλεξαρτώμενα δίκτυα ξεκίνησε πολύ πρόσφατα ως μέρος της εφαρμογής της μηχανικής ανθεκτικότητας (Jahn, 2020).

Η μοντελοποίηση και προσομοίωση των αλληλεξαρτώμενων συστημάτων κρίσιμων υποδομών (ICIS) είναι ένα σχετικά νέο και αναπτυσσόμενο ερευνητικό πεδίο, διαπίστωσε ότι υπάρχουν πολλές προκλήσεις και ευκαιρίες που σχετίζονται με αυτό το πεδίο σε μια ανασκόπηση της βιβλιογραφίας. Μία από τις σημαντικότερες προκλήσεις είναι η πολυπλοκότητα των αλληλεξαρτήσεων μεταξύ των στοιχείων των ICIS. Οι Alkhaleel et al., (2022) υπογραμμίζουν τη σημασία των ασφαλών και ανθεκτικών αλληλεξαρτώμενων υποδομών, καθώς τα συστήματα αυτά χαρακτηρίζονται από συνδεδεμένα και αλληλοεξαρτώμενα δίκτυα, όπου, όταν ένα μόνο δίκτυο τίθεται σε κίνδυνο, ολόκληρο το σύστημα κινδυνεύει. Επιπλέον, οι Wang et al., (2021), υποστηρίζουν ότι η κατανόηση του τρόπου με τον οποίο αυτά τα αλληλεξαρτώμενα συστήματα αλληλεπιδρούν και λειτουργούν από κοινού είναι το κλειδί για τη διασφάλιση του ασφαλούς και ανθεκτικού χαρακτήρα τους. Επιπλέον, οι Korkmaz & Correia (2019), διαπίστωσαν ότι η έλλειψη δεδομένων, η ποικιλομορφία των στοιχείων του συστήματος και τα διαφορετικά επίπεδα αβεβαιότητας που σχετίζονται με τα αλληλεξαρτώμενα συστήματα συμβάλλουν στην πολυπλοκότητα της μοντελοποίησης και προσομοίωσης των ICISs για εφαρμογές ανθεκτικότητας.

Στην πραγματικότητα, είναι αξιοσημείωτο ότι απαιτείται η ανάπτυξη νέων εργαλείων, στρατηγικών και μεθοδολογιών (π.χ. μηχανική μάθηση) για την αποτελεσματική μοντελοποίηση, προσομοίωση και βελτίωση της ανθεκτικότητας των ICIS. Είναι απαραίτητο να αναπτυχθούν πιο αποτελεσματικοί και αποδοτικοί τρόποι συλλογής και ανάλυσης δεδομένων που σχετίζονται με τα στοιχεία των ICISs και η μελλοντική έρευνα θα πρέπει να επικεντρωθεί στις προκλήσεις και τις ευκαιρίες που σχετίζονται με αυτόν τον τομέα.

Τα δίκτυα υποδομών δεν είναι απομονωμένα μεταξύ τους, αλλά μάλλον εξαρτώνται το ένα από το άλλο με διαφορετικούς τρόπους για την εύρυθμη λειτουργία τους. Ως εκ τούτου, παρουσιάζουν αλληλεξάρτηση, όπου ένα ζεύγος δικτύων υποδομής λέγεται ότι είναι αλληλοεξαρτώμενο εάν υπάρχει μια αμφίδρομη σχέση μεταξύ τους μέσω της οποίας η κατάσταση κάθε υποδομής εξαρτάται από την κατάσταση της άλλης. Οι αλληλεξαρτήσεις διαδραματίζουν κρίσιμο ρόλο στην

ανθεκτικότητα των ΚΠ, όχι μόνο συμβάλλοντας στην ευρεία διάδοση της αποτυχίας (π.χ. αλυσιδωτές αποτυχίες), αλλά και διευκολύνοντας ή περιπλέκοντας την όλη διαδικασία αποκατάστασης. Η ανάγκη περιγραφής των σχέσεων μεταξύ των συστημάτων υποδομής και της αντίστοιχης διάδοσης των διαταραχών των συστημάτων οδήγησε στον ορισμό διαφόρων ταξινομήσεων της φύσης των αλληλεξαρτήσεων των υποδομών. Ωστόσο, η ταξινόμηση του μπορεί να περιγραφεί ως μια "αυτοτελής ταξινόμηση" που μπορεί να αποτυπώσει τη διαφορετική φύση των αλληλεξαρτήσεων (Tizghadam et al., 2019).

Ένα σύστημα ελέγχου (ΣΕ) είναι μια συσκευή ή ένα σύνολο συσκευών που εκτελεί τη διαχείριση και τη ρύθμιση της συμπεριφοράς άλλων συσκευών ή συστημάτων. Παραδείγματα τέτοιων είναι τα συστήματα εποπτικού ελέγχου και συλλογής δεδομένων (SCADA) και τα κατανεμημένα συστήματα ελέγχου. (DCS). Τα CS αναπτύσσονται σε πολλούς τύπους περιβαλλόντων, αλλά όταν χρησιμεύουν ως βοήθεια σε υποδομές απαραίτητες για την ευημερία της κοινωνίας, θεωρούνται κρίσιμα συστήματα ελέγχου και η καλή λειτουργία τους είναι υψίστης σημασίας.

Για την προστασία αυτών των συστημάτων, διάφοροι οργανισμοί έχουν αναπτύξει κατευθυντήριες γραμμές προστασίας για τις κρίσιμες υποδομές (ΚΥ) (COM, 2011), οι οποίες εκδηλώνουν την ανάγκη προσφοράς προηγμένων λύσεων άμυνας για τις ΚΥ στον τομέα της πρόληψης. Μία από τους κύριες πυλώνες αυτής της έρευνας είναι οι δυναμικές λύσεις πρόληψης, όπως τα συστήματα ανίχνευσης εισβολών (IDS) (Scarfone & Mell, 2012), ικανά να συμπληρώνουν την πρόληψη με αυτοματοποιημένη δράση σε σενάρια κρίσης. Παραδοσιακά, τα IDS σχεδιάστηκαν για δίκτυα γενικής χρήσης και η εφαρμογή τους για το CIP δεν είναι πάντα επαρκής λόγω της παρουσίας αυστηρών απαιτήσεων και πρωτοκόλλων επικοινωνίας ιδιοκτησίας στα ΚΠ. Ωστόσο, η ανάγκη εισαγωγής τέτοιων στοιχείων σε κρίσιμα περιβάλλοντα έχει την υγιή υποστήριξη της επιστημονικής κοινότητας (Chertoff, 2008) και των ιδρυμάτων. Επιπλέον, σύμφωνα με τους Alcaraz & Lopez (2013), υπάρχει μεγάλη ανάγκη για την παροχή αυτών των IDS με ευφυΐα και αυτόματες δυνατότητες, ώστε να μπορούν να ανταποκρίνονται γρήγορα και αποτελεσματικά σε καταστάσεις έκτακτης ανάγκης, ιδίως σε απομονωμένα σενάρια.

Οι τεχνικές μάθησης ποικίλλουν και προέρχονται από πολύ διαφορετικά πεδία γνώσης (π.χ. βελτιστοποίηση, στατιστική, λογική κ.λπ.). Είναι ενδιαφέρον να μελετήσουμε αυτές τις μεθόδους λαμβάνοντας υπόψη χαρακτηριστικά που επιβάλλουν περιορισμούς στο υποκείμενο σύστημα και επηρεάζουν τη δυνατότητα εισαγωγής τους στο πλαίσιο του ΚΠΑ. Είναι σημαντικό να συζητηθεί

το σχήμα γνώσης και το επίπεδο εποπτείας του συστήματος (Burbeck & Nadjm-Tehrani, 2007). Το σχήμα γνώσης υποδεικνύει το επίπεδο γνώσης που τροφοδοτείται στο σύστημα πριν από την εκπαίδευση: τα συστήματα που βασίζονται σε προηγούμενη γνώση τροφοδοτούνται με τη γνώση και την εμπειρία ενός εμπειρογνώμονα, τα συστήματα χωρίς προηγούμενη γνώση βασίζονται στη γνώση που εξάγεται μέσω μιας αυτόματης (ή ημιαυτόματης) διαδικασίας εκπαίδευσης, τα υβριδικά συστήματα βασισμένα στη γνώση συστημάτων προσθέτουν τη γνώση ενός εμπειρογνώμονα στο μοντέλο του συστήματος που προκύπτει μέσω εκπαίδευσης. Το επίπεδο εποπτείας του συστήματος μπορεί να διαχωριστεί σε: επιβλεπόμενη μάθηση, όπου το σύστημα έχει γνώση για τις μεταβλητές που μαθαίνει, και μη επιβλεπόμενη μάθηση, όπου δεν παρέχεται καμία γνώση στο σύστημα κατά την εκπαίδευσή του (Burbeck & Nadjm-Tehrani, 2007).

Ανάλογα με το επίπεδο εποπτείας ή εμπλοκής του χειριστή στη διαδικασία μάθησης, είναι δυνατόν να κατηγοριοποιηθούν οι κύριες τεχνικές ML που θα μπορούσαν να οδηγήσουν στο να χρησιμοποιηθούν στο πλαίσιο της ανίχνευσης εισβολών για CIP σε εποπτευόμενες και μη εποπτευόμενες μεθόδους μάθησης. Στην πρώτη κατηγορία, οι κύριες προσεγγίσεις είναι οι αλγόριθμοι που βασίζονται στη λογική, όπως τα δέντρα αποφάσεων, οι μαθητευόμενοι κανόνες- οι αλγόριθμοι στατιστικής μάθησης, όπως τα Μπεϋζιανά δίκτυα, ο Naïve Bayes και οι μαθητευόμενοι με βάση το παράδειγμα- και τα τεχνητά νευρωνικά δίκτυα, όπως οι τεχνικές που βασίζονται σε perceptron. Όσον αφορά τις μεθόδους μάθησης χωρίς επίβλεψη, οι κύριες τεχνικές είναι η μάθηση κανόνων συσχέτισης (π.χ. αλγόριθμος Apriori και αλγόριθμος FP-growth), τεχνικές ομαδοποίησης (π.χ. ο k-means), και οι αλυσίδες Markov.

Οι τεχνικές εκμάθησης μπορούν να εφαρμοστούν σε διάφορα συστήματα που παρέχουν προστασία για CI, ιδίως σε λύσεις IDS (Witten et al., 2006). Οι Dussel et al. (2010), παρουσιάζουν ένα IDS δικτύου με βάση το ωφέλιμο φορτίο και με βάση τις ανωμαλίες για CIs, ικανό- να παρακολουθεί την κυκλοφορία σε πραγματικό χρόνο. Το IDS κάνει χρήση διαφορετικών τεχνικών, το σύστημα εξάγει τις πληροφορίες με τη μορφή διανυσμάτων, υπολογίζει τα μέτρα απόστασης ομοιότητας και τα συγκρίνει με ένα προηγούμενως διδαγμένο μοντέλο κανονικότητας, υποδεικνύοντας την παρουσία ή την απουσία μιας ανωμαλίας με την εμφάνιση συναγερμών. Οι Roosta et al., (2008) εισάγουν ένα IDS με βάση τις ανωμαλίες για ασύρματα συστήματα ελέγχου διεργασιών. Το IDS που παρουσιάζεται είναι θεωρητικό, όπου η ανίχνευση βασίζεται σε κανόνες πολιτικής σχεδιασμένους από ειδικούς. Οι Yang et al. (2006), παρουσιάζουν ένα IDS που

βασίζεται στην αντιστοίχιση προτύπων, ικανό να ανιχνεύει ανωμαλίες αναλύοντας την απόκλιση από την κανονική συμπεριφορά. Χρησιμοποιούν αυτοσυσχετική παλινδρόμηση μοντέλα πυρήνα και ένα διαδοχικό τεστ αναλογίας πιθανοτήτων για να διακρίνουν μια επίθεση από την κανονική συμπεριφορά. Το MELISSA (Hadiosmanovic et al., 2011) είναι ένα IDS σημασιολογικού επιπέδου βασισμένο σε FP-Trees (Witten et al., 2011) που αναζητά ανεπιθύμητες ενέργειες του χρήστη μέσω της επεξεργασίας των αρχείων καταγραφής στο κέντρο ελέγχου SCADA. Οι Carcano et al. (2010), προτείνουν ένα IDS βασισμένο στην κατάσταση που χρησιμοποιεί κανόνες για την ανίχνευση σύνθετων σεναρίων επίθεσης με βάση αλυσίδες παράνομων πακέτων δικτύου.

Το Autoscopy (Reeves et al., 2012), είναι ένα IDS ικανό να ανιχνεύει κακόβουλο λογισμικό που προσπαθεί να "καταλάβει" δείκτες και ρουτίνες του συστήματος. Η συμπεριφορά του συστήματος και κατά τη διάρκεια της λειτουργίας του, χρησιμοποιεί στατιστικά στοιχεία για να διακρίνει ανώμαλες συμπεριφορές και ενεργοποιεί συναγερμούς. Οι D'Antonio et al. (2006), παρουσιάζουν ένα IDS που υλοποιεί έναν κανόνα μαθητή για την ταξινόμηση των τιμών σε κανονικά δεδομένα και δεδομένα επίθεσης, διαθέτει ένα στοιχείο παρακολούθησης ροής που εξάγει στατιστικές σχέσεις μεταξύ διαφορετικών συνόδων για να βελτιώσει το μοντέλο που μαθαίνει. Οι Cheung et al. (2007), προτείνουν ένα IDS τριών επιπέδων που βασίζεται σε μοντέλα και πρότυπα του συστήματος που έχει σχεδιαστεί από έναν εμπειρογνώμονα. Ένα από τα στρώματα εφαρμόζει μια μονάδα Bayesian learning για την ανίχνευση αλλαγών στη διαθεσιμότητα του επιτηρούμενου συστήματος. Οι Lin et al. (2013), προτείνουν ένα IDS βασισμένο σε προδιαγραφές, το οποίο χρησιμοποιεί τις τυπικές προδιαγραφές του υπό επιτήρηση συστήματος για να επαληθεύσει τη σωστή χρήση των πακέτων δικτύου. Οι Raciti et al. (2013) ένα IDS για έξυπνα δίκτυα που βασίζεται σε τεχνικές ομαδοποίησης, για την ανίχνευση ανωμαλιών στο κυβερνοχώρο και στο φυσικό επίπεδο.

Είναι ενδιαφέρον να παρατηρήσουμε ότι τα σημερινά συστήματα εφαρμόζουν μόνο παθητικές μεθόδους αντίδρασης, συνήθως απευθύνουν προειδοποιήσεις στους χειριστές και πρέπει να εκτελούν χειροκίνητα τις επιθεωρήσεις, τις επισκευές των συστημάτων και να βοηθούν σε καταστάσεις κρίσης. (π.χ. αιχμές τάσης σε πυλώνες, υψηλές πιέσεις σε φράγματα κ.λπ.). Το κύριο μειονέκτημα των μη αυτόματων συστημάτων είναι ότι η βοήθεια μπορεί να φτάσει πολύ αργά, και οι βλάβες του συστήματος μπορεί να μεταδοθούν σε άλλα εξαρτώμενα συστήματα, συμπεριλαμβανομένων άλλων αλληλεξαρτώμενων ΚΠ, προκαλώντας κάθε είδους καταστροφή.

Συνεπώς, θεωρούμε ζωτικής σημασίας τη λήψη αποτελεσματικών μέτρων, για την αποφυγή της πιθανής κοινωνικής και οικονομικής ζημίας που προκύπτει από μια αλυσιδωτή αποτυχία. Πρέπει να υπάρχουν αυτόματες ενεργές διαδικασίες αντίδρασης, ίσως βασισμένες σε τεχνικές ML, οι οποίες είναι ικανές να παρέχουν αποτελεσματικά αντίμετρα μπροστά σε κάθε είδους ανωμαλία ή επίθεση.

2.4. Ανάλυση των υπαρχόντων εργαλείων TN που χρησιμοποιούνται σε αυτόν τον τομέα

Οι μέθοδοι τεχνητής νοημοσύνης (AI) επιτρέπουν στις μηχανές να μαθαίνουν και να συμπεραίνουν από μεγάλο όγκο δεδομένων (Ertel, 2017). Καθώς τα συστήματα υποδομής γίνονται όλο και πιο διασυνδεδεμένα, πολύπλοκα και ψηφιοποιημένα, η τεχνητή νοημοσύνη θα είναι ζωτικής σημασίας για την παροχή και τη διατήρηση υπηρεσιών που όλο και περισσότεροι αριθμός ανθρώπων εξαρτάται καθημερινά από αυτές (Luckey et al., 2021). Ωστόσο, καθώς το ενδιαφέρον για την τεχνητή νοημοσύνη συνεχίζει να αυξάνεται, η έρευνα για την εφαρμογή της σε συστήματα υποδομής παραμένει σε μεγάλο βαθμό απομονωμένη.

Ο Άλαν Τούρινγκ πρότεινε το "τεστ Τούρινγκ" για να προσφέρει έναν λειτουργικό ορισμό της τεχνητής νοημοσύνης, δηλώνοντας ότι ένα πραγματικά ευφυές σύστημα πρέπει να είναι ικανό να ταιριάζει με τις ανθρώπινες γνωστικές επιδόσεις σε βαθμό που ένας άνθρωπος ανακριτής δεν μπορεί να διακρίνει τη διαφορά μεταξύ ανθρώπου και μηχανής. κατά την αλληλεπίδραση μέσω ενός συστήματος τηλετύπου (Russell and Norvig, 2002). Σε μια "ολική" δοκιμή Turing, η αντίληψη και οι φυσικές ικανότητες εξετάζονται παράλληλα με τη γνωστική λειτουργία. Κάθε ένα από τα ακόλουθα στοιχεία αντιπροσωπεύει ένα τομέα της TN που συμβάλλει στην επίτευξη ενός ή περισσότερων από τους στόχους του τεστ Turing:

- αναπαράσταση γνώσης, για την αποθήκευση δεδομένων
- αυτοματοποιημένη συλλογιστική, για την εξαγωγή και χρήση συμπερασμάτων από τα αποθηκευμένα δεδομένα
- μηχανική μάθηση, για τον εντοπισμό προτύπων και την τροποποίηση της συμπεριφοράς
- όραση υπολογιστών, για την αντίληψη του περιβάλλοντος

- ρομποτική, για την αλληλεπίδραση με το φυσικό περιβάλλον
- επεξεργασία φυσικής γλώσσας, για την επικοινωνία στην ανθρώπινη γλώσσα

Ενώ αυτά τα στοιχεία μπορούν να περιγράψουν χαρακτηριστικά ενός ιδανικού ευφυούς συστήματος υπολογιστή, μπορούν επίσης να θεωρηθούν θέματα στον τομέα της Τεχνητής Νοημοσύνης

έρευνας, καθένα από τα οποία ασχολείται με τεχνικές που συμβάλλουν σε ένα στοιχείο της νοημοσύνης του συστήματος.

Υπάρχει σημαντική αλληλεπικάλυψη μεταξύ των πεδίων, με την αυτοματοποιημένη συλλογιστική να εξαρτάται εγγενώς από τη βάση γνώσης στην οποία στηρίζεται, καθώς και με τις τεχνικές μηχανικής μάθησης - ιδίως τα συνελκτικά νευρωνικά δίκτυα - που χρησιμοποιούνται όλο και περισσότερο στα συστήματα όρασης υπολογιστών. Τα εν λόγω συστήματα όρασης συχνά ενσωματώνονται σε ευφυή ρομπότ. Μοντέλα που συνδυάζουν τόσο το στοιχείο της συλλογιστικής όσο και της μηχανικής μάθησης, όπως τα προσαρμοστικά νευρο-ασαφή συστήματα συμπερασμού (ANFIS), βρίσκουν επίσης αυξανόμενη χρήση. Ως τα πιο ευρέως υιοθετημένα πεδία της Τεχνητής Νοημοσύνης (TN) στην έρευνα για τις υποδομές, η μηχανική μάθηση και οι μέθοδοι υπολογιστικής όρασης εξετάζονται παρακάτω. Οι μέθοδοι που περιγράφονται στην παρούσα ενότητα δεν αποτελούν μια εξαντλητική ανασκόπηση όλων των τεχνικών μηχανικής μάθησης και υπολογιστικής όρασης, αλλά επικεντρώνονται στις πιο συχνά χρησιμοποιούμενες μεθόδους στα ερευνητικά έργα που εξετάστηκαν, προκειμένου να παρέχουν πλαίσιο για περαιτέρω συζήτηση.

Για κρίσιμα συστήματα, τα μοντέλα TN μπορούν να χρησιμοποιηθούν σε συλλογισμό γνώσης, σχεδιασμό, επεξεργασία φυσικής γλώσσας, υπολογιστική όραση, ρομποτική και τεχνητή γενική νοημοσύνη, καθιστώντας τα ιδιαίτερα κατάλληλα για ένα ευρύ φάσμα κρίσιμων εφαρμογών. Η TN μπορεί να συμβάλει στη βελτιστοποίηση της απόδοσης, στην ασφαλή λειτουργία, στην πρόληψη και ανίχνευση απάτης και εισβολής, καθώς και σε προγνωστικά μοντέλα για την υγεία και στην ανάλυση μετά την αποτυχία. Τυπικοί τομείς εφαρμογής για κρίσιμα συστήματα περιλαμβάνουν την υγειονομική περίθαλψη, τις μεταφορές, τις υποδομές πληροφορικής και υπηρεσιών κοινής ωφέλειας, το εμπόριο και τα οικονομικά. Ο Πίνακας 1 παρουσιάζει τους τομείς εφαρμογών για υποδομές ζωτικής σημασίας, προσαρμοσμένους από μελέτη του Κογκρέσου των

ΗΠΑ για την αναγνώριση υποδομών ζωτικής σημασίας [Moteff], μαζί με παραδείγματα συστημάτων και τυπικές χρήσεις της TN.

Application Domain	Example Systems	Typical AI Uses
telecommunication infrastructure	public telephone network, local branch exchange	ID, FA
water supply systems	water treatment plant, dam control	FA, FS, PHM
electrical power systems	nuclear power plant, regional electrical grid	FA, FS, PHM, PO
oil and gas generation and distribution	gas pipeline, gas powered power plant	FA, FS, PHM, PO
roadway transportation systems	Smart interstate highway, traffic monitoring and control	FA, FS, PHM, PO
railway transportation systems	high speed train line, metropolitan train network control	FA, FS, PHM, PO
air transportation systems	air traffic control system network, passenger aircraft autopilot	FA, FS, PHM, PO
banking and financial services	pension fund management, stock market management	FD, ID
public safety services	air passenger screening, police dispatch	FD, ID, PO
healthcare systems	robotic surgery, healthcare record management	FD, ID, PO
administration and public services	employee personnel database, retirement management	FD, ID

Figure 1 Πίνακας 1.

Η Ανάλυση Αποτυχίας (Failure Analysis, FA) καθορίζεται από σαφώς ορισμένους κανόνες και πρωτόκολλα. Οι μηχανές που βασίζονται στην Τεχνητή Νοημοσύνη (AI) χρησιμοποιούν τη γνώση που προκύπτει από περιστατικά (συλλογή αποδεικτικών στοιχείων) για να εξάγουν νέες πληροφορίες και να προτείνουν τα επόμενα βήματα. Η ανίχνευση απάτης (Fraud Detection, FD) και εισβολής (Intrusion Detection, ID) αντιμετωπίζεται με τη χρήση μεθόδων ανίχνευσης ανωμαλιών. Σήμερα, αυτές οι μέθοδοι είναι συνήθως είτε βασισμένες σε στατιστικές προσεγγίσεις είτε σε τεχνικές μηχανικής μάθησης. Κάθε μοντέλο ανίχνευσης ανωμαλιών λαμβάνει αποφάσεις με βάση έναν περιορισμένο αριθμό πηγών πληροφόρησης. Στο μέλλον, τα συστήματα θα βασίζονται σε μεθόδους βαθιάς μάθησης, όπως οι αυτοκωδικοποιητές, και θα μπορούν να εντοπίζουν ανωμαλίες αναλύοντας ταυτόχρονα μια μεγάλη ποικιλία σημάτων εισόδου.

Οι λειτουργίες ασφαλούς αποτυχίας (Failsafe, FS) θα επιτυγχάνονται με τη χρήση συνόλων μοντέλων TN, διασφαλίζοντας ότι, σε περίπτωση αποτυχίας ενός μοντέλου, τα υπόλοιπα θα συνεχίζουν να λαμβάνουν αποφάσεις. Αυτά τα μοντέλα θα ελέγχονται από ευφυείς πράκτορες, οι οποίοι θα επικοινωνούν μεταξύ τους σύμφωνα με συγκεκριμένα πρωτόκολλα για να συντονίζουν τις αποφάσεις τους. Για παράδειγμα, εάν ένας πράκτορας διαπιστώσει ότι δεν μπορεί να λάβει αξιόπιστη απόφαση λόγω απρόβλεπτων συνθηκών, θα ζητήσει βοήθεια από άλλους πράκτορες, αναζητώντας τις αποφάσεις τους για την τρέχουσα κατάσταση. Τα προγνωστικά συστήματα και η διαχείριση υγείας (Prognostics and Health Management, PHM) θα επιτυγχάνονται μέσω ενός αριθμού μοντέλων TN που αναπτύσσονται σε αντίστοιχες συσκευές ή συστήματα, τα οποία θα

παρακολουθούν συνεχώς τη λειτουργική τηλεμετρία για να προβλέψουν ποια στοιχεία είναι πιθανό να αποτύχουν εντός του επόμενου λεπτού, ώρας, ημέρας κ.λπ.

Η βελτιστοποίηση απόδοσης (Performance Optimization, PO) των συστημάτων θα επιτευχθεί μέσω της δημιουργίας «Ψηφιακών Διδύμων» (Digital Twins). Πρόκειται για μοντέλα που βασίζονται σε Τεχνητή Νοημοσύνη (AI), όπως τα βαθιά νευρωνικά δίκτυα, τα οποία αναπαριστούν την απόκριση ενός συστήματος σε σήματα εισόδου. Μέσω της τροφοδότησης διαφόρων σημάτων και συνδυασμών τους, θα είναι δυνατό να καθοριστεί η βέλτιστη διαμόρφωση που ελαχιστοποιεί συγκεκριμένα κριτήρια. Ωστόσο, η χρήση της TN σε κρίσιμα συστήματα δεν περιορίζεται απαραίτητα σε αυτούς τους τομείς. Άλλα πεδία εφαρμογής κρίσιμων συστημάτων, όπου η αστοχία θα μπορούσε να οδηγήσει σε τραυματισμό ή ακόμα και θάνατο, περιλαμβάνουν καταναλωτικά συστήματα (π.χ. μηχανήματα αυτόματης πώλησης ζεστών φαγητών, έξυπνα συστήματα πλυσίματος αυτοκινήτων), ψυχαγωγία (π.χ. ηλεκτρονικά παιχνίδια) και οικιακές εφαρμογές (π.χ. έξυπνα σπίτια, έξυπνες συσκευές).

2.5. Αξιολόγηση εργαλείων TN που χρησιμοποιούν αλγορίθμους Μηχανικής Μάθησης (Machine Learning- ML) για τη διαχείριση κρίσιμων υποδομών

Η Τεχνητή Νοημοσύνη (TN) επιτρέπει την προγνωστική ανάλυση, την αναγνώριση προτύπων και την ανίχνευση απειλών σε πραγματικό χρόνο, αξιοποιώντας αλγόριθμους μηχανικής μάθησης για την ανάλυση μεγάλου όγκου δεδομένων και εντοπισμό μοτίβων και ανωμαλιών που μπορεί να υποδηλώνουν πιθανή κυβερνοεπίθεση. Με την αυτοματοποίηση αυτών των διαδικασιών, η TN μπορεί να παρέχει στους οργανισμούς δυνατότητες ανίχνευσης απειλών σε πραγματικό χρόνο, προγνωστικής ανάλυσης και συνεχούς βελτίωσης, επιτρέποντάς τους να παραμείνουν μπροστά από την καμπύλη όσον αφορά την ασφάλεια στον κυβερνοχώρο. Ωστόσο, τα οφέλη και οι περιορισμοί της τεχνητής νοημοσύνης σε αυτούς τους τομείς εξαρτώνται από την ποιότητα και την ακρίβεια των δεδομένων που χρησιμοποιούνται για την εκπαίδευση των μοντέλων TN, καθώς και από τη διαφάνεια και την ερμηνεία των αλγορίθμων που χρησιμοποιούνται (Ramos et al., 2020). Επιπλέον, η εξάρτηση από τα συστήματα TN μπορεί να οδηγήσει σε υπερβολική εξάρτηση από την τεχνολογία και παραμέληση άλλων κρίσιμων μέτρων ασφαλείας, όπως η εκπαίδευση των εργαζομένων και ο σχεδιασμός αντιμετώπισης περιστατικών.

Η αξιολόγηση των εργαλείων πληροφορικής, τα οποία ενσωματώνουν αλγορίθμους μηχανικής μάθησης (ML) για τη διαχείριση κρίσιμων υποδομών, αποτελεί ένα σημαντικό ζήτημα στον τομέα

της τεχνολογίας. Κρίσιμες υποδομές, όπως τα ενεργειακά και τηλεπικοινωνιακά δίκτυα, τα συστήματα ύδρευσης και τα μεταφορικά συστήματα, απαιτούν πολύπλοκη και αξιόπιστη διαχείριση λόγω της μεγάλης τους σημασίας για την κοινωνία και την οικονομία. Τα εργαλεία ML διευκολύνουν την προληπτική συντήρηση μέσω της ανίχνευσης προτύπων και ανωμαλιών στα δεδομένα λειτουργίας των συστημάτων. Αυτό επιτρέπει την πρόβλεψη βλαβών πριν αυτές συμβούν, μειώνοντας έτσι το κόστος και τον χρόνο αδράνειας. Οι αλγόριθμοι ML ενισχύουν την ασφάλεια των κρίσιμων συστημάτων μέσω προηγμένων τεχνικών ανίχνευσης εισβολών και αντίδρασης σε κυβερνοεπιθέσεις. Η διαρκής εκπαίδευση και αναβάθμιση των αλγορίθμων εξασφαλίζει την προσαρμογή τους σε νέες απειλές. Επιπλέον, η χρήση ML διευκολύνει την αυτοματοποίηση των ρουτίνων διαδικασιών και την ταχύτερη λήψη αποφάσεων μέσω της ανάλυσης μεγάλων όγκων δεδομένων, βελτιώνοντας την αποδοτικότητα και περιορίζοντας τα λάθη από ανθρώπινο παράγοντα.

Είναι σημαντικό τα συστήματα μηχανικής μάθησης (ML) να είναι διαφανή και να παρέχουν εξηγήσεις για τις αποφάσεις που λαμβάνουν, ιδιαίτερα σε κρίσιμες εφαρμογές όπου απαιτείται υψηλό επίπεδο ευθύνης και λογοδοσίας. Η αποδοτικότητα των αλγορίθμων ML εξαρτάται άμεσα από την ποιότητα, την πληρότητα και την ακρίβεια των δεδομένων που χρησιμοποιούνται για την εκπαίδευσή τους. Παράλληλα, η ασφάλεια των ίδιων των ML συστημάτων πρέπει να είναι εξασφαλισμένη, καθώς αυτά είναι ευάλωτα σε κυβερνοεπιθέσεις. Επιπροσθέτως, η χρήση της ML οφείλει να συμμορφώνεται με ηθικά πρότυπα και νομικές διατάξεις, ειδικά σε τομείς που επηρεάζουν τη δημόσια ασφάλεια και τα προσωπικά δεδομένα. Η αξιολόγηση και η ενσωμάτωση της ML στη διαχείριση κρίσιμων υποδομών πρέπει να πραγματοποιηθεί με ιδιαίτερη προσοχή, ενσωματώνοντας συνεχή παρακολούθηση και αξιολόγηση της αποδοτικότητας και της ασφάλειας των συστημάτων, για να διασφαλιστεί η βέλτιστη απόδοση και προστασία.

Για την αποτελεσματική χρήση της τεχνητής νοημοσύνης για την πρόβλεψη και την πρόληψη κυβερνοεπιθέσεων σε συστήματα υποδομών ζωτικής σημασίας, συνιστάται η υιοθέτηση βέλτιστων πρακτικών, όπως η συνεχής παρακολούθηση και δοκιμή των μοντέλων τεχνητής νοημοσύνης, ενσωμάτωση λύσεων TN με άλλα μέτρα ασφαλείας, όπως τείχη προστασίας, ανίχνευση εισβολής και μηχανισμοί ελέγχου πρόσβασης, και η καθιέρωση σαφών διαδικασιών για την αντιμετώπιση περιστατικών και την υποβολή εκθέσεων. Οι οργανισμοί θα πρέπει επίσης να διασφαλίζουν τη διαφάνεια και την ερμηνεία των AI αλγορίθμων, να επανεξετάζουν και να

ενημερώνουν τακτικά τις πολιτικές ασφαλείας τους και να παρέχουν συνεχείς υπηρεσίες στους εργαζόμενους και εκπαίδευση για να διασφαλιστεί ότι το προσωπικό κατανοεί πώς να χρησιμοποιεί και να ανταποκρίνεται σε εργαλεία ασφαλείας που βασίζονται στην TN (Al Dabbagh et al., 2021). Επιπλέον, οι οργανισμοί θα πρέπει να αξιολογούν τακτικά την αποτελεσματικότητα των λύσεων τεχνητής νοημοσύνης τους και να εργάζονται για την αντιμετώπιση τυχόν εντοπισμένων αδυναμιών. Ακολουθώντας αυτές τις βέλτιστες πρακτικές, οι οργανισμοί μπορούν να μεγιστοποιήσουν τα δυνητικά οφέλη της TN όσον αφορά τη βελτίωση της ασφάλειας στον κυβερνοχώρο και τη μείωση του κινδύνου κυβερνοεπιθέσεων στα συστήματα κρίσιμων υποδομών τους.

Οι οργανισμοί μπορούν να αξιοποιήσουν αποτελεσματικά την Τεχνητή Νοημοσύνη για να βελτιώσουν τις συνολικές ικανότητές τους για την αντιμετώπιση κρίσεων, χρησιμοποιώντας προγνωστικές αναλύσεις με βάση την Τεχνητή Νοημοσύνη για τον εντοπισμό πιθανών σεναρίων κρίσης και χρησιμοποιώντας μηχανικές αλγορίθμων μηχανικής μάθησης για την ανάλυση ιστορικών δεδομένων και τον εντοπισμό μοτίβων και τάσεων που μπορεί να υποδηλώνουν μια επικείμενη κρίση (Chang et al., 2021). Η τεχνητή νοημοσύνη μπορεί επίσης να χρησιμοποιηθεί για την αυτοματοποίηση των διαδικασιών αντιμετώπισης κρίσεων και την πραγματοποίηση αποφάσεων σε πραγματικό χρόνο με βάση τεράστιες ποσότητες δεδομένων και πληροφοριών. Ωστόσο, οι οργανισμοί ενδέχεται να αντιμετωπίσουν προκλήσεις στην αποτελεσματική αξιοποίηση της TN για την αντιμετώπιση κρίσεων, συμπεριλαμβανομένης της ανάγκης για υψηλής ποιότητας δεδομένων και πληροφοριών, η διαφάνεια και η ερμηνεία των αλγορίθμων TN και ο κίνδυνος υπερβολικής εξάρτησης από την τεχνολογία. Επιπλέον, οι οργανισμοί ενδέχεται να δυσκολεύονται να ενσωματώσουν τα συστήματα αντιμετώπισης κρίσεων που βασίζονται στην TN στις υπάρχουσες διαδικασίες και να διασφαλίσουν ότι το προσωπικό θα εκπαιδευτεί για τον τρόπο χρήσης και αντιμετώπισης αυτών των συστημάτων. Με την αντιμετώπιση αυτών των προκλήσεων, οι οργανισμοί μπορούν να αξιοποιήσουν αποτελεσματικά την TN για να βελτιώσουν τις συνολικές τους δυνατότητες αντιμετώπισης κρίσεων και να προετοιμαστούν καλύτερα για πιθανά σενάρια κρίσης και να ανταποκριθούν σε αυτά.

Κεφάλαιο 3: Θέματα ασφάλειας σχετικά με τη TN σε κρίσιμες υποδομές

3.1. TN στην επιθετική ασφάλεια

Η TN μπορεί να χρησιμοποιηθεί για την ανάπτυξη προηγμένων στρατιωτικών συστημάτων και όπλων, δημιουργώντας νέες απειλές για την εθνική και παγκόσμια ασφάλεια. Τα αυτόνομα όπλα, για παράδειγμα, μπορούν να λειτουργούν χωρίς ανθρώπινη επίβλεψη, εγείροντας ανησυχίες για την ηθική και τη νομιμότητα της χρήσης τους σε ένοπλες συγκρούσεις. Αν και η Τεχνητή Νοημοσύνη προσφέρει τεράστιες δυνατότητες και οφέλη, είναι απαραίτητο να αναγνωρίσουμε και να αντιμετωπίσουμε τις απειλές που τη συνοδεύουν. Οι οργανισμοί, οι κυβερνήσεις και η κοινωνία στο σύνολό της πρέπει να συνεργαστούν για την ανάπτυξη κανονισμών, πολιτικών και τεχνολογικών λύσεων που θα μετριάσουν αυτούς τους κινδύνους και θα διασφαλίσουν ότι η TN χρησιμοποιείται υπεύθυνα και ηθικά (Kaloudi & Li, 2020).

Η εξέλιξη του τομέα της επιθετικής ασφάλειας έχει προχωρήσει σημαντικά. Αρχικά, η εστίαση ήταν κυρίως σε απλές σαρώσεις ευπαθειών, αλλά πλέον περιλαμβάνει προηγμένες τεχνικές που προσομοιώνουν πραγματικές κυβερνοεπιθέσεις. Αυτή η μετατόπιση αναδεικνύει την αυξανόμενη πολυπλοκότητα των κυβερνοαπειλών και το γεγονός ότι οι επιτιθέμενοι γίνονται πιο εξελιγμένοι, χρησιμοποιώντας προηγμένες τεχνικές και τακτικές. Στο παρελθόν, οι προσπάθειες ασφάλειας επικεντρώνονταν κυρίως στην ανταπόκριση σε περιστατικά μετά την εμφάνισή τους. Ωστόσο, με προσεγγίσεις όπως το "red teaming" και το ηθικό hacking, υπάρχει πλέον έμφαση στην προληπτική αντιμετώπιση αυτών των ζητημάτων, υιοθετώντας την οπτική του επιτιθέμενου. Αυτή η στάση συμβάλλει στην ενίσχυση των αμυντικών μας μηχανισμών έναντι των κυβερνοαπειλών.

Η πρόιμη ιστορία του penetration testing ανάγεται στη δεκαετία του 1960, όταν αναδείχθηκαν οι πρώτες ανησυχίες σχετικά με την ασφάλεια των υπολογιστικών επικοινωνιών. Κυβερνήσεις και επιχειρήσεις άρχισαν να σχηματίζουν ομάδες για να δοκιμάσουν και να εντοπίσουν τρωτά σημεία στα δίκτυά τους, λειτουργώντας ως άμυνα έναντι πιθανών επιθέσεων. Σημαντικές συνεισφορές σε αυτόν τον τομέα έγιναν από πρωτοπόρους όπως ο James P. Anderson, ο οποίος ανέπτυξε μεθοδολογίες που εξακολουθούν να χρησιμοποιούνται σήμερα, όπως αναφέρεται στο Anderson Report Institute [2019].

Επιπλέον, η ενσωμάτωση του συνεχούς penetration testing με συστήματα όπως το SIEM (Security Incident Event Management) επαναπροσδιορίζει τον τρόπο με τον οποίο οι ευπάθειες εντοπίζονται

και αντιμετωπίζονται, προσφέροντας μια πιο ομαλή προσέγγιση στην κυβερνοασφάλεια. Αυτή η ενσωμάτωση συμβάλλει στην αυτοματοποίηση των διαδικασιών απόκρισης και αποκατάστασης, ενισχύοντας την αποδοτικότητα των ομάδων ασφαλείας και μειώνοντας τον χρόνο που απαιτείται για την αντιμετώπιση των ευπαθειών Security [2022].

Η τεχνητή νοημοσύνη ασκεί σημαντική επίδραση στον τομέα της επιθετικής ασφάλειας, μετασχηματίζοντας τον τρόπο με τον οποίο οι ειδικοί εντοπίζουν και εκμεταλλεύονται τα αδύναμα σημεία σε συστήματα και δίκτυα. Η επιθετική ασφάλεια, η οποία περιλαμβάνει προληπτικές μεθόδους όπως η δοκιμή διείσδυσης και οι ομάδες "red teaming", αποσκοπεί στον εντοπισμό αδυναμιών πριν αυτές αξιοποιηθούν από κακόβουλους παράγοντες. Η ΤΝ ενισχύει αυτές τις προσπάθειες μέσω της αυτοματοποίησης εργασιών, της αύξησης της αποδοτικότητας και της αποκάλυψης σύνθετων ευπαθειών που ενδέχεται να παραβλέπονται από παραδοσιακές μεθόδους.

Μία βασική εφαρμογή της ΤΝ στην επιθετική ασφάλεια είναι η αυτοματοποιημένη σάρωση για αδυναμίες. Τα εργαλεία που βασίζονται στην ΤΝ μπορούν να εκτελούν συστηματικές σαρώσεις σε δίκτυα και εφαρμογές, αξιοποιώντας αλγορίθμους μηχανικής μάθησης για τον εντοπισμό προτύπων και ανωμαλιών που υποδηλώνουν προβλήματα ασφαλείας. Αυτή η αυτοματοποίηση μειώνει τον χρόνο και τους πόρους που απαιτούνται για τη χειροκίνητη σάρωση, ενώ αυξάνει την πιθανότητα εντοπισμού κρυφών ευπαθειών.

Επιπλέον, η ΤΝ συμβάλλει στην ανάπτυξη προηγμένων τεχνικών εκμετάλλευσης ευπαθειών. Μέσω της ανάλυσης διαφόρων διανυσμάτων επίθεσης έναντι συγκεκριμένων αμυντικών μηχανισμών, τα μοντέλα μηχανικής μάθησης μπορούν να προβλέψουν τους πιο αποτελεσματικούς τρόπους παραβίασης ενός συστήματος. Αυτό επιτρέπει στους επαγγελματίες της ασφάλειας να προσαρμόζουν τις στρατηγικές τους με μεγαλύτερη ακρίβεια, βελτιώνοντας την αποτελεσματικότητα των δοκιμών διείσδυσης (Mirsky et al., 2023).

Στον τομέα της κοινωνικής μηχανικής, η ΤΝ αξιοποιεί την Επεξεργασία Φυσικής Γλώσσας (NLP) για τη δημιουργία εξαιρετικά πειστικών μηνυμάτων ηλεκτρονικού "ψαρέματος" και άλλων απατηλών επικοινωνιών. Αναλύοντας δημόσια διαθέσιμα δεδομένα, η ΑΙ μπορεί να προσωποποιήσει μηνύματα, αυξάνοντας την πιθανότητα εξαπάτησης των στόχων. Αυτή η αυξημένη πολυπλοκότητα αναδεικνύει την ανάγκη για ενισχυμένη εκπαίδευση στην ευαισθητοποίηση σχετικά με την ασφάλεια, προκειμένου να αντιμετωπιστούν οι ολοένα και πιο

εξελιγμένες τακτικές κοινωνικής μηχανικής. Η TN μπορεί να δημιουργήσει δυναμικό κακόβουλο λογισμικό που προσαρμόζεται σε πραγματικό χρόνο για να αποφύγει τον εντοπισμό. Τέτοιο λογισμικό είναι ικανό να τροποποιεί τον κώδικα ή τη συμπεριφορά του όταν αντιμετωπίζει μέτρα ασφαλείας, καθιστώντας πιο δύσκολη την ανίχνευση και την εξουδετέρωσή του από παραδοσιακές λύσεις αντιιών.

Η TN επιτρέπει επίσης την προσομοίωση σύνθετων σεναρίων επίθεσης που μιμούνται Προηγμένες Επίμονες Απειλές (APTs). Μέσω αυτών των προσομοιώσεων, οι οργανισμοί μπορούν να κατανοήσουν βαθύτερα τον τρόπο με τον οποίο τέτοιες επιθέσεις μπορεί να εξελιχθούν και να αναπτύξουν πιο αποτελεσματικές αμυντικές στρατηγικές.

Ωστόσο, η χρήση της TN στην επιθετική ασφάλεια εγείρει σημαντικά ηθικά ζητήματα και κινδύνους διπλής χρήσης. Τα ίδια εργαλεία και τεχνικές που ενισχύουν την ασφάλεια μπορούν να χρησιμοποιηθούν κακόβουλα για την πραγματοποίηση κυβερνοεπιθέσεων. Είναι απαραίτητο οι επαγγελματίες της ασφάλειας να τηρούν ηθικές κατευθυντήριες γραμμές και νομικούς κανονισμούς, διασφαλίζοντας ότι η TN χρησιμοποιείται υπεύθυνα. Η αυξανόμενη ενσωμάτωση της στην επιθετική ασφάλεια αναδεικνύει επίσης τη σημασία της αξιοποίησής της σε αμυντικές στρατηγικές. Οι αμυντικοί μηχανισμοί που βασίζονται στην TN μπορούν να εντοπίζουν και να ανταποκρίνονται σε απειλές πιο άμεσα, αναλύοντας πρότυπα και προβλέποντας πιθανές επιθέσεις, ενισχύοντας έτσι τη συνολική ασφάλεια ενός οργανισμού. Η συνεργασία μεταξύ ειδικών στην TN και επαγγελματιών της ασφάλειας είναι κρίσιμη. Διεπιστημονικές ομάδες και εκπαιδευτικά προγράμματα μπορούν να γεφυρώσουν το χάσμα μεταξύ της τεχνολογίας AI και των πρακτικών εφαρμογών ασφαλείας, προωθώντας την καινοτομία ενώ διατηρούν υψηλά ηθικά πρότυπα.

Μια κοινή προσπάθεια της Microsoft και της OpenAI αποκάλυψε ότι οι απειλητικοί παράγοντες αξιοποιούν ενεργά την TN για να ενισχύσουν τις επιχειρήσεις τους. Η κλίμακα, η ταχύτητα και η πολυπλοκότητα των κυβερνοεπιθέσεων έχουν αυξηθεί παράλληλα με την ταχεία πρόοδο της TN. Οι επαγγελματίες της ασφάλειας πρέπει να επιδιώκουν να παραμένουν ένα βήμα μπροστά στη μάχη κατά των κυβερνοαπειλών, κατανοώντας και αντιμετωπίζοντας τον τρόπο με τον οποίο οι κακόβουλοι παράγοντες χρησιμοποιούν την TN (Malatji & Tolah, 2024).

Η TN ενισχύει την επιθετική ασφάλεια μέσω της προσομοίωσης προηγμένων κυβερνοεπιθέσεων, επιτρέποντας στους ειδικούς της ασφάλειας να εντοπίζουν και να αντιμετωπίζουν ευαλωτότητες

προτού τις εκμεταλλευτούν οι κακόβουλοι παράγοντες. Επιπλέον, οι δυνατότητες της TN μπορούν να βελτιστοποιήσουν τις διαδικασίες σάρωσης, να αυτοματοποιήσουν την αναγνώριση, να παράγουν ολοκληρωμένες αναφορές κυβερνοασφάλειας και ακόμη και να εκμεταλλευτούν αυτόνομα αδυναμίες για τη δοκιμή της ανθεκτικότητας του συστήματος.

Η εφαρμογή της TN στην επιθετική ασφάλεια βελτιώνει την κλιμακωσιμότητα, ενισχύει την αποδοτικότητα, αποκαλύπτει πιο σύνθετες ευαλωτότητες και, τελικά, ενισχύει τη συνολική αμυντική στάση. Ωστόσο, παρά τα οφέλη, καμία μεμονωμένη λύση TN δεν επαρκεί από μόνη της. Οι οργανισμοί πρέπει να ενθαρρύνουν ένα περιβάλλον μάθησης και ανάπτυξης, όπου τα μέλη της ομάδας μπορούν να πειραματίζονται με διάφορα εργαλεία TN για την εξεύρεση αποτελεσματικών λύσεων.

3.2. TN στην αμυντική ασφάλεια

Η χρήση TN σε κρίσιμες υποδομές ενέχει τον κίνδυνο κυβερνοεπιθέσεων, οι οποίες μπορεί να εκμεταλλευτούν τις αδυναμίες των αλγορίθμων ή των δεδομένων που χρησιμοποιούνται για τη λειτουργία των συστημάτων. Ειδικά σε συστήματα όπως τα ενεργειακά δίκτυα και οι υπηρεσίες κοινής ωφέλειας, η παραβίαση αυτών των αλγορίθμων μπορεί να οδηγήσει σε εκτεταμένες διακοπές ή ακόμα και σε φυσικές καταστροφές. Η προστασία από τέτοιες απειλές απαιτεί την εφαρμογή προληπτικών μέτρων, όπως η συνεχής παρακολούθηση των συστημάτων, η χρήση προηγμένων τεχνικών ανίχνευσης ανωμαλιών και η ανάπτυξη ασφαλών αλγορίθμων που μπορούν να αντέξουν σε κακόβουλες επιθέσεις.

Η Τεχνητή Νοημοσύνη (TN) για την αμυντική ασφάλεια αποτελεί έναν σύγχρονο και εξελισσόμενο τομέα που χρησιμοποιεί προηγμένους αλγορίθμους και τεχνικές για την αναγνώριση, εκμετάλλευση και διείσδυση σε ευάλωτα συστήματα και δίκτυα. Σε αντίθεση με την παραδοσιακή άμυνα, η αμυντική ασφάλεια επικεντρώνεται στην ανάπτυξη επιθέσεων για να εντοπίσει και να διορθώσει αδυναμίες προτού αυτές εκμεταλλευτούν από κακόβουλους παράγοντες. Η χρήση της TN στην επιθετική ασφάλεια επιτρέπει την αυτοματοποίηση και τη βελτίωση της αποτελεσματικότητας των επιθετικών εργαλείων και τεχνικών (Kaloudi & Li, 2020).

Ένα από τα κύρια πλεονεκτήματα της TN στην αμυντική ασφάλεια είναι η ικανότητά της να ανιχνεύει αδυναμίες και να αναγνωρίζει μοτίβα που μπορεί να διαφύγουν από την παραδοσιακή ανάλυση. Χρησιμοποιώντας μηχανική μάθηση και βαθιά μάθηση, τα συστήματα TN μπορούν να

αναλύσουν τεράστιες ποσότητες δεδομένων για να εντοπίσουν ανωμαλίες. Για παράδειγμα, μπορούν να εξετάσουν τον πηγαίο κώδικα λογισμικού, να αναγνωρίσουν ευπαθή σημεία και να προτείνουν τρόπους εκμετάλλευσης αυτών των ευπαθειών, επιτρέποντας στους ειδικούς ασφαλείας να αναπτύξουν διορθωτικά μέτρα. Επιπλέον, η TN μπορεί να χρησιμοποιηθεί για την ανάπτυξη προηγμένων εργαλείων αμυντικής ασφάλειας που προσομοιώνουν σύνθετες επιθέσεις. Αυτά τα εργαλεία μπορούν να μιμηθούν τη συμπεριφορά κακόβουλων παραγόντων, δοκιμάζοντας την ανθεκτικότητα των συστημάτων ασφαλείας σε επιθέσεις τύπου DDoS, phishing, και zero-day. Χρησιμοποιώντας τεχνικές όπως η δημιουργία κακόβουλου λογισμικού με βάση την TN, οι επιτιθέμενοι μπορούν να δημιουργήσουν πιο αποτελεσματικά και δυσκολότερα ανιχνεύσιμα κακόβουλα προγράμματα.

Η TN μπορεί επίσης να βελτιώσει την αποτελεσματικότητα των εργαλείων ανίχνευσης και απόκρισης. Χρησιμοποιώντας αλγορίθμους μηχανικής μάθησης, τα συστήματα TN μπορούν να ανιχνεύσουν και να αντιδράσουν σε επιθέσεις σε πραγματικό χρόνο, προσαρμόζοντας τις τακτικές τους βάσει των ανιχνευόμενων μοτίβων. Αυτό επιτρέπει τη δυναμική ανάλυση και την προσαρμογή της στρατηγικής επιθετικής ασφάλειας για την αντιμετώπιση των πιο εξελιγμένων απειλών.

Καθώς οι επιτιθέμενοι εξελίσσουν τις τακτικές τους, τα παραδοσιακά συστήματα ασφαλείας που βασίζονται σε κανόνες αντιμετωπίζουν αυξανόμενες προκλήσεις στην προσπάθεια τους να συμβαδίσουν. Η αμυντική Τεχνητή Νοημοσύνη (TN) ενσωματώνει μια σειρά από λειτουργίες και χαρακτηριστικά που έχουν σχεδιαστεί για να ενισχύσουν τα μέτρα κυβερνοασφάλειας.

Ικανότητες ανίχνευσης απειλών

Η αμυντική TN χρησιμοποιεί ανίχνευση ανωμαλιών, όπου μοντέλα μηχανικής μάθησης εντοπίζουν αποκλίσεις από τη συνήθη συμπεριφορά του δικτύου, επισημαίνοντας αποτελεσματικά ύποπτες δραστηριότητες, όπως ασυνήθιστα μοτίβα εισόδου ή μεταφορές δεδομένων, ακόμα και σε περιπτώσεις άγνωστων επιθέσεων.

Προγνωστική ανάλυση

Μέσω της προγνωστικής ανάλυσης, η TN αξιοποιεί ιστορικά δεδομένα για να προβλέψει πιθανά κενά ασφαλείας, επιτρέποντας στους οργανισμούς να ενισχύσουν προληπτικά τις άμυνές τους πριν από την εκδήλωση μιας επίθεσης.

Ανάλυση συμπεριφοράς

Η ανάλυση συμπεριφοράς περιλαμβάνει την παρακολούθηση της συμπεριφοράς των χρηστών και των οντοτήτων από την TN, προκαλώντας ειδοποιήσεις για ενέργειες που αποκλίνουν από τα αναμενόμενα μοτίβα.

Γρήγορη απόκριση σε απειλές

Με την ταχεία απόκριση σε απειλές, η TN μπορεί να αυτοματοποιήσει ενέργειες όπως την απομόνωση συμβιβασμένων συσκευών, το μπλοκάρισμα κακόβουλης κίνησης ή την ειδοποίηση των ανθρώπινων ομάδων ασφαλείας. Η ενσωμάτωση αυτών των δυνατοτήτων της TN στα συστήματα κυβερνοασφάλειας μπορεί να ενισχύσει σημαντικά την ικανότητα ενός οργανισμού να ανιχνεύει, να προβλέπει και να ανταποκρίνεται αποτελεσματικά σε κυβερνοαπειλές, δημιουργώντας ένα πιο ανθεκτικό και ασφαλές περιβάλλον δικτύου.

Τα Συστήματα Ανίχνευσης Εισβολών (IDS) που βασίζονται στην Τεχνητή Νοημοσύνη (TN) αναλύουν τα μοτίβα κυκλοφορίας δικτύου για να ανιχνεύσουν ανωμαλίες, συμπεριλαμβανομένων των ακανόνιστων πακέτων TCP/IP, των επιθέσεων Άρνησης Υπηρεσίας (DOS) και των σαρών θυρών. Επίσης, είναι ικανά να παρακολουθούν την επικοινωνία εντός οχημάτων, όπως το δίκτυο Controller Area Network (CAN), για να εντοπίζουν απειλές όπως η έγχυση δεδομένων αισθητήρων ή οι επιθέσεις πλημμύρας στον δίαυλο σε συνδεδεμένα οχήματα. Επιπλέον, μέσω της παρακολούθησης της απόδοσης του υλικού, η αμυντική TN εντοπίζει απειλές όπως οι Spectre και Meltdown.

Πληροφορίες απειλών

Μέσω εκτεταμένης συλλογής δεδομένων από πηγές ασφαλείας, φόρουμ και αναφορές συμβάντων, η TN εντοπίζει αποτελεσματικά νέες απειλές στην κυβερνοασφάλεια. Επιπλέον, αυτοματοποιεί την αξιολόγηση των απειλών κατηγοριοποιώντας γνωστές απειλές ενώ αναγνωρίζει νέες βάσει μοναδικών προτύπων, ενισχύοντας τα προληπτικά μέτρα άμυνας.

Επεξεργασία Φυσικής Γλώσσας (NLP)

Η αμυντική TN εξετάζει το περιεχόμενο των email, τα συνημμένα αρχεία και τη συμπεριφορά των αποστολέων για να εντοπίσει προσπάθειες phishing και κακόβουλους συνδέσμους, ενισχύοντας την ασφάλεια των email. Ταυτόχρονα, τα chatbots που βασίζονται στην TN αναγνωρίζουν και

μετριάζουν επιβλαβές περιεχόμενο, προστατεύοντας τους χρήστες από διαδικτυακές απειλές κατά την επικοινωνία σε πραγματικό χρόνο.

Παρακολούθηση σκοτεινού δικτύου

Η παρακολούθηση του σκοτεινού δικτύου περιλαμβάνει ενεργή σάρωση κρυφών διαδικτυακών δικτύων για να ανιχνεύσει τυχόν αναφορές σε ευαίσθητα δεδομένα ή διαπιστευτήρια ενός οργανισμού. Έτσι, οι απειλές εντοπίζονται νωρίτερα, σηματοδοτώντας πιθανές παραβιάσεις πριν αξιοποιηθούν.

Ωστόσο, η χρήση της TN για αμυντική ασφάλεια συνοδεύεται από ηθικές και νομικές προκλήσεις. Η ανάπτυξη και η χρήση εργαλείων επιθετικής ασφάλειας πρέπει να γίνεται με υπευθυνότητα, διασφαλίζοντας ότι χρησιμοποιούνται για την προστασία και όχι για την εκμετάλλευση των συστημάτων. Οι ειδικοί ασφαλείας πρέπει να τηρούν αυστηρά ηθικά πρότυπα και νομικές κατευθυντήριες γραμμές για να αποτρέψουν την κατάχρηση αυτών των τεχνολογιών. Η Τεχνητή Νοημοσύνη προσφέρει σημαντικές δυνατότητες για την αμυντική ασφάλεια, παρέχοντας προηγμένα εργαλεία και τεχνικές για την ανίχνευση, την εκμετάλλευση και την διείσδυση σε ευάλωτα συστήματα. Η αυτοματοποίηση και η βελτίωση της αποτελεσματικότητας των επιθετικών εργαλείων μέσω της TN μπορεί να ενισχύσει την ασφάλεια των συστημάτων, επιτρέποντας την προληπτική διόρθωση αδυναμιών. Παρά τα οφέλη αυτά, η υπεύθυνη χρήση της TN για αμυντική ασφάλεια είναι απαραίτητη για να διασφαλιστεί ότι οι τεχνολογίες αυτές χρησιμοποιούνται για το καλό και όχι για κακόβουλους σκοπούς.

3.3.Τρωτά σημεία στην TN

Ενώ η TN προσφέρει σημαντικά οφέλη, παρουσιάζει επίσης τρωτά σημεία που μπορούν να εκμεταλλευτούν κακόβουλοι παράγοντες, οδηγώντας σε αρνητικές συνέπειες. Η κατανόηση αυτών των ευπαθειών είναι κρίσιμη για την ανάπτυξη ανθεκτικών συστημάτων TN και τη διασφάλιση της ασφαλούς εφαρμογής τους.

Οι επιθέσεις αντιπάλων περιλαμβάνουν τη χειραγώγηση των δεδομένων εισόδου με σκοπό να εξαπατηθούν τα μοντέλα TN και να παράγουν λανθασμένες προβλέψεις ή ταξινομήσεις. Με την εισαγωγή ανεπαίσθητων διαταραχών στα δεδομένα εισόδου—συχνά αόρατων στον άνθρωπο—οι επιτιθέμενοι μπορούν να προκαλέσουν στα μοντέλα να παρερμηνεύσουν τις πληροφορίες. Για παράδειγμα, η προσθήκη θορύβου σε μια εικόνα μπορεί να οδηγήσει ένα σύστημα υπολογιστικής

όρασης να ταξινομήσει λανθασμένα ένα σήμα "STOP" ως "Περιορισμό Ταχύτητας", θέτοντας σε κίνδυνο την ασφάλεια στην αυτόνομη οδήγηση (Spring et al., 2020).

Η εισαγωγή κακόβουλων δεδομένων, μπορεί να επηρεάσει τη συμπεριφορά του μοντέλου, προκαλώντας κακή απόδοση ή ανεπιθύμητες ενέργειες. Αυτή η ευαλωτότητα είναι ιδιαίτερα ανησυχητική για μοντέλα που εκπαιδεύονται σε δεδομένα από πλήθος ή δημόσια διαθέσιμα, όπου δεν μπορεί να διασφαλιστεί η ακεραιότητα του συνόλου δεδομένων.

Αυτές οι επιθέσεις στοχεύουν στην εξαγωγή ευαίσθητων πληροφοριών από τα μοντέλα TN. Η αντιστροφή μοντέλου επιτρέπει στους επιτιθέμενους να ανακατασκευάσουν τα δεδομένα εισόδου—όπως εικόνες ή κείμενο—που χρησιμοποιήθηκαν κατά την εκπαίδευση. Οι επιθέσεις εξαγωγής μελών καθορίζουν εάν ένα συγκεκριμένο δεδομένο περιλαμβανόταν στο σύνολο εκπαίδευσης. Και οι δύο τύποι επιθέσεων θέτουν σημαντικούς κινδύνους για την ιδιωτικότητα, ιδίως όταν πρόκειται για εμπιστευτικές ή προσωπικές πληροφορίες.

Τα συστήματα TN βασίζονται σε πολύπλοκα λογισμικά και υλικό που μπορεί να περιέχουν κενά ασφαλείας. Σφάλματα λογισμικού, μη ασφαλή APIs και ευπάθειες σε υποκείμενες βιβλιοθήκες μπορούν να εκμεταλλευτούν για να παραβιαστούν εφαρμογές TN. Επιπλέον, επιθέσεις που στοχεύουν στο υλικό, όπως οι επιθέσεις πλευρικών καναλιών (side-channel attacks), μπορούν να εξάγουν ευαίσθητες πληροφορίες αναλύοντας τις εκπομπές του υλικού ή τα πρότυπα χρήσης πόρων.

Τα μοντέλα AI μπορεί να αποδίδουν καλά στα δεδομένα εκπαίδευσης αλλά να αποτυγχάνουν να γενικεύσουν σε νέα, άγνωστα δεδομένα. Αυτή η έλλειψη ανθεκτικότητας τα καθιστά ευάλωτα σε απρόβλεπτες εισόδους ή αλλαγές στο περιβάλλον. Για παράδειγμα, ένα σύστημα AI εκπαιδευμένο για αναγνώριση εικόνων υπό συγκεκριμένες συνθήκες φωτισμού μπορεί να έχει κακή απόδοση όταν αυτές οι συνθήκες αλλάξουν, οδηγώντας σε λανθασμένες εξόδους (Spring et al., 2020).

Από το 2023, όλες οι μεγάλες εταιρείες λογισμικού έχουν ενσωματώσει την τεχνητή νοημοσύνη στις βασικές τους προσφορές λογισμικού. Για παράδειγμα, η Microsoft παρουσίασε το CoPilot για το GitHub και το PowerBI, επιτρέποντας στους προγραμματιστές και τους αναλυτές δεδομένων να λαμβάνουν υποστήριξη από την TN. Ωστόσο, αυτή η ενσωμάτωση δημιουργεί ευπάθειες στο λογισμικό τους. Οι επιτιθέμενοι μπορούν να εκμεταλλευτούν αυτές τις αδυναμίες για να διαταράξουν τη λειτουργικότητα των συστημάτων. Καθώς η χρήση της TN συνεχίζει να

επεκτείνεται μέσω της υιοθέτησης από το ευρύ κοινό, μοντέλα που έχουν σχεδιαστεί ή εκπαιδευτεί ανεπαρκώς ενδέχεται να διαρρεύσουν εταιρικές ή προσωπικές πληροφορίες μέσω συμπερασματικών επιθέσεων και επιθέσεων προτροπής.

Ας εξετάσουμε ένα σενάριο όπου μια επιχείρηση χρησιμοποιεί ιδιόκτητα δεδομένα για την εκπαίδευση ενός αλγορίθμου TN. Ελλείψει κατάλληλων ελέγχων ή διακυβέρνησης, ο αλγόριθμος μπορεί να έχει πρόσβαση σε υπερβολικά πολλές εταιρικές πληροφορίες, συμπεριλαμβανομένων της επιχειρηματικής στρατηγικής, των πληροφοριών πελατών, των προγραμμάτων και των εμπορικών μυστικών ή της πνευματικής ιδιοκτησίας. Ένας επιτιθέμενος που αποκτά πρόσβαση στην TN μπορεί να την προτρέπει με στοχευμένες ερωτήσεις, οδηγώντας την ακούσια στην αποκάλυψη ευαίσθητων εταιρικών πληροφοριών.

Σε ένα άλλο σενάριο, ένας επιτιθέμενος μπορεί να τροφοδοτήσει το σύστημα TN με εσφαλμένα ή ελλιπή σύνολα δεδομένων, προκαλώντας το να παράγει λανθασμένες, μεροληπτικές ή ανακριβείς προβλέψεις. Ανάλογα με την ταχύτητα του κύκλου λήψης αποφάσεων της εταιρείας, μια τέτοια επίθεση μπορεί να οδηγήσει σε σημαντικές οικονομικές ζημιές. Αυτές οι επιθέσεις βασίζονται στα δεδομένα εκπαίδευσης που παρέχονται στο σύστημα TN, καθώς αυτό μπορεί να κάνει προβλέψεις μόνο με βάση τις πληροφορίες που έχει λάβει. Εάν το σύστημα τροφοδοτηθεί με ανακριβή ή ψευδή δεδομένα, οι προβλέψεις του θα είναι επίσης λανθασμένες ή ακόμη και κακόβουλες.

Η αντιμετώπιση των απειλών ασφάλειας δεδομένων που προκύπτουν από την TN απαιτεί μια πολυεπίπεδη προσέγγιση. Δεν πρόκειται για υποθετικά σενάρια, αλλά για πραγματικές πιθανότητες. Για παράδειγμα, το GitHub CoPilot ή το Google Gemini Code Assist μπορούν να ρυθμιστούν ώστε να έχουν πρόσβαση σε ολόκληρη τη βάση κώδικα μιας εταιρείας. Ενώ αυτό μπορεί να είναι επωφελές για τους προγραμματιστές, παρέχει επίσης βαθιά πρόσβαση στην κύρια πνευματική ιδιοκτησία της εταιρείας. Αυτό είναι ιδιαίτερα επικίνδυνο για εταιρείες λογισμικού και τεχνολογίας που αντλούν το μεγαλύτερο μέρος των εσόδων τους από την πνευματική ιδιοκτησία. Για αυτές τις εταιρείες, η εφαρμογή ισχυρών μεθόδων διακυβέρνησης και ασφάλειας δεδομένων για τον έλεγχο της πρόσβασης της TN είναι ζωτικής σημασίας.

Η υπερβολική εξάρτηση από την TN χωρίς επαρκή ανθρώπινη εποπτεία μπορεί να οδηγήσει σε ευαλωτότητες. Τα συστήματα της ενδέχεται να λαμβάνουν αποφάσεις βάσει ελλιπών ή εσφαλμένων δεδομένων, και χωρίς ανθρώπινη παρέμβαση, αυτά τα λάθη μπορούν να

διαγωνίζονται. Η διασφάλιση μιας προσέγγισης με τον άνθρωπο στον βρόχο (human-in-the-loop) μπορεί να μετριάσει τους κινδύνους που σχετίζονται με την υπερβολική εξάρτηση από την ΤΝ.

Με την κατανόηση των πιθανών αδυναμιών και την εφαρμογή προληπτικών στρατηγικών μετριασμού, οι προγραμματιστές και οι οργανισμοί μπορούν να ενισχύσουν την ασφάλεια και την αξιοπιστία των συστημάτων ΤΝ.

3.4. Απειλές που οφείλονται στην ΤΝ

Η ανάπτυξη και χρήση της ΤΝ εγείρει σημαντικά ηθικά και κοινωνικά ζητήματα. Η αυτοματοποίηση μέσω της ΤΝ μπορεί να οδηγήσει σε απώλεια θέσεων εργασίας και οικονομική ανισότητα. Επιπλέον, οι αποφάσεις που λαμβάνονται από αλγορίθμους ΤΝ μπορεί να περιέχουν προκαταλήψεις, οι οποίες αντικατοπτρίζουν τις προκαταλήψεις στα δεδομένα εκπαίδευσης. Αυτό μπορεί να οδηγήσει σε διακρίσεις και άδικες πρακτικές, ιδιαίτερα σε τομείς όπως η απασχόληση, η υγειονομική περίθαλψη και το ποινικό σύστημα.

Τα τελευταία χρόνια, η αξιοσημείωτη πρόοδος της Τεχνητής Νοημοσύνης (ΤΝ) έχει φέρει επανάσταση στο τεχνολογικό τοπίο μας. Τα οφέλη της έχουν όχι μόνο βελτιώσει την αποδοτικότητα των καθημερινών μας λειτουργιών, αλλά έχουν επίσης προκαλέσει μετασχηματιστικές αλλαγές σε διάφορες βιομηχανίες. Η επίδραση της ΤΝ έχει κάνει τη ζωή μας πιο βολική, δημιουργώντας νέες ευκαιρίες στον ψηφιακό κόσμο. Προσβλέποντας στο μέλλον, η επιρροή της ΤΝ υπόσχεται ένα μέλλον γεμάτο καινοτομία και δυναμικό (Mauri & Damiani, 2022).

Ωστόσο, όπου υπάρχει θετική ανάπτυξη, συχνά αναδύεται και ένα αρνητικό αντίστοιχο. Οι κυβερνοεγκληματίες έχουν επιδέξια εκμεταλλευτεί την ΤΝ για κακόβουλους σκοπούς. Τα σημαντικά οφέλη της ΤΝ, όταν χρησιμοποιούνται από κακόβουλους παράγοντες, μπορούν να οδηγήσουν σε καταστροφικά αποτελέσματα. Μια μελέτη που πραγματοποιήθηκε από την Darktrace αποκάλυψε ότι το 88% των ηγετών ασφάλειας προβλέπουν την αναπόφευκτη εμφάνιση επιθετικής ΤΝ. Επιπλέον, πάνω από το 80% των υπεύθυνων λήψης αποφάσεων στην κυβερνοασφάλεια συμφώνησαν ότι οι οργανισμοί χρειάζονται προηγμένες άμυνες κυβερνοασφάλειας για να αποτρέψουν αποτελεσματικά τις απειλές από την επιθετική ΤΝ.

Αυτό αναδεικνύει τη σημασία της αμυντικής Τεχνητής Νοημοσύνης. Η αμυντική ΤΝ περιλαμβάνει την αξιοποίηση τεχνικών τεχνητής νοημοσύνης και μηχανικής μάθησης για την προστασία συστημάτων, δικτύων και δεδομένων από κυβερνοαπειλές. Αφορά την εφαρμογή

τεχνολογιών που βασίζονται στην TN για την αναγνώριση, πρόληψη και απόκριση σε ένα ευρύ φάσμα επιθέσεων και ευπαθειών. Επιπλέον, μια πρόσφατη έρευνα που δημοσιεύθηκε από το MIT Technology Review Insights και την Darktrace υποδεικνύει ότι το 96% των εκτελεστικών στελεχών εξετάζουν την υιοθέτηση της “αμυντικής TN” για την αντιμετώπιση των κυβερνοεπιθέσεων.

Ασφάλεια τελικών σημείων

Οι λύσεις Endpoint Detection and Response (EDR) που υποστηρίζονται από την TN παρακολουθούν συνεχώς τη δραστηριότητα των συσκευών και αναγνωρίζουν, ανταποκρίνονται και περιλαμβάνουν δυνητικές απειλές άμεσα.

Τα επόμενης γενιάς προγράμματα Antivirus που βασίζονται στην TN επιτρέπουν την ανίχνευση και την άμβλυνση προηγουμένως άγνωστου κακόβουλου λογισμικού μέσω μιας λεπτομερούς ανάλυσης της συμπεριφοράς των αρχείων, μειώνοντας την εξάρτηση από την ανίχνευση βάσει υπογραφών. Η αμυντική TN διαδραματίζει κρίσιμο ρόλο στην προστασία έναντι της επιθετικής TN. Όχι μόνο αντεπιτίθεται στις επιθέσεις, αλλά επίσης ενισχύει και βελτιστοποιεί τα υπάρχοντα εργαλεία κυβερνοασφάλειας. Χρησιμοποιώντας προηγμένους αλγορίθμους και προσαρμοστική μάθηση, εντοπίζει και μετριάζει ενεργά τις απειλές, ενισχύοντας τις ψηφιακές μας άμυνες (Mauri & Damiani, 2022).

Η ενσωμάτωση της Τεχνητής Νοημοσύνης (TN) σε διάφορους τομείς της κοινωνίας και της οικονομίας προσφέρει σημαντικά οφέλη, αλλά παράλληλα εισάγει νέες και αυξανόμενες απειλές. Οι κίνδυνοι αυτοί ποικίλουν από κυβερνοεπιθέσεις και παραβιάσεις προσωπικών δεδομένων έως ηθικές και κοινωνικές συνέπειες. Αναλύοντας τις κυριότερες απειλές που απορρέουν από την TN, γίνεται σαφές ότι οι οργανισμοί πρέπει να αναπτύξουν στρατηγικές για την αντιμετώπισή τους.

Μία από τις πιο άμεσες απειλές είναι η χρήση της TN από κυβερνοεγκληματίες για την ανάπτυξη προηγμένων και δυσκολότερα ανιχνεύσιμων επιθέσεων. Οι επιτιθέμενοι μπορούν να χρησιμοποιήσουν αλγορίθμους μηχανικής μάθησης για να δημιουργήσουν κακόβουλο λογισμικό που προσαρμόζεται και μεταλλάσσεται για να αποφύγει τα συστήματα ανίχνευσης. Επίσης, η TN μπορεί να χρησιμοποιηθεί για την εκτέλεση αυτοματοποιημένων επιθέσεων, όπως οι επιθέσεις DDoS, με αυξημένη αποτελεσματικότητα και ταχύτητα.

Κακόβουλοι παράγοντες μπορούν να χρησιμοποιήσουν αλγορίθμους TN για να αυτοματοποιήσουν και να ενισχύσουν τις κυβερνοεπιθέσεις, καθιστώντας τες πιο περίπλοκες και δυσκολότερες στην ανίχνευση. Για παράδειγμα, κακόβουλο λογισμικό που υποστηρίζεται από TN μπορεί να προσαρμόζει τη συμπεριφορά του σε πραγματικό χρόνο για να παρακάμπτει τα μέτρα ασφαλείας, ενώ επιθέσεις phishing με τη χρήση TN μπορούν να δημιουργούν εξαιρετικά εξατομικευμένα μηνύματα για την εξαπάτηση χρηστών. Η αυξημένη αυτή απειλή απαιτεί την ανάπτυξη προηγμένων αμυντικών μηχανισμών που μπορούν να ανταποκριθούν στον ρυθμό των εξελιγμένων επιθέσεων που ενισχύονται από την TN (Bécue et al., 2021).

Η TN μπορεί να επεξεργαστεί τεράστιες ποσότητες δεδομένων, καθιστώντας την ιδιαίτερα αποτελεσματική στην εξαγωγή προσωπικών πληροφοριών από διάφορες πηγές. Αυτό αυξάνει τον κίνδυνο παραβιάσεων προσωπικών δεδομένων, καθώς οι επιτιθέμενοι μπορούν να χρησιμοποιήσουν τεχνικές TN για να εντοπίσουν και να εκμεταλλευτούν ευαίσθητα δεδομένα. Επιπλέον, η ικανότητα της TN να αναγνωρίζει πρότυπα μπορεί να οδηγήσει σε ακούσιες διαρροές προσωπικών πληροφοριών ακόμα και σε ανώνυμα δεδομένα.

Η TN διευκολύνει τη δημιουργία ψευδών ειδήσεων και την παραπληροφόρηση μέσω της δημιουργίας ρεαλιστικών αλλά ψευδών εικόνων, βίντεο και κειμένων (deepfakes). Οι τεχνικές αυτές μπορούν να χρησιμοποιηθούν για να χειραγωγήσουν τη δημόσια γνώμη, να επηρεάσουν εκλογές και να προκαλέσουν κοινωνική αναταραχή. Η ταχύτητα και η κλίμακα με την οποία μπορεί να διαδοθεί η παραπληροφόρηση μέσω της TN καθιστούν την αντιμετώπιση της μια ιδιαίτερα δύσκολη πρόκληση.

Οι αλγόριθμοι TN εκπαιδεύονται σε υπάρχοντα δεδομένα, τα οποία ενδέχεται να περιέχουν ιστορικές προκαταλήψεις και μεροληψίες. Εάν δεν αντιμετωπιστούν κατάλληλα, αυτές οι προκαταλήψεις μπορούν να διατηρηθούν ή ακόμη και να ενισχυθούν από αυτά τα συστήματα, οδηγώντας σε διακριτική μεταχείριση σε τομείς όπως η πρόσληψη, η χορήγηση δανείων και η επιβολή του νόμου. Για παράδειγμα, συστήματα αναγνώρισης προσώπου έχουν παρουσιάσει αυξημένα ποσοστά σφάλματος για συγκεκριμένες δημογραφικές ομάδες, με αποτέλεσμα άδικες συνέπειες. Η αντιμετώπιση της αλγοριθμικής μεροληψίας απαιτεί τη χρήση ποικιλόμορφων και αντιπροσωπευτικών συνόλων δεδομένων, καθώς και συνεχή αξιολόγηση και προσαρμογή των μοντέλων TN για την προώθηση της δικαιοσύνης και της ισότητας (Bécue et al., 2021).

Τα συστήματα ΤΝ δεν είναι απαλλαγμένα από κενά ασφαλείας. Επιθέσεις αντιπάλων μπορούν να τροποποιήσουν τα δεδομένα εισόδου με ανεπαίσθητους τρόπους, εξαπατώντας τα μοντέλα ΤΝ και οδηγώντας σε λανθασμένα αποτελέσματα. Για παράδειγμα, η αλλαγή ελάχιστων εικονοστοιχείων σε μια εικόνα μπορεί να προκαλέσει εσφαλμένη ταξινόμηση από το σύστημα ΤΝ. Τέτοιες αδυναμίες μπορούν να αξιοποιηθούν για να διαταράξουν υπηρεσίες ή να προκαλέσουν απρόβλεπτη συμπεριφορά στα συστήματα ΤΝ. Η ενίσχυση της ανθεκτικότητας των μοντέλων αυτών απέναντι σε τέτοιες επιθέσεις είναι κρίσιμη για τη διατήρηση της αξιοπιστίας τους.

Κεφάλαιο 4: Πλεονεκτήματα και Μειονεκτήματα της χρήσης ΤΝ σε κρίσιμες υποδομές

Η ενσωμάτωση της τεχνητής νοημοσύνης (ΑΙ) σε κρίσιμες υποδομές προσφέρει μεγάλες δυνατότητες μετασχηματισμού, ενισχύοντας την αποτελεσματικότητα, την παραγωγικότητα και τις διαδικασίες λήψης αποφάσεων. Η ικανότητα της ΑΙ να αυτοματοποιεί εργασίες ρουτίνας και να βελτιστοποιεί τη διαχείριση των πόρων μειώνει σημαντικά το λειτουργικό κόστος και αυξάνει την αποδοτικότητα του συστήματος. Τα προγνωστικά αναλυτικά στοιχεία που παρέχονται από την τεχνητή νοημοσύνη μπορούν να προβλέψουν τάσεις και πιθανά ζητήματα, επιτρέποντας την προληπτική συντήρηση και διαχείριση κινδύνου. Αυτές οι δυνατότητες όχι μόνο απλοποιούν τις λειτουργίες αλλά και διασφαλίζουν μια πιο αξιόπιστη και ανθεκτική υποδομή, ικανή να αντέχει σε διάφορες προκλήσεις και να διατηρεί συνεχή παροχή υπηρεσιών.

Παρά τα πλεονεκτήματα αυτά, η εφαρμογή της τεχνητής νοημοσύνης σε υποδομές ζωτικής σημασίας εμπεριέχει σημαντικούς κινδύνους για την ασφάλεια. Η εξάρτηση από συστήματα τεχνητής νοημοσύνης εκθέτει την υποδομή σε απειλές στον κυβερνοχώρο, όπου εξελιγμένες επιθέσεις μπορούν να θέσουν σε κίνδυνο την ακεραιότητα και τη λειτουργικότητα αυτών των συστημάτων. Επιπλέον, ο τεράστιος όγκος δεδομένων που επεξεργάζεται η τεχνητή νοημοσύνη εγείρει σημαντικές ανησυχίες σχετικά με το απόρρητο των δεδομένων και την προστασία των ευαίσθητων πληροφοριών. Η διασφάλιση ισχυρών μέτρων κυβερνοασφάλειας και πρωτοκόλλων προστασίας δεδομένων είναι πρωταρχικής σημασίας για την αποτροπή πιθανών παραβιάσεων και τη διατήρηση της εμπιστοσύνης των ενδιαφερομένων (Sakhnini and Karimirpour, 2023) .

Οι ηθικές και νομικές εκτιμήσεις παρουσιάζουν επίσης σοβαρές προκλήσεις στην ανάπτυξη της τεχνητής νοημοσύνης σε υποδομές ζωτικής σημασίας. Η πολυπλοκότητα του καθορισμού της λογοδοσίας για αποφάσεις και ενέργειες που βασίζονται στην τεχνητή νοημοσύνη περιπλέκει το τοπίο, ειδικά σε περίπτωση αστοχιών ή ατυχημάτων του συστήματος. Επιπλέον, η δυνατότητα των συστημάτων τεχνητής νοημοσύνης να διαιωνίζουν και να ενισχύουν τις υπάρχουσες προκαταλήψεις εγείρει ανησυχίες σχετικά με τη δικαιοσύνη και τις διακρίσεις. Η αντιμετώπιση αυτών των ηθικών ζητημάτων απαιτεί ολοκληρωμένα ρυθμιστικά πλαίσια και ηθικές κατευθυντήριες γραμμές που θα διέπουν τη χρήση και την ανάπτυξη τεχνολογιών τεχνητής νοημοσύνης σε κρίσιμους τομείς.

Επιπλέον, οι οικονομικές και κοινωνικές επιπτώσεις της ενσωμάτωσης της τεχνητής νοημοσύνης σε υποδομές ζωτικής σημασίας δεν μπορούν να παραβλεφθούν. Ενώ η αυτοματοποίηση και η βελτιστοποιημένη απόδοση μπορεί να οδηγήσουν σε εξοικονόμηση κόστους, ενέχουν επίσης τον κίνδυνο μετατόπισης θέσεων εργασίας, απαιτώντας την επανεκπαίδευση και την προσαρμογή του εργατικού δυναμικού. Τα οφέλη της τεχνητής νοημοσύνης ενδέχεται να μην κατανέμονται ομοιόμορφα, επιδεινώνοντας ενδεχομένως τις οικονομικές ανισότητες. Οι τεχνικές προκλήσεις που σχετίζονται με την εφαρμογή, τη συντήρηση και την αναβάθμιση συστημάτων τεχνητής νοημοσύνης προσθέτουν ένα άλλο επίπεδο πολυπλοκότητας, απαιτώντας σημαντικές επενδύσεις και τεχνογνωσία. Η εξισορρόπηση του μετασχηματιστικού δυναμικού της τεχνητής νοημοσύνης με αυτές τις πολύπλευρες προκλήσεις είναι απαραίτητη για την επιτυχή και δίκαιη ανάπτυξή της σε υποδομές ζωτικής σημασίας.

Ένα από τα σημαντικότερα πλεονεκτήματα της τεχνητής νοημοσύνης είναι η ικανότητά της να αυτοματοποιεί επαναλαμβανόμενες εργασίες, απελευθερώνοντας έτσι το ανθρώπινο δυναμικό ώστε να επικεντρωθεί σε πιο σύνθετες και δημιουργικές προσπάθειες. Μέσω της απλοποίησης των διαδικασιών, η τεχνητή νοημοσύνη μπορεί να βελτιώσει σημαντικά την αποτελεσματικότητα και την παραγωγικότητα σε διάφορους τομείς, ειδικότερα σε εργασίες που απαιτούν επαναληπτικότητα.

Σε αντίθεση με τους ανθρώπους, οι μηχανές δεν χρειάζονται διαλείμματα για να αναρρώσουν από την κούραση, γεγονός που αυξάνει την παραγωγικότητά τους. Οι μηχανές που βασίζονται σε τεχνητή νοημοσύνη μπορούν να εκτελούν επαναλαμβανόμενες εργασίες για μεγάλα χρονικά διαστήματα χωρίς επιβράδυνση. Για παράδειγμα, στον τομέα της κατασκευής, τα ρομπότ με τεχνητή νοημοσύνη μπορούν να λειτουργούν αδιάκοπα, αυξάνοντας τους ρυθμούς παραγωγής και μειώνοντας τα σφάλματα. Η τεχνητή νοημοσύνη διευκολύνει τη λειτουργία μηχανών για απεριόριστο χρονικό διάστημα χωρίς μείωση της παραγωγικότητας. Αυτό είναι ένα από τα βασικά πλεονεκτήματα της τεχνητής νοημοσύνης, το οποίο έχει οδηγήσει στην ευρεία αποδοχή της σε κάθε τομέα. Η τεχνητή νοημοσύνη χρησιμοποιείται από τους κατασκευαστές για να παράγουν συνεχώς προϊόντα, ανταποκρινόμενοι στις απαιτήσεις της αγοράς και επιτυγχάνοντας υψηλά κέρδη (Guzman et al., 2016).

Επιπλέον, η συνεχής λειτουργία των μηχανών με τεχνητή νοημοσύνη επιτρέπει τη διατήρηση υψηλών επιπέδων παραγωγής, μειώνοντας ταυτόχρονα τα λειτουργικά κόστη και βελτιώνοντας

την ποιότητα των παραγόμενων προϊόντων. Αυτός ο συνδυασμός αποτελεσματικότητας και αυξημένης παραγωγικότητας καθιστά την τεχνητή νοημοσύνη αναπόσπαστο εργαλείο για τη σύγχρονη βιομηχανία, ενισχύοντας την ανταγωνιστικότητα και την καινοτομία.

Οι αλγόριθμοι τεχνητής νοημοσύνης διακρίνονται για την ικανότητά τους να αναλύουν μεγάλες ποσότητες δεδομένων σε πολύ σύντομο χρονικό διάστημα, σε σύγκριση με τον χρόνο που απαιτείται από τους ανθρώπους. Αυτή η δυνατότητα επιτρέπει στις επιχειρήσεις να αποκτούν πολύτιμες πληροφορίες και να λαμβάνουν αποφάσεις βάσει δεδομένων, οδηγώντας σε βελτιωμένες στρατηγικές, εξατομικευμένες εμπειρίες πελατών και βελτιστοποιημένες λειτουργίες.

Στον τομέα της υγειονομικής περίθαλψης, η τεχνητή νοημοσύνη μπορεί να συμβάλλει στη διάγνωση ασθενειών, στην πρόβλεψη των αποτελεσμάτων των ασθενών και στον εντοπισμό προτύπων στα ιατρικά δεδομένα. Αυτές οι δυνατότητες τελικά βελτιώνουν την αποτελεσματικότητα της φροντίδας και της θεραπείας των ασθενών, προσφέροντας πιο έγκαιρες και ακριβείς διαγνώσεις.

Η χρήση της τεχνητής νοημοσύνης στην ανάλυση δεδομένων επιτρέπει επίσης την ανίχνευση τάσεων και ανωμαλιών που ενδέχεται να μην είναι άμεσα αντιληπτές από τους ανθρώπους. Αυτές οι πληροφορίες μπορούν να αξιοποιηθούν για τη βελτίωση των επιχειρησιακών διαδικασιών και την ανάπτυξη νέων προϊόντων και υπηρεσιών, ενισχύοντας έτσι την καινοτομία και την ανταγωνιστικότητα.

Οι οργανισμοί μπορούν να αυτοματοποιήσουν διάφορες διαδικασίες αξιοποιώντας την τεχνολογία ΑΙ, οδηγώντας σε σημαντική εξοικονόμηση κόστους. Η ικανότητα της ΑΙ να αυτοματοποιεί διάφορες εργασίες που εκτελούνται επί του παρόντος από ανθρώπους αποτελεί ένα από τα κύρια πλεονεκτήματά της. Παραδείγματα τέτοιων διαδικασιών περιλαμβάνουν την εισαγωγή δεδομένων, την ανάλυση δεδομένων και την εξυπηρέτηση πελατών. Με την αυτοματοποίηση αυτών των λειτουργιών, οι επιχειρήσεις μπορούν να εξοικονομήσουν χρόνο, να αυξήσουν την αποδοτικότητα και να αναθέσουν το ανθρώπινο δυναμικό σε πιο απαιτητικές και δημιουργικές εργασίες (Guzman et al., 2016). Η αξιοποίηση της τεχνητής νοημοσύνης για την αυτοματοποίηση επιτρέπει στις επιχειρήσεις να μειώσουν τα λειτουργικά κόστη, βελτιώνοντας παράλληλα την

ακρίβεια και την ταχύτητα εκτέλεσης των διαδικασιών. Αυτό έχει ως αποτέλεσμα την παροχή καλύτερων υπηρεσιών στους πελάτες και τη βελτιστοποίηση των εσωτερικών λειτουργιών.

Η Τεχνητή Νοημοσύνη (AI) έχει τη δυνατότητα να οδηγήσει και να συμβάλει στη δημιουργία καινοτομιών με ποικίλους τρόπους. Η τεχνητή νοημοσύνη πυροδοτεί πολλές εφευρέσεις σε σχεδόν όλους τους τομείς, διευκολύνοντας τους ανθρώπους να επιλύουν τα πιο σύνθετα προβλήματα. Για παράδειγμα, οι γιατροί έχουν πρόσφατα αξιοποιήσει την τεχνολογία που βασίζεται στην τεχνητή νοημοσύνη για την πρόβλεψη του καρκίνου του μαστού σε πρώιμο στάδιο. Αυτή η δυνατότητα πρόβλεψης οφείλεται στην ανάλυση μεγάλων όγκων δεδομένων και την ανίχνευση προτύπων που δεν είναι άμεσα αντιληπτά από τους ανθρώπους. Με αυτόν τον τρόπο, η τεχνητή νοημοσύνη συμβάλλει σημαντικά στην έγκαιρη διάγνωση και θεραπεία, αυξάνοντας τις πιθανότητες επιτυχίας και βελτιώνοντας την ποιότητα ζωής των ασθενών.

Επιπλέον, η τεχνητή νοημοσύνη ενισχύει την καινοτομία σε τομείς όπως η φαρμακευτική έρευνα, η κατασκευή και η χρηματοοικονομική ανάλυση. Με την ικανότητά της να επεξεργάζεται τεράστιες ποσότητες δεδομένων και να παρέχει ακριβείς προβλέψεις, η AI βοηθά τις επιχειρήσεις να αναπτύσσουν νέα προϊόντα και υπηρεσίες, να βελτιστοποιούν τις διαδικασίες τους και να λαμβάνουν καλύτερες αποφάσεις.

Η τεχνητή νοημοσύνη βασίζεται σε μεγάλο βαθμό στα δεδομένα, και η συλλογή και η χρήση προσωπικών πληροφοριών εγείρει σημαντικές ανησυχίες σχετικά με την ασφάλεια και το απόρρητο. Τα συστήματα τεχνητής νοημοσύνης πρέπει να σχεδιάζονται με ισχυρά μέτρα ασφαλείας για την προστασία ευαίσθητων δεδομένων από μη εξουσιοδοτημένη πρόσβαση και παραβιάσεις. Αυτό απαιτεί την υιοθέτηση προηγμένων τεχνολογιών κρυπτογράφησης και αυστηρών πρωτοκόλλων ασφαλείας που εξασφαλίζουν την ακεραιότητα και την εμπιστευτικότητα των δεδομένων.

Η αντιμετώπιση αυτών των κινδύνων απαιτεί την ανάπτυξη ρυθμιστικών πλαισίων και κατευθυντήριων γραμμών για τη χρήση της τεχνητής νοημοσύνης. Οι κανονισμοί πρέπει να διασφαλίζουν ότι τα συστήματα τεχνητής νοημοσύνης χρησιμοποιούνται με υπευθυνότητα και διαφάνεια, προλαμβάνοντας τη διάδοση ψευδών πληροφοριών και προστατεύοντας τα δικαιώματα των ατόμων. Η διεθνής συνεργασία και η ανταλλαγή βέλτιστων πρακτικών είναι κρίσιμη για την επίτευξη αυτών των στόχων.

Η ευαισθητοποίηση του κοινού σχετικά με τις δυνατότητες και τους κινδύνους της τεχνητής νοημοσύνης είναι απαραίτητη. Οι πολίτες πρέπει να ενημερώνονται για τις πιθανές απειλές και να εκπαιδεύονται στην αναγνώριση και αντιμετώπιση παραπληροφόρησης. Μόνο μέσα από συντονισμένες προσπάθειες όλων των εμπλεκόμενων φορέων μπορεί η τεχνητή νοημοσύνη να αξιοποιηθεί με ασφάλεια και υπευθυνότητα, μεγιστοποιώντας τα οφέλη της για την κοινωνία.

Η ενσωμάτωση και ανάπτυξη της τεχνητής νοημοσύνης στην ασφάλεια των Κυβερνο-Φυσικών Συστημάτων (CPS) ολοκληρώνεται μέσω της χρήσης μεθόδων βαθιάς μάθησης (Deep Learning - DL). Η βαθιά μάθηση παρέχει μια πλούσια συλλογή μοντέλων που μπορούν να προσεγγίσουν οποιαδήποτε συνάρτηση και να προσαρμοστούν σε οποιαδήποτε δεδομένα. Επιπλέον, η DL προσφέρει επεκτασιμότητα, η οποία είναι απαραίτητη για μεγάλες υποδομές ζωτικής σημασίας. Αυτές οι ιδιότητες είναι ιδιαίτερα επιθυμητές, διότι τα CPS συνήθως επεξεργάζονται δεδομένα υψηλών διαστάσεων που προέρχονται από μεγάλο αριθμό αισθητήρων. Επιπλέον, τα CPS ενημερώνονται συνεχώς με νέα δεδομένα, γεγονός που καθιστά τη χρήση της DL ευρέως διαδεδομένη στην ασφάλεια στον κυβερνοχώρο των CPS (Pinto et al., 2023).

Ωστόσο, υπάρχουν πολλές προκλήσεις που σχετίζονται με τη χρήση της DL. Ένα από τα σημαντικότερα προβλήματα των προσεγγίσεων DL είναι η έλλειψη θεωρητικού υποβάθρου στο θέμα. Οι περισσότερες μελέτες στο αντικείμενο δείχνουν την εντυπωσιακή απόδοση των μεθόδων DL χωρίς να παρέχουν λεπτομερείς εξηγήσεις για το πώς αυτές γενικεύονται καλά. Αυτή η έλλειψη διαφάνειας έχει μειώσει την εμπιστοσύνη των πολιτών στις προσεγγίσεις DL. Επιπλέον, έχει καταστήσει δύσκολη την εφαρμογή των κατάλληλων τεχνικών προσαρμογής σε πραγματικές εφαρμογές. Η απουσία θεωρητικής βάσης και διαφάνειας στις μεθόδους βαθιάς μάθησης δημιουργεί εμπόδια στην πλήρη αποδοχή και εμπιστοσύνη των πολιτών και των επαγγελματιών. Η έλλειψη λεπτομερούς κατανόησης των μηχανισμών λειτουργίας αυτών των μεθόδων περιορίζει την αποτελεσματική εφαρμογή τους σε πραγματικούς κόσμους.

Για να αντιμετωπιστούν αυτές οι προκλήσεις, απαιτείται περαιτέρω έρευνα και ανάπτυξη θεωρητικών προτύπων που να εξηγούν τις αρχές λειτουργίας της βαθιάς μάθησης. Μόνο μέσω της επίλυσης αυτών των ζητημάτων θα μπορέσει η βαθιά μάθηση να εφαρμοστεί με επιτυχία και διαφάνεια στην ασφάλεια των CPS, εξασφαλίζοντας την εμπιστοσύνη και την αποδοχή των τεχνολογιών αυτών από την κοινωνία.

Η τακτοποίηση αποτελεί αναπόσπαστο μέρος της εφαρμογής μεθόδων τεχνητής νοημοσύνης σε οποιαδήποτε εφαρμογή, περιλαμβάνοντας τροποποιήσεις παραμέτρων για να διασφαλιστεί η επεκτασιμότητα στα συστήματα του πραγματικού κόσμου. Για παράδειγμα, στις κλασικές θεωρίες μηχανικής μάθησης (ML), η τακτοποίηση συνήθως περιλαμβάνει αλλαγές στις παραμέτρους έτσι ώστε ο αριθμός των δειγμάτων εκπαίδευσης να υπερβαίνει τον αριθμό των παραμέτρων στους αλγόριθμους ML. Στις μεθόδους βαθιάς μάθησης (DL), ωστόσο, έχει επιτευχθεί καλή γενίκευση ακόμη και με υπερβολικά παραμετροποιημένες ρυθμίσεις. Ορισμένες μέθοδοι DL απαιτούν διαφορετικές τεχνικές για καλή γενίκευση, όπως μοίρασμα βάρους, βάρος αποσύνθεσης, πτώση επιπέδων και αύξηση δεδομένων. Παρά τη συχνή χρήση αυτών των τεχνικών, η σημασία τους σπάνια επισημαίνεται και εκφράζεται επαρκώς.

Μια άλλη κρίσιμη πρόκληση για την εφαρμογή της τεχνητής νοημοσύνης, και ειδικότερα των μεθόδων DL, είναι η επιλογή της σωστής παραμέτρου και η ρύθμιση ισορροπίας του αλγορίθμου. Η αυξημένη ανάπτυξη της GPU έχει διευρύνει τις δυνατότητες των αλγορίθμων DL, ενθαρρύνοντας τους ερευνητές να εκπαιδεύουν ευρύτερα και βαθύτερα δίκτυα. Αυτά τα βαθύτερα δίκτυα έχουν τη δυνατότητα να μαθαίνουν πιο πολύπλοκα μοτίβα. Ωστόσο, έχει αποδειχθεί ότι τα δίκτυα με μεγαλύτερη χωρητικότητα μειώνουν την πρακτικότητα των μεθόδων τακτοποίησης, όπως η εγκατάλειψη ή η μείωση του βάρους, και απαιτούν μεγαλύτερο χρόνο εκπαίδευσης. Ως εκ τούτου, οι ερευνητές και οι προγραμματιστές AI πρέπει να βρουν την ισορροπία μεταξύ της εφαρμογής βαθύτερων αρχιτεκτονικών και της χρήσης μεθόδων τακτοποίησης. Αυτό είναι το κλειδί για την αύξηση της ευρωστίας των μεθόδων DL, οι οποίες μπορεί να είναι ευαίσθητες σε αντίθετα δείγματα. Η εύρεση αυτής της ισορροπίας είναι απαραίτητη για την ανάπτυξη αποτελεσματικών και αξιόπιστων συστημάτων τεχνητής νοημοσύνης που μπορούν να ανταποκριθούν στις απαιτήσεις των σύγχρονων εφαρμογών.

Η συνεχής βελτίωση και εξέλιξη των μεθόδων DL, σε συνδυασμό με την κατάλληλη τακτοποίηση, θα επιτρέψει την αξιοποίηση της πλήρους δυναμικής των τεχνολογιών αυτών. Με αυτόν τον τρόπο, θα εξασφαλιστεί η ανθεκτικότητα και η αποτελεσματικότητα των συστημάτων AI, προάγοντας την καινοτομία και τη βελτίωση σε διάφορους τομείς της επιστήμης και της βιομηχανίας.

Η εκθετική ανάπτυξη και ενσωμάτωση της τεχνητής νοημοσύνης στην ασφάλεια στον κυβερνοχώρο και σε όλες τις πτυχές της τεχνολογίας έχει εγείρει ιδεολογικούς και ηθικούς

προβληματισμούς. Αναφορές υποδεικνύουν ότι η εξάρτηση από την τεχνητή νοημοσύνη μπορεί να δημιουργήσει νέα τρωτά σημεία που θα μπορούσαν να αξιοποιηθούν από αντιπάλους. Συνεπώς, πρέπει να δοθεί ιδιαίτερη προσοχή στην ανάπτυξη τεχνολογιών ή πολιτικών λύσεων που μπορούν να μετριάσουν τις κακόβουλες εφαρμογές αυτών των εργαλείων ΑΙ.

Η ενσωμάτωση της τεχνητής νοημοσύνης εισάγει επίσης νέες δυσκολίες στις αλληλεπιδράσεις μεταξύ μηχανών και ανθρώπων. Η ερμηνεία πληροφοριών από πολύπλοκα αυτόνομα συστήματα μπορεί να γίνει όλο και πιο δύσκολη. Υπάρχουν επίσης εικασίες για τη δημιουργία αυτόνομων μηχανισμών λήψης αποφάσεων, οι οποίοι μπορεί να επηρεαστούν από υπερφόρτωση πληροφοριών ή ανθρώπινα συναισθήματα, επιφέροντας πρόσθετες προκλήσεις. Ορισμένες τεχνικές συστάσεις που έχουν προταθεί για το σχεδιασμό και την ανάπτυξη ευφύων συστημάτων περιλαμβάνουν: εξασφάλιση κατάλληλου ελέγχου υπό οποιεσδήποτε συνθήκες, καθορισμό αυστηρών περιορισμών για τη συμπεριφορά του αλγορίθμου, διεξαγωγή προσεκτικών δοκιμών και επικύρωσης για να ικανοποιηθούν οι ανησυχίες για την ασφάλεια, και περιορισμό του περιβάλλοντος στο οποίο επιτρέπεται η λειτουργία του ευφυούς συστήματος σε γνωστές πλατφόρμες.

Αυτές οι συστάσεις αποσκοπούν στη διασφάλιση της ασφάλειας και της αξιοπιστίας των συστημάτων τεχνητής νοημοσύνης, ενώ ταυτόχρονα αντιμετωπίζουν τις ιδεολογικές και ηθικές ανησυχίες που προκύπτουν από την ευρεία εφαρμογή τους. Η ισορροπία μεταξύ της αξιοποίησης των δυνατοτήτων της τεχνητής νοημοσύνης και της προστασίας από τους κινδύνους που μπορεί να προκύψουν είναι ζωτικής σημασίας για την επιτυχή ενσωμάτωση αυτής της τεχνολογίας στην κοινωνία.

Ένα άλλο θέμα που πρέπει να εξεταστεί περαιτέρω είναι το πρωτόκολλο επικοινωνίας. Ένα κύριο μέλημα είναι αν η επικοινωνία πρέπει να είναι μονόδρομη, επιτρέποντας μόνο τη συλλογή πληροφοριών, ή αμφίδρομη, επιτρέποντας την ανάληψη γρήγορων ενεργειών. Η αμφίδρομη επικοινωνία μπορεί να προσφέρει πλεονεκτήματα όπως ταχύτερους χρόνους αντίδρασης, αλλά είναι επίσης επιρρεπής σε πρόσθετα ζητήματα ασφάλειας, όπως η απώλεια ελέγχου του ευφυούς συστήματος.

Τα ολοκληρωμένα πρωτόκολλα επικοινωνίας έχουν επίσης μεγαλύτερες δυνατότητες αντιμετώπισης απειλών πολλαπλών πρακτόρων ή απειλών που αφορούν πιθανή συνεργατική

συμπεριφορά μεταξύ διάφορων πρακτόρων. Τέτοιοι σχηματισμοί πολλαπλών πρακτόρων έχουν υψηλές δυνατότητες εφαρμογής σε επιχειρήσεις στον κυβερνοχώρο. Οι επικοινωνίες μεταξύ πρακτόρων μπορούν να χρησιμοποιηθούν για πιο ολοκληρωμένη συνεργασία με σκοπό την επίτευξη γενικότερων στόχων. Ωστόσο, ανεπιθύμητοι συνασπισμοί μεταξύ πρακτόρων μπορεί να προκύψουν εάν δοθεί στους πράκτορες υπερβολική αυτονομία στη λήψη αποφάσεων. Αυτό μπορεί να έχει σοβαρές συνέπειες που είναι δύσκολο να αναστραφούν. Οι τεχνικές λύσεις που προτείνονται για την αντιμετώπιση τέτοιων προβλημάτων περιλαμβάνουν τη χρήση διασφαλίσεων όπως οι πίσω πόρτες και η αναγκαστική καταστροφή. Αυτές οι μέθοδοι μπορούν να προσφέρουν έναν μηχανισμό ασφαλείας σε περιπτώσεις που οι πράκτορες αποκτούν υπερβολική αυτονομία ή αλληλεπιδρούν με μη επιθυμητό τρόπο.

Η ανάπτυξη κατάλληλων πρωτοκόλλων επικοινωνίας και τεχνικών διασφαλίσεων είναι κρίσιμη για την ασφαλή και αποτελεσματική χρήση των ευφυών συστημάτων. Η ισορροπία μεταξύ ταχύτητας αντίδρασης και ασφάλειας πρέπει να διατηρηθεί για να διασφαλιστεί ότι τα συστήματα μπορούν να ανταποκριθούν αποτελεσματικά στις απαιτήσεις του σύγχρονου ψηφιακού περιβάλλοντος χωρίς να εκτίθενται σε περιττούς κινδύνους.

4.1. Αντιμετώπιση πιθανών μειονεκτημάτων και προκλήσεων

Η ενσωμάτωση της τεχνητής νοημοσύνης (AI) σε κρίσιμες υποδομές προσφέρει δυνατότητες μετασχηματισμού, αλλά παρουσιάζει επίσης πολλά μειονεκτήματα και προκλήσεις. Η αντιμετώπιση αυτών των ανησυχιών είναι απαραίτητη για να διασφαλιστεί ότι η ανάπτυξη της τεχνητής νοημοσύνης ενισχύει αντί να υπονομεύει την ανθεκτικότητα και την ασφάλεια των ζωτικών συστημάτων.

Μία από τις σημαντικότερες προκλήσεις της χρήσης της τεχνητής νοημοσύνης σε υποδομές ζωτικής σημασίας είναι ο αυξημένος κίνδυνος τρωτών σημείων ασφαλείας. Τα συστήματα τεχνητής νοημοσύνης μπορούν να γίνουν στόχοι για εξελιγμένες επιθέσεις στον κυβερνοχώρο, συμπεριλαμβανομένων των αντίθετων επιθέσεων που χειραγωγούν αλγόριθμους τεχνητής νοημοσύνης, προκαλώντας πιθανώς δυσλειτουργία ή λήψη λανθασμένων αποφάσεων. Η εξάρτηση από την τεχνητή νοημοσύνη αυξάνει την επιφάνεια επίθεσης, παρέχοντας περισσότερα σημεία εισόδου για κακόβουλους παράγοντες. Επιπλέον, η φύση του μαύρου κουτιού πολλών μοντέλων τεχνητής νοημοσύνης μπορεί να δυσκολέψει τον έγκαιρο εντοπισμό και τον μετριασμό αυτών των απειλών (Taylor et al., 2017).

Τα συστήματα τεχνητής νοημοσύνης, αν και ισχυρά, μπορεί μερικές φορές να είναι αναξιόπιστα και να μην έχουν στιβαρότητα, ιδιαίτερα σε απρόβλεπτες ή ακραίες συνθήκες. Για παράδειγμα, τα μοντέλα τεχνητής νοημοσύνης που έχουν εκπαιδευτεί σε ιστορικά δεδομένα ενδέχεται να δυσκολεύονται να προβλέψουν με ακρίβεια ή να ανταποκριθούν σε νέες καταστάσεις ή σπάνια γεγονότα, οδηγώντας σε πιθανές αστοχίες σε κρίσιμα συστήματα. Η διασφάλιση της αξιοπιστίας και της ευρωστίας της τεχνητής νοημοσύνης υπό όλες τις λειτουργικές συνθήκες απαιτεί εκτεταμένες δοκιμές, επικύρωση και συνεχή παρακολούθηση, η οποία μπορεί να απαιτεί ένταση πόρων.

Η ανάπτυξη της τεχνητής νοημοσύνης σε υποδομές ζωτικής σημασίας εγείρει σημαντικές ηθικές και νομικές ανησυχίες. Τα συστήματα τεχνητής νοημοσύνης μπορούν να διαιωνίσουν, ακόμη και να ενισχύσουν τις υπάρχουσες προκαταλήψεις, εάν δεν σχεδιαστούν και εκπαιδευτούν σωστά, οδηγώντας σε άδικα ή μεροληπτικά αποτελέσματα. Ο καθορισμός της λογοδοσίας για αποφάσεις και ενέργειες που βασίζονται στην τεχνητή νοημοσύνη είναι ένα άλλο πολύπλοκο ζήτημα, ειδικά σε περιπτώσεις αστοχιών συστήματος ή ατυχημάτων. Τα νομικά πλαίσια και οι δεοντολογικές κατευθυντήριες γραμμές είναι απαραίτητα για να διέπουν την υπεύθυνη χρήση της τεχνητής νοημοσύνης, διασφαλίζοντας τη διαφάνεια, τη δικαιοσύνη και τη λογοδοσία (Markopoulou, D., & Papakonstantinou, 2021).

Τα συστήματα τεχνητής νοημοσύνης απαιτούν συνήθως μεγάλες ποσότητες δεδομένων για να λειτουργήσουν αποτελεσματικά, τα οποία μπορεί να περιλαμβάνουν ευαίσθητες προσωπικές πληροφορίες. Η συλλογή, αποθήκευση και επεξεργασία αυτών των δεδομένων εγείρει σημαντικές ανησυχίες για το απόρρητο. Η διασφάλιση του απορρήτου και της ασφάλειας των δεδομένων είναι πρωταρχικής σημασίας για την αποτροπή κακής χρήσης και μη εξουσιοδοτημένης πρόσβασης. Η συμμόρφωση με τους κανονισμούς προστασίας δεδομένων, όπως ο GDPR, προσθέτει άλλο ένα επίπεδο πολυπλοκότητας στην ανάπτυξη της τεχνητής νοημοσύνης σε υποδομές ζωτικής σημασίας.

Η ενσωμάτωση της τεχνητής νοημοσύνης σε υπάρχοντα συστήματα υποδομής μπορεί να είναι πρόκληση λόγω ζητημάτων διαλειτουργικότητας. Πολλά κρίσιμα συστήματα βασίζονται σε τεχνολογίες παλαιού τύπου που ενδέχεται να μην διασυνδέονται απρόσκοπτα με σύγχρονες λύσεις τεχνητής νοημοσύνης. Η επίτευξη ομαλής ενοποίησης απαιτεί σημαντική προσπάθεια για τον επανασχεδιασμό και την αναβάθμιση των υποδομών, η οποία μπορεί να είναι δαπανηρή και

χρονοβόρα. Επιπλέον, τα ζητήματα διαλειτουργικότητας μπορούν να οδηγήσουν σε λειτουργικές διακοπές και αναποτελεσματικότητα.

Οι οικονομικές και κοινωνικές επιπτώσεις της ενσωμάτωσης της τεχνητής νοημοσύνης σε υποδομές ζωτικής σημασίας είναι σημαντικές. Ενώ η τεχνητή νοημοσύνη μπορεί να βελτιώσει την αποτελεσματικότητα και να μειώσει το κόστος, μπορεί επίσης να οδηγήσει σε μετατόπιση θέσεων εργασίας, καθώς τα αυτοματοποιημένα συστήματα αντικαθιστούν τους ανθρώπινους εργαζομένους. Αυτό απαιτεί επενδύσεις για την αναβάθμιση των δεξιοτήτων και την αναβάθμιση του εργατικού δυναμικού ώστε να προσαρμοστεί σε νέους ρόλους. Επιπλέον, τα οφέλη της τεχνητής νοημοσύνης ενδέχεται να μην κατανέμονται ομοιόμορφα, επιδεινώνοντας ενδεχομένως τις υπάρχουσες οικονομικές ανισότητες.

Η υπερβολική εξάρτηση από την τεχνητή νοημοσύνη μπορεί να οδηγήσει σε απώλεια της ανθρώπινης επίβλεψης και ελέγχου. Τα συστήματα υποδομής ζωτικής σημασίας που διαχειρίζονται κυρίως η τεχνητή νοημοσύνη μπορεί να μειώσουν τον ρόλο των ανθρώπινων χειριστών, οι οποίοι είναι ουσιαστικοί για την ερμηνεία περίπλοκων καταστάσεων και τη λήψη διαφοροποιημένων αποφάσεων. Αυτή η εξάρτηση μπορεί να είναι επικίνδυνη, ιδιαίτερα σε καταστάσεις όπου η ανθρώπινη διαίσθηση και εμπειρία είναι ζωτικής σημασίας. Η διατήρηση μιας ισορροπίας μεταξύ της αυτοματοποίησης τεχνητής νοημοσύνης και της ανθρώπινης επίβλεψης είναι ζωτικής σημασίας για τη διασφάλιση της ανθεκτικότητας και της υπευθυνότητας του συστήματος.

Το ρυθμιστικό τοπίο για την τεχνητή νοημοσύνη σε υποδομές ζωτικής σημασίας είναι περίπλοκο και συνεχώς εξελίσσεται. Οι οργανισμοί πρέπει να πλοηγηθούν σε μια μυριάδα κανονισμών και προτύπων για να εξασφαλίσουν την ασφαλή και συμβατή ανάπτυξη τεχνολογιών τεχνητής νοημοσύνης. Αυτή μπορεί να είναι μια προκλητική και έντασης πόρων διαδικασία, που απαιτεί συνεχή προσαρμογή σε νέες κανονιστικές απαιτήσεις και βέλτιστες πρακτικές.

Ενώ η τεχνητή νοημοσύνη προσφέρει σημαντικές δυνατότητες για τη βελτίωση της αποτελεσματικότητας, της ασφάλειας και της ανθεκτικότητας της υποδομής ζωτικής σημασίας, εισάγει επίσης σημαντικές προκλήσεις και κινδύνους. Η αντιμετώπιση αυτών των προκλήσεων απαιτεί μια συνολική προσέγγιση, συμπεριλαμβανομένων ισχυρών μέτρων ασφαλείας, δεοντολογικών πλαισίων, αξιόπιστων και διαφανών συστημάτων τεχνητής νοημοσύνης και

διαρκούς κανονιστικής συμμόρφωσης. Με την προληπτική διαχείριση αυτών των ζητημάτων, οι οργανισμοί μπορούν να αξιοποιήσουν τη μετασχηματιστική δύναμη της τεχνητής νοημοσύνης, προστατεύοντας παράλληλα την ακεραιότητα και τη λειτουργικότητα της υποδομής ζωτικής σημασίας.

Για να αξιοποιηθούν τα οφέλη της τεχνητής νοημοσύνης (AI) στην κυβερνοασφάλεια, ενώ παράλληλα μετριάζονται οι προκλήσεις, είναι απαραίτητη μια ισορροπημένη προσέγγιση. Οι οργανισμοί πρέπει να προτεραιοποιούν την ανθρώπινη εξειδίκευση. Παρόλο που η AI μπορεί να ενισχύσει τις προσπάθειες κυβερνοασφάλειας, η ανθρώπινη εξειδίκευση παραμένει κρίσιμη. Εξειδικευμένοι επαγγελματίες στην κυβερνοασφάλεια πρέπει να συνεργάζονται με τα συστήματα AI για να ερμηνεύουν τα αποτελέσματα, να επικυρώνουν τα ευρήματα και να λαμβάνουν ενημερωμένες αποφάσεις (Zio, 2016).

Απαραίτητη είναι η τακτική παρακολούθηση και αξιολόγηση των μοντέλων και των αλγορίθμων AI για την αναγνώριση και αντιμετώπιση ζητημάτων όπως ψευδώς θετικά, επιθετικές επιθέσεις και προκαταλήψεις. Η συνεχής βελτίωση και προσαρμογή είναι το κλειδί για τη διατήρηση της αποτελεσματικότητας των λύσεων κυβερνοασφάλειας που βασίζονται στην AI. Επιπλέον, η ενθάρρυνση της συνεργασίας και της κοινοποίησης γνώσεων μεταξύ επαγγελματιών κυβερνοασφάλειας, ερευνητών AI και επιστημόνων δεδομένων μπορεί να οδηγήσει σε καινοτόμες λύσεις και βέλτιστες πρακτικές στην κυβερνοασφάλεια που υποστηρίζονται από την AI.

Η προώθηση της διαφάνειας στους αλγορίθμους AI, η γνωστοποίηση των περιορισμών και η διασφάλιση της υπευθυνότητας για τη χρήση τους σε εφαρμογές κυβερνοασφάλειας είναι ζωτικής σημασίας. Αυτό οικοδομεί εμπιστοσύνη μεταξύ των ενδιαφερομένων και προάγει υπεύθυνες και ηθικές πρακτικές AI. Η AI παρουσιάζει σημαντικά οφέλη και προκλήσεις στον τομέα της κυβερνοασφάλειας. Η ικανότητά της να ανιχνεύει απειλές, να αυτοματοποιεί αντιδράσεις, να αναλύει τεράστιους όγκους δεδομένων και να προσαρμόζεται σε εξελισσόμενες απειλές αποτελεί τεράστια υπόσχεση για την ενίσχυση της κυβερνοάμυνάς μας (Markopoulou, D., & Papakonstantinou, 2021).

Ωστόσο, τα πιθανά μειονεκτήματα των ψευδώς θετικών, των επιθετικών επιθέσεων, της έλλειψης εξηγησιμότητας, των απαιτήσεων πόρων και των ηθικών ανησυχιών πρέπει να αντιμετωπιστούν προσεκτικά. Η εύρεση της σωστής ισορροπίας μεταξύ AI και ανθρώπινης εξειδίκευσης, καθώς

και η προώθηση της διαφάνειας, της υπευθυνότητας και της συνεχούς παρακολούθησης, είναι απαραίτητη για την πλήρη αξιοποίηση του δυναμικού της ΑΙ στην κυβερνοασφάλεια, ενώ παράλληλα μετριάζονται οι ενδεχόμενοι κίνδυνοι. Με την υπεύθυνη και προσεκτική αξιοποίηση της ΑΙ, μπορούμε να ενισχύσουμε τη στάση μας απέναντι στην κυβερνοασφάλεια και να παραμείνουμε ένα βήμα μπροστά στο συνεχώς εξελισσόμενο τοπίο απειλών (Taylor et al., 2017).

4.2. Ανάλυση Πλεονεκτημάτων και Προκλήσεων στη χρήση ML σε συστήματα κρίσιμων υποδομών

Η ενσωμάτωση της μηχανικής μάθησης (ML) στα συστήματα υποδομών ζωτικής σημασίας προσφέρει πολλά πλεονεκτήματα. Ένα από τα πρωταρχικά οφέλη είναι η αυξημένη αποδοτικότητα και παραγωγικότητα. Οι αλγόριθμοι ML μπορούν να επεξεργάζονται τεράστιες ποσότητες δεδομένων ταχύτερα και ακριβέστερα από τους ανθρώπους, βελτιστοποιώντας τις λειτουργίες και βελτιώνοντας σημαντικά την αποδοτικότητα. Για παράδειγμα, στη διαχείριση ενέργειας, το ML μπορεί να προβλέψει τις ώρες αιχμής της χρήσης και να προσαρμόσει ανάλογα την κατανομή των πόρων, μειώνοντας τη σπατάλη και βελτιώνοντας την αξιοπιστία των υπηρεσιών.

Ένα άλλο πλεονέκτημα είναι η προγνωστική συντήρηση. Τα μοντέλα ML μπορούν να αναλύσουν δεδομένα από αισθητήρες και άλλες συσκευές παρακολούθησης για να προβλέψουν πότε ο εξοπλισμός είναι πιθανό να αποτύχει. Αυτή η ικανότητα πρόβλεψης επιτρέπει στους οργανισμούς να εκτελούν προληπτική συντήρηση, μειώνοντας τον χρόνο διακοπής λειτουργίας και παρατείνοντας τη διάρκεια ζωής των κρίσιμων στοιχείων υποδομής. Η προσέγγιση αυτή όχι μόνο εξοικονομεί κόστος, αλλά εξασφαλίζει επίσης συνεχή και αξιόπιστη εξυπηρέτηση (Pinto et al., 2023).

Η ML βελτιώνει επίσης τη λήψη αποφάσεων αναλύοντας ιστορικά δεδομένα και δεδομένα σε πραγματικό χρόνο για να παρέχει πληροφορίες που μπορούν να αξιοποιηθούν. Σε κρίσιμες υποδομές, όπως τα δίκτυα μεταφορών, η ML μπορεί να βελτιστοποιήσει τη ροή της κυκλοφορίας, να μειώσει τη συμφόρηση και να ενισχύσει την ασφάλεια προβλέποντας και αντιδρώντας σε πιθανούς κινδύνους σε πραγματικό χρόνο. Επιπλέον, η ML ενισχύει σημαντικά τα μέτρα ασφάλειας στον κυβερνοχώρο ανιχνεύοντας ανωμαλίες και πιθανές απειλές πιο αποτελεσματικά από τις παραδοσιακές μεθόδους. Η συνεχής παρακολούθηση της δραστηριότητας του δικτύου

επιτρέπει στα συστήματα ML να εντοπίζουν και να ανταποκρίνονται γρήγορα στις απειλές στον κυβερνοχώρο, προστατεύοντας τις υποδομές ζωτικής σημασίας από επιθέσεις.

Παρά τα πλεονεκτήματα αυτά, η ανάπτυξη της ML σε υποδομές ζωτικής σημασίας συνοδεύεται από σημαντικές προκλήσεις. Η προστασία της ιδιωτικής ζωής και η ασφάλεια των δεδομένων αποτελούν μείζονες προβληματισμούς, καθώς η συλλογή και η ανάλυση τεράστιων όγκων δεδομένων συχνά περιλαμβάνουν ευαίσθητες πληροφορίες. Η διασφάλιση του απορρήτου και της ασφάλειας αυτών των δεδομένων είναι ζωτικής σημασίας για την αποτροπή λειτουργικών διαταραχών και τη διατήρηση της εμπιστοσύνης του κοινού. Επιπλέον, η εφαρμογή του ML προσθέτει επίπεδα πολυπλοκότητας στα συστήματα, απαιτώντας συνεχή συντήρηση, ενημερώσεις και παρακολούθηση για να διασφαλιστεί η αποτελεσματικότητα. Αυτή η πολυπλοκότητα μπορεί επίσης να καταστήσει την αντιμετώπιση προβλημάτων και την επίλυση ζητημάτων πιο δύσκολη, απαιτώντας εξειδικευμένο εργατικό δυναμικό (Handa et al., 2019).

Η αξιοπιστία και η ανθεκτικότητα αποτελούν πρόσθετες ανησυχίες. Τα μοντέλα ML ενδέχεται να μην αποδίδουν πάντα αξιόπιστα υπό όλες τις συνθήκες, ιδίως σε καταστάσεις για τις οποίες δεν εκπαιδεύτηκαν. Αυτή η έλλειψη ευρωστίας μπορεί να είναι προβληματική σε κρίσιμες υποδομές όπου η αξιοπιστία είναι υψίστης σημασίας. Η διασφάλιση ότι τα μοντέλα ML έχουν δοκιμαστεί διεξοδικά και μπορούν να χειριστούν ακραίες περιπτώσεις και απροσδόκητα σενάρια είναι ζωτικής σημασίας. Επιπλέον, η χρήση της ML εγείρει ηθικά και νομικά ζητήματα σχετικά με τη λογοδοσία και τη διαφάνεια. Οι αποφάσεις που λαμβάνονται από αλγορίθμους ML μπορούν να επηρεάσουν σημαντικά τη δημόσια ασφάλεια και ευημερία, γεγονός που καθιστά αναγκαία την ύπαρξη σαφών κατευθυντήριων γραμμών και κανονιστικών πλαισίων για τη ρύθμιση της χρήσης τους (Sheikh et al., 2022).

Τα ζητήματα διαλειτουργικότητας αποτελούν επίσης προκλήσεις, καθώς τα συστήματα υποδομών ζωτικής σημασίας συχνά περιλαμβάνουν ένα μείγμα παλαιών και σύγχρονων τεχνολογιών. Η ενσωμάτωση λύσεων ML σε υφιστάμενα συστήματα μπορεί να είναι δύσκολη λόγω ζητημάτων συμβατότητας, απαιτώντας σημαντική προσπάθεια και επένδυση για τη διασφάλιση της απρόσκοπτης λειτουργίας. Οι οικονομικές και κοινωνικές επιπτώσεις της ενσωμάτωσης του ML είναι επίσης σημαντικές. Ενώ το ML μπορεί να οδηγήσει σε εξοικονόμηση κόστους και αύξηση της αποδοτικότητας, μπορεί επίσης να οδηγήσει σε εκτόπιση θέσεων εργασίας, απαιτώντας στρατηγικές για την επανεκπαίδευση και την προσαρμογή του εργατικού δυναμικού. Επιπλέον, η

άνιση κατανομή των οφελών του ML μπορεί να επιδεινώσει τις κοινωνικές ανισότητες (Alghamdi, 2020).

Η χρήση της μηχανικής μάθησης σε συστήματα υποδομών ζωτικής σημασίας προσφέρει σημαντικά πλεονεκτήματα, όπως αυξημένη αποδοτικότητα, προληπτική συντήρηση, βελτιωμένη λήψη αποφάσεων και αυξημένη ασφάλεια. Ωστόσο, επιφέρει επίσης σημαντικές προκλήσεις, όπως ανησυχίες σχετικά με την προστασία της ιδιωτικής ζωής και την ασφάλεια των δεδομένων, την πολυπλοκότητα του συστήματος, ζητήματα αξιοπιστίας, ηθικές και νομικές επιπτώσεις, προκλήσεις διαλειτουργικότητας και οικονομικές και κοινωνικές επιπτώσεις. Η αντιμετώπιση αυτών των προκλήσεων απαιτεί μια ισορροπημένη προσέγγιση που περιλαμβάνει ισχυρά μέτρα ασφαλείας, σαφή κανονιστικά πλαίσια, συνεχή παρακολούθηση και στρατηγικές για την προσαρμογή του εργατικού δυναμικού. Με τον τρόπο αυτό, οι οργανισμοί μπορούν να αξιοποιήσουν πλήρως το δυναμικό της ML για να ενισχύσουν την ανθεκτικότητα και την αποδοτικότητα των υποδομών ζωτικής σημασίας, μειώνοντας παράλληλα τους συναφείς κινδύνους (Geetha,& Thilagam, 2021).

Η εκπαίδευση των μοντέλων AI απαιτεί μεγάλο όγκο δεδομένων υψηλής ποιότητας με ετικέτες, τα οποία είναι ακριβά και χρονοβόρα για να συλλεχθούν και να επισημανθούν. Τα συστήματα υποδομής είναι περίπλοκα, οπότε τα μοντέλα AI που χρησιμοποιούνται για την ανάλυσή τους πρέπει να είναι προηγμένα ώστε να αντικατοπτρίζουν με ακρίβεια τις συνθήκες του πραγματικού κόσμου. Η επιτυχής χρήση της μηχανικής μάθησης (ML) στη διαχείριση υποδομών απαιτεί προσεκτικό σχεδιασμό και βαθιά κατανόηση του τρόπου λειτουργίας αυτών των τεχνολογιών για τη βελτίωση της απόδοσης και της αποτελεσματικότητας. Για να αποφευχθούν αυτές οι προκλήσεις, οι μηχανικοί ML και άλλοι τεχνικοί ρόλοι θα πρέπει να επικεντρωθούν στις τεχνικές λεπτομέρειες που προωθούν αυτές τις εξελίξεις. Ακολουθούν ορισμένα τεχνικά χαρακτηριστικά και λειτουργίες που είναι ιδιαίτερα σημαντικά:

Αυτοματοποιημένη εξαγωγή χαρακτηριστικών: Παραδοσιακές προσεγγίσεις όρασης υπολογιστή, όπως η εξαγωγή χαρακτηριστικών μέσω συνελίξεων, επιταχύνουν το στάδιο προεπεξεργασίας και βελτιώνουν τον εντοπισμό ελαττωμάτων σε περιουσιακά στοιχεία υποδομής.

Τεχνικές εμπλουτισμού δεδομένων: Απαραίτητες για τη διαχείριση περιορισμένων συνόλων δεδομένων, οι μέθοδοι εμπλουτισμού δεδομένων, όπως η περιστροφή, η αναστροφή, η κλιμάκωση

κ.λπ., αυξάνουν την ανθεκτικότητα των μοντέλων προσφέροντας ποικιλία αντίστοιχη με τα δεδομένα του πραγματικού κόσμου.

Προβλέψεις σειρών χρόνου και προγνωστική συντήρηση: Χρησιμοποιώντας μοντέλα όπως τα Long Short-Term Memory (LSTM) δίκτυα, οι μηχανικοί ML μπορούν να αναλύσουν δεδομένα αισθητήρων με την πάροδο του χρόνου για να προβλέψουν ανάγκες συντήρησης, επιτρέποντας την προληπτική διαχείριση περιουσιακών στοιχείων.

Αποδοτικότητα μέσω της μεταφοράς μάθησης: Με την προσαρμογή προεκπαιδευμένων μοντέλων σε εργασίες συγκεκριμένων τομέων, οι μηχανικοί ML μπορούν να εξοικονομήσουν πόρους και να επιτύχουν υψηλή ακρίβεια χωρίς να ξεκινούν από το μηδέν, κάτι που είναι ιδιαίτερα χρήσιμο για την ανάλυση εικόνων επιθεώρησης υποδομών.

Ανάλυση σε πραγματικό χρόνο μέσω υπολογιστικής περιβάλλοντος: Η υπολογιστική περιβάλλοντος διευκολύνει την άμεση επεξεργασία και ανάλυση δεδομένων στην πηγή, κάτι που είναι ζωτικής σημασίας για έγκαιρες ενέργειες συντήρησης σε απομακρυσμένα περιουσιακά στοιχεία υποδομής, μειώνοντας την καθυστέρηση και την εξάρτηση από συνεχή συνδεσιμότητα στο cloud. Με τη σωστή διαχείριση αυτών των τεχνικών προκλήσεων, οι οργανισμοί μπορούν να αξιοποιήσουν τα οφέλη της τεχνητής νοημοσύνης για τη βελτίωση της αποδοτικότητας και της ανθεκτικότητας των συστημάτων υποδομής τους.

Συμπεράσματα

Η τεχνητή νοημοσύνη (AI) αποτελεί κρίσιμο στοιχείο σε κάποια από τα πιο πολύπλοκα συστήματα της ανθρωπότητας, ειδικότερα σε εκείνα που είναι ζωτικής σημασίας για την ασφάλεια. Σε τέτοια κρίσιμα συστήματα, το λογισμικό γενικά ελέγχει τη συμπεριφορά των ηλεκτρομηχανικών εξαρτημάτων και παρακολουθεί τις αλληλεπιδράσεις τους. Ωστόσο, η AI χρησιμοποιείται και με πολλούς άλλους τρόπους, όπως στην αντιστοίχιση προτύπων, τη λήψη αποφάσεων, την προγνωστική ανάλυση, την ανίχνευση ανωμαλιών και άλλα.

Σε απλούστερα σενάρια, η AI προσφέρει σημαντικά οφέλη στην αυτοματοποίηση εργασιών που παλαιότερα απαιτούσαν ανθρώπινη παρέμβαση, όπως την ανάλυση τεράστιων όγκων δεδομένων για την εξαγωγή πληροφοριών που οδηγούν σε αποφάσεις. Μάλιστα, αν είναι κατάλληλα εκπαιδευμένη, η AI μπορεί να αναλάβει και τη λήψη αυτών των αποφάσεων. Αν και η εφαρμογή της μπορεί να γίνει σε υλικό, υλικολογισμικό ή λογισμικό, ο σχεδιασμός, η υλοποίηση και οι δοκιμές της πρέπει να διέπονται από αυστηρά πρότυπα ασφαλείας και αξιοπιστίας. Η AI για κρίσιμα συστήματα οφείλει επίσης να συνδυάζει την ανάλυση σε πραγματικό χρόνο με ισχυρές δομές επικοινωνίας που προσαρμόζονται στις μεταβαλλόμενες συνθήκες.

Η σημερινή AI διαφέρει από το γενικό λογισμικό κυρίως λόγω της ανάγκης για εκπαίδευση των αλγορίθμων, με την προοπτική της αυτομάθησης και βαθύτερης κατανόησης. Οι εκπαιδευμένοι αλγόριθμοι, που λειτουργούν ως "μαύρα κουτιά", εισάγουν μια έλλειψη διαφάνειας στη λειτουργία τους. Αυτή η εκπαίδευση μπορεί να περιλαμβάνει προκαταλήψεις, εξαιτίας της ποιότητας των δεδομένων που χρησιμοποιούνται. Συνεπώς, στην τεχνητή νοημοσύνη υπάρχει μεγαλύτερη επιτακτικότητα για ηθικές σκοπιμότητες σε σύγκριση με το παραδοσιακό λογισμικό.

Σε κρίσιμα συστήματα, οι εσωτερικές και εξωτερικές αλληλεπιδράσεις, οι χρονισμοί και τα γενικά σφάλματα επεξεργασίας μπορούν να οδηγήσουν το λογισμικό σε μη ασφαλείς καταστάσεις ή ακόμα και σε αποτυχίες του συστήματος. Η τεχνητή νοημοσύνη (AI) μπορεί να ενσωματωθεί για την αποφυγή ή την ανάκαμψη από τέτοιες καταστάσεις. Σε περιπτώσεις όπου η AI σε κρίσιμα συστήματα δεν λειτουργεί ως αναμενόμενο—είτε αποτυγχάνοντας να προλάβει μια αποτυχία είτε συμβάλλοντας σε αυτή—οι συνέπειες μπορεί να είναι σοβαρές. Αυτές κυμαίνονται από μικρές ανωμαλίες στην απόδοση έως καταστροφικές αποτυχίες που προκαλούν σημαντικές οικονομικές ζημιές, υλικές απώλειες, τραυματισμούς ή ακόμα και ανθρώπινες απώλειες.

Για αυτό το λόγο, είναι κρίσιμο να υπάρχει λογισμικό ΑΙ που παρέχει εξηγήσιμα αποτελέσματα. Οι συστάσεις που παρέχονται από το σύστημα πρέπει να είναι προβλέψιμες και να επαναλαμβάνονται αξιόπιστα υπό μια ευρεία ποικιλία εισροών, προκαταλήψεων και αποτελεσμάτων.

Η ανάλυση των υφιστάμενων εργαλείων τεχνητής νοημοσύνης (ΑΙ) στον τομέα των κρίσιμων υποδομών αποτελεί ένα σημαντικό πεδίο μελέτης, καθώς η ενσωμάτωση τέτοιων τεχνολογιών συμβάλλει στη βελτίωση της αποτελεσματικότητας, της ασφάλειας και της ανθεκτικότητας των υποδομών αυτών. Οι εφαρμογές ΑΙ στον τομέα αυτό καλύπτουν ένα ευρύ φάσμα λειτουργιών, από την πρόγνωση και την προειδοποίηση για πιθανά προβλήματα έως την αυτόματη ανταπόκριση σε επείγουσες καταστάσεις.

Πρώτον, στη διαχείριση ενεργειακών δικτύων, η ΑΙ χρησιμοποιείται για την ανάλυση τεράστιων όγκων δεδομένων προκειμένου να προβλέψει ζητήματα ζήτησης και προσφοράς ενέργειας, βοηθώντας στην οπтимιστική διαχείριση πόρων. Επιπλέον, συστήματα ΑΙ συνεργάζονται με ανιχνευτές και άλλα ηλεκτρονικά συστήματα για την ανίχνευση και την επισκευή βλαβών σε πραγματικό χρόνο. Δεύτερον, στον τομέα των μεταφορών, τα ΑΙ συστήματα ενισχύουν την ασφάλεια και την αποδοτικότητα μέσω της ανάλυσης δεδομένων κίνησης και της διαχείρισης συστημάτων ελέγχου τροχαίας κίνησης. Αυτό επιτρέπει την ταχύτερη ανταπόκριση σε απρόσμενες καταστάσεις και τη βελτίωση της ροής της κίνησης. Στον τομέα της υγείας, η ΑΙ χρησιμοποιείται για την ανάλυση ιατρικών δεδομένων και την αναγνώριση πιθανών απειλών για τη δημόσια υγεία, παρέχοντας τη δυνατότητα για προληπτικές δράσεις και γρήγορη ανταπόκριση σε επιδημίες. Η ένταξη των εργαλείων ΑΙ στον τομέα των κρίσιμων υποδομών έχει άμεσο αντίκτυπο στην ενίσχυση της ασφάλειας, μέσω της βελτίωσης της παρακολούθησης και της αντίδρασης σε απειλές. Η συνεχής εξέλιξη και βελτίωση των ΑΙ εργαλείων αναμένεται να οδηγήσει σε ακόμη πιο αποτελεσματική και αξιόπιστη διαχείριση των κρίσιμων υποδομών στο μέλλον.

Οι κυριότεροι κίνδυνοι για την ταχεία ανάπτυξη της τεχνητής νοημοσύνης περιλαμβάνουν την αργή υλοποίηση των προσδοκώμενων οφελών και την κοινωνική και κανονιστική επιφυλακτικότητα. Υπερβολικές προσδοκίες και παραπλανητικές αντιλήψεις σχετικά με τις πραγματικές δυνατότητες και κινδύνους, επηρεασμένες από ταινίες επιστημονικής φαντασίας και βιβλία, καθώς και αναρτήσεις στα μέσα κοινωνικής δικτύωσης, μπορούν να προκαλέσουν αβάσιμο φόβο και αβεβαιότητα.

Συστήματα τεχνητής νοημοσύνης με αυτογνωσία, όπως το HAL από την ταινία "2001, Οδύσσεια του Διαστήματος", δημιουργούν μια αναληθή εικόνα για την τεχνητή νοημοσύνη στο κοινό. Μια σοβαρή καταστροφή θα μπορούσε να οδηγήσει σε κλήσεις για αυστηρό περιορισμό της χρήσης της τεχνητής νοημοσύνης στα κρίσιμα συστήματα.

Ένα άλλο ζήτημα είναι η εύρεση κατάλληλων δεδομένων εκπαίδευσης όπως βιομετρικά δεδομένα, πληροφορίες συμπεριφοράς και μοτίβα χρήσης για την ενίσχυση των συστημάτων τεχνητής νοημοσύνης. Ζητήματα απορρήτου, σχετικά με τα δεδομένα αυτά, θα μπορούσαν να εμποδίσουν την πρόοδο και ανάπτυξη.

Τέλος, τα προβλήματα ενσωμάτωσης παλαιών συστημάτων, η υπερφόρτωση προτύπων και η σύγχυση μπορεί να επιβραδύνουν την πρόοδο. Η τεχνητή νοημοσύνη για κρίσιμα συστήματα θα απαιτήσει επίσης στενό συντονισμό μεταξύ της βιομηχανίας και των ρυθμιστικών αρχών. Θα χρειαστεί να δοθεί μεγαλύτερη προσοχή στις επενδύσεις από το κράτος, την κοινότητα, τα πανεπιστήμια και τη βιομηχανία για συνεργασίες και για τη διασφάλιση της επαγγελματικής ευθύνης και αξιοπιστίας αυτών που κατασκευάζουν τα συστήματα αυτά, προκειμένου να αξιοποιηθούν οι πραγματικές τους δυνατότητες.

Επίλογος

Η επανάσταση στον τομέα της τεχνητής νοημοσύνης (AI) προσφέρει σημαντικές ευκαιρίες για την αξιοποίηση της ψηφιοποίησης και μπορεί να ενεργοποιήσει την προσέγγιση "συστημάτων" που απαιτείται για πιο πολύπλοκες υποδομές. Αυτό το έγγραφο εξετάζει τον τρόπο με τον οποίο η έρευνα στην οικονομική υποδομή συνδυάζεται με την τεχνητή νοημοσύνη, εστιάζοντας στις μεθόδους και τους σκοπούς της εφαρμογής της AI σε αυτούς τους τομείς. Η μηχανική μάθηση είναι η κυρίαρχη τεχνική, με μεθόδους όπως τα τεχνητά νευρωνικά δίκτυα, οι μηχανές υποστήριξης διανυσμάτων και τα τυχαία δάση να ξεχωρίζουν. Επιπλέον, η αυτοματοποιημένη τεχνική συλλογιστικής ασαφούς λογικής έχει ευρεία εφαρμογή λόγω της ικανότητάς της να αντιμετωπίζει αβεβαιότητες. Στους τομείς της ενέργειας, του νερού, των μεταφορών και των τηλεπικοινωνιών, η τεχνητή νοημοσύνη χρησιμοποιείται για διάφορους σκοπούς, όπως η πρόβλεψη, η δρομολόγηση, η συντήρηση και η διαχείριση της ποιότητας. Παρόλα αυτά, πρέπει να υπάρξει μεγαλύτερη συνεργασία και ενοποίηση στο σχεδιασμό και την πολιτική προκειμένου να επιτευχθούν λύσεις που θα αντιμετωπίζουν τις πραγματικές ανάγκες των υποδομών.

Η εντατικοποίηση της ενσωμάτωσης προηγμένων τεχνολογιών σε υποδομές κρίσιμης σημασίας επιφέρει πολλαπλά οφέλη. Παρ' όλα αυτά, η αυξανόμενη εξάρτηση από δίκτυα επικοινωνίας καλλιεργεί αυξημένες ανησυχίες όσον αφορά την ασφάλεια. Επιπρόσθετα, η χρήση κυβερνοφυσικών συστημάτων σε ζωτικές υποδομές, όπως η υγειονομική περίθαλψη, οι μεταφορές και η παραγωγή ενέργειας, εγείρει σοβαρές προκλήσεις που ενισχύουν τις ανησυχίες για την ασφάλεια. Με την εξέλιξη της τεχνολογίας προς δικτυωμένες συσκευές, η ασφάλεια αυτών των φυσικών συστημάτων στον κυβερνοχώρο έχει καταστεί κεντρικό ζήτημα. Κατά συνέπεια, πλήθος μέτρων ασφαλείας και αμυντικών μηχανισμών έχουν εισαχθεί για την προστασία από αυτά τα ζητήματα.

Με την αυτοματοποίηση των καθηκόντων ρουτίνας και την παροχή υποστήριξης λήψης αποφάσεων μέσω των γνώσεων της τεχνητής νοημοσύνης, οι τεχνολογίες αυτές δίνουν στους ανθρώπους τη δυνατότητα να επικεντρωθούν σε καθήκοντα υψηλότερου επιπέδου που απαιτούν δημιουργικότητα, κριτική σκέψη και λήψη στρατηγικών αποφάσεων.

Αντί να φοβόμαστε τον εκτοπισμό, είναι ζωτικής σημασίας να κατανοήσουμε πώς η τεχνητή νοημοσύνη και η αυτοματοποίηση χρησιμεύουν ως εργαλεία για την ανύψωση των ανθρώπινων ικανοτήτων και την ενίσχυση της ανθεκτικότητας και της αποτελεσματικότητας των υποδομών

ζωτικής σημασίας. Για τα κρίσιμα συστήματα, τα μοντέλα TN μπορούν να χρησιμοποιηθούν στη συλλογιστική της γνώσης, τον προγραμματισμό, την επεξεργασία φυσικής γλώσσας, την όραση υπολογιστών, τη ρομποτική και την TN, καθιστώντας τα ιδιαίτερα κατάλληλα για ένα ευρύ φάσμα εφαρμογών κρίσιμων συστημάτων. Η τεχνητή νοημοσύνη μπορεί να βελτιώσει τη βελτιστοποίηση των επιδόσεων, την ασφαλή λειτουργία, την πρόληψη και ανίχνευση απάτης και εισβολής, την πρόγνωση της υγείας και την ανάλυση μετά την αποτυχία.

Η τεχνητή νοημοσύνη μπορεί να διαδραματίσει καθοριστικό ρόλο στην εξάλειψη των επαχθών, τετριμμένων καθηκόντων σε αυτά τα περίπλοκα κρίσιμα για την ασφάλεια πλαίσια. Η αυτοματοποίηση με βάση την TN μπορεί να βελτιώσει τις εργασίες ρουτίνας, όπως η συλλογή δεδομένων, η ανάλυση και η υποβολή εκθέσεων. Οι αλγόριθμοι προληπτικής συντήρησης μπορούν να εντοπίζουν προληπτικά τις βλάβες του εξοπλισμού πριν από την εμφάνισή τους, ελαχιστοποιώντας τον χρόνο διακοπής λειτουργίας και βελτιστοποιώντας τη χρήση των πόρων.

Η αυτοματοποίηση μπορεί επίσης να βελτιώσει τη διαχείριση αποθεμάτων, τις διαδικασίες προμηθειών και τις εργασίες κανονιστικής συμμόρφωσης, απελευθερώνοντας ανθρώπινους πόρους για πιο στρατηγικές προσπάθειες. Οι αλγόριθμοι βελτιστοποίησης με τεχνητή νοημοσύνη μπορούν να βελτιώσουν την αποδοτικότητα των δικτύων διανομής ενέργειας, των συστημάτων μεταφοράς και των αλυσίδων εφοδιασμού, μειώνοντας τη σπατάλη και το κόστος και ενισχύοντας παράλληλα την αξιοπιστία. Αναθέτοντας επαναλαμβανόμενες εργασίες σε ΑΙ και αυτοματοποιημένα συστήματα, τα περιβάλλοντα OT μπορούν να μειώσουν το λειτουργικό κόστος και να βελτιώσουν τη συνολική παραγωγικότητα και την ευελιξία λήψης αποφάσεων, ξεκλειδώνοντας νέα επίπεδα αποδοτικότητας και ανταγωνιστικότητας. Επιπλέον, τα εργαλεία εκτίμησης κινδύνου με βάση την TN μπορούν να βοηθήσουν στην ιεράρχηση των επενδύσεων σε αναβαθμίσεις υποδομών και μέτρα ανθεκτικότητας για τον μετριασμό των επιπτώσεων φυσικών καταστροφών και άλλων καταστάσεων έκτακτης ανάγκης.

Ωστόσο, αυτά τα κρίσιμα συστήματα αντιμετωπίζουν πρωτοφανείς προκλήσεις στην ψηφιακή εποχή. Η εξάρτησή μας από τις κρίσιμες υποδομές πρέπει να ληφθεί υπόψη, και η διασύνδεσή τους μεγεθύνει τον πιθανό αντίκτυπο των διαταραχών, είτε αυτές προκαλούνται από φυσικές καταστροφές, απειλές στον κυβερνοχώρο ή ανθρώπινα λάθη. Οι παραδοσιακές μέθοδοι διαχείρισης των υποδομών ζωτικής σημασίας χρειάζονται συχνά βοήθεια για να συμβαδίσουν με τις σύγχρονες προκλήσεις. Η γήρανση των υποδομών, η αυξανόμενη ζήτηση, οι απειλές για την

κυβερνοασφάλεια και η ανάγκη για βιωσιμότητα θέτουν πολύπλοκες προκλήσεις. Επιπλέον, το ανθρώπινο εργατικό δυναμικό που υποστηρίζει αυτά τα συστήματα πρέπει να βελτιωθεί όσον αφορά την ικανότητα, την ταχύτητα και την ακρίβεια, ιδίως σε καταστάσεις υψηλής πίεσης ή όταν πρόκειται για τεράστιες ποσότητες δεδομένων.

Καμία συζήτηση για την τεχνητή νοημοσύνη δεν μπορεί να γίνει χωρίς να συζητηθούν οι ηθικές επιπτώσεις. Η ευρεία υιοθέτηση της τεχνητής νοημοσύνης και της αυτοματοποίησης στις κρίσιμες υποδομές εγείρει ηθικές και κοινωνικές επιπτώσεις που πρέπει να εξεταστούν προσεκτικά. Η προστασία της ιδιωτικής ζωής των δεδομένων, η μεροληψία των αλγορίθμων και οι ανησυχίες σχετικά με τη μετατόπιση θέσεων εργασίας απαιτούν ισχυρά ρυθμιστικά πλαίσια και ηθικές κατευθυντήριες γραμμές για να διασφαλιστεί ότι οι τεχνολογίες αυτές αναπτύσσονται υπεύθυνα και δίκαια. Επιπλέον, η παροχή πρόσβασης σε ευκαιρίες κατάρτισης και εκπαίδευσης είναι απαραίτητη για τον μετριασμό του κινδύνου διεύρυνσης των κενών δεξιοτήτων και των κοινωνικοοικονομικών ανισοτήτων. Τελικά, η εισαγωγή της τεχνητής νοημοσύνης και της αυτοματοποίησης σε πολύπλοκα συστήματα μπορεί να έχει απρόβλεπτες συνέπειες που είναι δύσκολο να προβλεφθούν. Ακόμη και οι καλά σχεδιασμένοι αλγόριθμοι μπορεί να παρουσιάσουν απροσδόκητες συμπεριφορές ή αλληλεπιδράσεις με άλλα στοιχεία του συστήματος, οδηγώντας ενδεχομένως σε δυσλειτουργίες ή διαταραχές που δεν είχαν προβλεφθεί κατά τη φάση του σχεδιασμού.

Η αντιμετώπιση αυτών των προκλήσεων απαιτεί μια πολύπλευρη προσέγγιση που συνδυάζει ισχυρά μέτρα κυβερνοασφάλειας, αυστηρές δοκιμές και επικύρωση αλγορίθμων τεχνητής νοημοσύνης, διαφάνεια στις διαδικασίες λήψης αποφάσεων και προληπτικές στρατηγικές διαχείρισης κινδύνων. Με την προσεκτική εξέταση των πιθανών κινδύνων και την εφαρμογή των κατάλληλων διασφαλίσεων, οι φορείς εκμετάλλευσης υποδομών ζωτικής σημασίας μπορούν να αξιοποιήσουν τη δύναμη της ΤΝ και της αυτοματοποίησης για να ενισχύσουν την ανθεκτικότητα και την αποτελεσματικότητα των υποδομών ζωτικής σημασίας, ελαχιστοποιώντας παράλληλα τους συναφείς κινδύνους.

Ο ρόλος της ΤΝ και της αυτοματοποίησης στις υποδομές ζωτικής σημασίας θα συνεχίσει να εξελίσσεται, με γνώμονα τις τεχνολογικές εξελίξεις και τις μεταβαλλόμενες κοινωνικές ανάγκες. Από τη μεταποίηση και τα έξυπνα δίκτυα έως τα συστήματα υγειονομικής περίθαλψης με τεχνητή νοημοσύνη, οι πιθανές εφαρμογές αυτών των τεχνολογιών είναι τεράστιες και ποικίλες. Ωστόσο,

είναι σημαντικό να αναγνωρίσουμε ότι ο απώτερος στόχος δεν είναι η αντικατάσταση των ανθρώπων με μηχανές, αλλά η αξιοποίηση της τεχνολογίας για την ενίσχυση των ανθρώπινων ικανοτήτων και τη δημιουργία πιο ανθεκτικών, αποδοτικών και βιώσιμων συστημάτων υποδομής προς όφελος της κοινωνίας.

Οι προκλήσεις που συνδέονται με την ασφάλεια των κυβερνοφυσικών συστημάτων είναι πολυδιάστατες. Η διαφορετικότητα των απειλών και των ευπαθειών καθιστά τις συμβατικές τεχνικές ανεπαρκείς. Η ποικιλομορφία των τύπων των συστημάτων, καθώς και η ποικιλία των κυβερνοεπιθέσεων που μπορεί να τους προσβάλλουν, απαιτούν πιο ολοκληρωμένες τεχνικές, ικανές να ανιχνεύουν απειλές διαφορετικής φύσης. Αυτές οι ολοκληρωμένες μέθοδοι ανίχνευσης πρέπει να επικεντρωθούν στην εφαρμογή τεχνητής νοημοσύνης για την ακριβή ταξινόμηση των απειλών. Τα μη ευφυή συστήματα, που βασίζονται σε προκαθορισμένα μοντέλα, απαιτούν σημαντική πολυπλοκότητα και δυσκολεύονται να επιτύχουν τα απαραίτητα αποτελέσματα, κάτι που καθιστά την ΤΝ ιδανική λύση για την ασφάλεια των κυβερνοφυσικών συστημάτων και των κρίσιμων υποδομών.

Βιβλιογραφικές Αναφορές

- Al Dabbagh, A., Al-Sabbagh, M., & Alkhabbaz, A. (2021) Artificial Intelligence and Machine Learning in Cybersecurity: A Systematic Review.
- Alcaraz, C., & Lopez, J. (2013). Wide-area situational awareness for critical infrastructure protection. *Computer*, 46(4), 30-37.
- Alghamdi, M. I. (2020). Survey on Applications of Deep Learning and Machine Learning Techniques for Cyber Security. *International Journal of Interactive Mobile Technologies*, 14(16).
- Alkhaleel, B. A., Liao, H., & Sullivan, K. M. (2022). Risk and resilience-based optimal post-disruption restoration for critical infrastructures under uncertainty. *European Journal of Operational Research*, 296(1), 174-202.
- Ampratwum, G., Tam, V. W., & Osei-Kyei, R. (2023). Critical analysis of risks factors in using public-private partnership in building critical infrastructure resilience: A systematic review. *Construction Innovation*, 23(2), 360-382.
- Argyropoulos, N., Khodashenas, P. S., Mavropoulos, O., Karapistoli, E., Lytos, A., Karypidis, P. A., & Hofmann, K. P. (2021). Addressing cybersecurity in the next generation mobility ecosystem with CARMEL. *Transportation Research Procedia*, 52, 307-314.
- Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine*, 34(6), 26-38.
- Bahrammirzaee, A. (2010). A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. *Neural Computing and Applications*, 19(8), 1165-1195.
- Bakalos, N., Voulodimos, A., Doulamis, N., Doulamis, A., Ostfeld, A., Salomons, E., ... & Li, P. (2019). Protecting water infrastructure from cyber and physical threats: Using multimodal data fusion and adaptive deep learning to monitor critical systems. *IEEE Signal Processing Magazine*, 36(2), 36-48.
- Bécue, A., Praça, I., & Gama, J. (2021). Artificial intelligence, cyber-threats and Industry 4.0: Challenges and opportunities. *Artificial Intelligence Review*, 54(5), 3849-3886.

- Bello, A., Jahan, S., Farid, F., & Ahamed, F. (2022). A systemic review of the cybersecurity challenges in Australian water infrastructure management. *Water*, 15(1), 168.
- Berghout, T., Benbouzid, M., & Muyeen, S. M. (2022). Machine learning for cybersecurity in smart grids: A comprehensive review-based study on methods, solutions, and prospects. *International Journal of Critical Infrastructure Protection*, 38, 100547.
- Burbeck, K., & Nadjm-Tehrani, S. (2007). Adaptive real-time anomaly detection with incremental clustering. *information security technical report*, 12(1), 56-67.
- Carcano, A., Fovino, I. N., Masera, M., & Trombetta, A. (2010). State-based network intrusion detection systems for SCADA protocols: a proof of concept. In *Critical Information Infrastructures Security: 4th International Workshop, CRITIS 2009, Bonn, Germany, September 30-October 2, 2009. Revised Papers 4* (pp. 138-150). Springer Berlin Heidelberg.
- Chang, C. S., Lin, Y. S., & Lin, C. Y. (2021) AI-Enabled IT/OT Convergence: A Review of Trends, Challenges and Opportunities.
- Chertoff, M. (2008). One Team, One Mission, Securing Our Homeland: US Department of Homeland Security Strategic Plan, Fiscal Years 2008-2013. *Washington, DC Retrieved July, 12, 2009*.
- Cheung, S., Dutertre, B., Fong, M., Lindqvist, U., Skinner, K., & Valdes, A. (2007, January). Using model-based intrusion detection for SCADA networks. In *Proceedings of the SCADA security scientific symposium* (Vol. 46, pp. 1-12). SRI International.
- D'Antonio, S., Oliviero, F., & Setola, R. (2006). High-speed intrusion detection in support of critical infrastructure protection. In *Critical Information Infrastructures Security: First International Workshop, CRITIS 2006, Samos, Greece, August 31-September 1, 2006. Revised Papers 1* (pp. 222-234). Springer Berlin Heidelberg.
- Das, T. K., Adepu, S., & Zhou, J. (2020). Anomaly detection in industrial control systems using logical analysis of data. *Computers & Security*, 96, 101935.
- Deng, L., & Yu, D. (2014). Deep learning: methods and applications. *Foundations and trends® in signal processing*, 7(3-4), 197-387.

Dussel, P., Gehl, C., Laskov, P., Bußer, J., Störmann, C., Kasten, J.: Cyber- Critical Infrastructure Protection using Real-Time Payload-Based Anomaly Detection. *Critical Information Infrastructures Security* (2010) 85–97

Eom, J. H., Kim, N. U., Kim, S. H., & Chung, T. M. (2012, June). Cyber military strategy for cyberspace superiority in cyber warfare. In *Proceedings title: 2012 international conference on cyber security, cyber warfare and digital forensic (cybersec)* (pp. 295-299). IEEE.

Ermolieva, T., & Ermoliev, Y. (2005). *Catastrophic Risk Management: Flood and. Applications of Stochastic Programming*.

European Commission: COM(2011) 163 Achievements and Next Steps: Towards Global Cyber-Security. Publications Office (2011)

Fang, X., Misra, S., Xue, G., & Yang, D. (2011). Smart grid—The new and improved power grid: A survey. *IEEE communications surveys & tutorials*, 14(4), 944-980.

Geetha, R., & Thilagam, T. (2021). A review on the effectiveness of machine learning and deep learning algorithms for cyber security. *Archives of Computational Methods in Engineering*, 28(4), 2861-2879.

Graham, S., & Marvin, S. (1999). Planning cybercities? Integrating telecommunications into urban planning. *The Town Planning Review*, 89-114.

Groth, M., & Cortekar, J. (2015, April). Is there a need for government interventions to adapt energy infrastructures to climate change? A German case study. *EGU General Assembly Conference*

Guzman, A., Ishida, S., Choi, E., & Aoyama, A. (2016, December). Artificial intelligence improving safety and risk analysis: A comparative analysis for critical infrastructure. In *2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)* (pp. 471-475). IEEE.

Hadiosmanovic, D., Bolzoni, D., Hartel, P., & Etalle, S. (2011, September). MELISSA: Towards automated detection of undesirable user actions in critical infrastructures. In *2011 Seventh European Conference on Computer Network Defense* (pp. 41-48). IEEE.

- Haghighi, M. S., Farivar, F., & Jolfaei, A. (2020). A Machine-Learning-Based Approach to Build Zero-False-Positive IPSs for Industrial IoT and CPS With a Case Study on Power Grids Security. *IEEE Transactions on Industry Applications*, 60(1), 920-928.
- Handa, A., Sharma, A., & Shukla, S. K. (2019). Machine learning in cybersecurity: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(4), e1306.
- He, Y., Luo, C., Zheng, J., Wang, K., & Zhang, H. (2022, September). AI Based Directory Discovery Attack and Prevention of the Medical Systems. In *2022 Computing in Cardiology (CinC)* (Vol. 498, pp. 1-4). IEEE.
- Hollands, R. G. (2020). Will the real smart city please stand up?: Intelligent, progressive or entrepreneurial?. In *The Routledge companion to smart cities* (pp. 179-199). Routledge.
- Kaloudi, N., & Li, J. (2020). The ai-based cyber threat landscape: A survey. *ACM Computing Surveys (CSUR)*, 53(1), 1-34.
- Karimipour, H., & Dinavahi, V. (2015, May). Parallel domain decomposition based distributed state estimation for large-scale power systems. In *2015 IEEE/IAS 51st Industrial & Commercial Power Systems Technical Conference (I&CPS)* (pp. 1-5). IEEE.
- Karimipour, H., & Dinavahi, V. (2017, August). On false data injection attack against dynamic state estimation on smart power grids. In *2017 IEEE International Conference on Smart Energy Grid Engineering (SEGE)* (pp. 388-393). IEEE.
- Karimipour, H., Dehghantanha, A., Parizi, R. M., Choo, K. K. R., & Leung, H. (2019). A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids. *Ieee Access*, 7, 80778-80788.
- Kayan, H., Nunes, M., Rana, O., Burnap, P., & Perera, C. (2022). Cybersecurity of industrial cyber-physical systems: A review. *ACM Computing Surveys (CSUR)*, 54(11s), 1-35.
- Khaw, Y. M., Jahromi, A. A., Arani, M. F., Sanner, S., Kundur, D., & Kassouf, M. (2020). A deep learning-based cyberattack detection system for transmission protective relays. *IEEE Transactions on Smart Grid*, 12(3), 2554-2565.

- Kiwiia, D., Dehghantanha, A., Choo, K. K. R., & Slaughter, J. (2018). A cyber kill chain-based taxonomy of banking Trojans for evolutionary computational intelligence. *Journal of computational science*, 27, 394-409.
- Korkmaz, C., & Correia, A. P. (2019). A review of research on machine learning in educational technology. *Educational Media International*, 56(3), 250-267.
- Lee, I., & Sokolsky, O. (2010, June). Medical cyber physical systems. In *Proceedings of the 47th design automation conference* (pp. 743-748).
- Lehto, M. (2020). Cyber security in aviation, maritime and automotive. *Computation and Big Data for Transport: Digital Innovations in Surface and Air Transport Systems*, 19-32.
- Lin, H., Slagell, A., Di Martino, C., Kalbarczyk, Z., & Iyer, R. K. (2013, January). Adapting bro into scada: building a specification-based intrusion detection system for the dnp3 protocol. In *Proceedings of the Eighth Annual Cyber Security and Information Intelligence Research Workshop* (pp. 1-4).
- Malatji, M., & Tolah, A. (2024). Artificial intelligence (AI) cybersecurity dimensions: a comprehensive framework for understanding adversarial and offensive AI. *AI and Ethics*, 1-28.
- Markopoulou, D., & Papakonstantinou, V. (2021). The regulatory framework for the protection of critical infrastructures against cyberthreats: Identifying shortcomings and addressing future challenges: The case of the health sector in particular. *Computer law & security review*, 41, 105502.
- Milenković, A., Otto, C., & Jovanov, E. (2006). Wireless sensor networks for personal health monitoring: Issues and an implementation. *Computer communications*, 29(13-14), 2521-2533.
- Mirsky, Y., Demontis, A., Kotak, J., Shankar, R., Gelei, D., Yang, L., ... & Biggio, B. (2023). The threat of offensive ai to organizations. *Computers & Security*, 124, 103006.
- Pinto, A., Herrera, L. C., Donoso, Y., & Gutierrez, J. A. (2023). Survey on intrusion detection systems based on machine learning techniques for the protection of critical infrastructure. *Sensors*, 23(5), 2415.

- Pinto, A., Herrera, L. C., Donoso, Y., & Gutierrez, J. A. (2023). Survey on intrusion detection systems based on machine learning techniques for the protection of critical infrastructure. *Sensors*, 23(5), 2415.
- Raciti, M., & Nadjm-Tehrani, S. (2013). Embedded cyber-physical anomaly detection in smart meters. In *Critical Information Infrastructures Security: 7th International Workshop, CRITIS 2012, Lillehammer, Norway, September 17-18, 2012, Revised Selected Papers* (pp. 34-45). Springer Berlin Heidelberg.
- Radanliev, P., & De Roure, D. (2022). Advancing the cybersecurity of the healthcare system with self-optimising and self-adaptative artificial intelligence (part 2). *Health and Technology*, 12(5), 923-929.
- Ramos, J., Clavelli, L., & Rossi, L. F. (2020) The Role of Artificial Intelligence in Predicting and Preventing Cyber Attacks.
- Reeves, J., Ramaswamy, A., Locasto, M., Bratus, S., & Smith, S. (2012). Intrusion detection for resource-constrained embedded control systems in the power grid. *International Journal of Critical Infrastructure Protection*, 5(2), 74-83.
- Roosta, T., Nilsson, D. K., Lindqvist, U., & Valdes, A. (2008, September). An intrusion detection system for wireless process control systems. In *2008 5th IEEE international conference on mobile ad hoc and sensor systems* (pp. 866-872). IEEE.
- Ruzbahani, H. M., Rahimnejad, A., & Karimipour, H. (2019, February). Smart households demand response management with micro grid. In *2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)* (pp. 1-5). IEEE.
- Sakhnini, J., Karimipour, H., Dehghantanha, A., Yazdinejad, A., Gadekallu, T. R., Victor, N., & Islam, A. (2023). A generalizable deep neural network method for detecting attacks in industrial cyber-physical systems. *IEEE Systems Journal*.
- Scarfone, K., Mell, P., (2012).: Guide to Intrusion Detection and Prevention Systems (IDPS). NIST Special Publication SP 800-94

- Sheikh, Z. A., Singh, Y., Singh, P. K., & Ghafoor, K. Z. (2022). Intelligent and secure framework for critical infrastructure (CPS): Current trends, challenges, and future scope. *Computer Communications*, 193, 302-331.
- Shin, G., Yooun, H., Shin, D., & Shin, D. (2018). Hybrid Feature Selection Method Based on a Naïve Bayes Algorithm that Enhances the Learning Speed while Maintaining a Similar Error Rate in Cyber ISR. *KSII Transactions on Internet and Information Systems (TIIS)*, 12(12), 5685-5700.
- Shin, G., Yooun, H., Shin, D., & Shin, D. (2018). Incremental learning method for cyber intelligence, surveillance, and reconnaissance in closed military network using converged IT techniques. *Soft Computing*, 22(20), 6835-6844.
- Sobien, D., Yardimci, M. O., Nguyen, M. B., Mao, W. Y., Fordham, V., Rahman, A., ... & Batarseh, F. A. (2023). AI for cyberbiosecurity in water systems—A survey. In *Cyberbiosecurity: A new field to deal with emerging threats* (pp. 217-263). Cham: Springer International Publishing.
- Spring, J. M., Galyardt, A., Householder, A. D., & VanHoudnos, N. (2020, October). On managing vulnerabilities in AI/ML systems. In *Proceedings of the New Security Paradigms Workshop 2020* (pp. 111-126).
- Such, M. C., & Hill, C. (2012, January). Battery energy storage and wind energy integrated into the Smart Grid. In *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)* (pp. 1-4). IEEE.
- Taylor, J. M., & Sharif, H. R. (2017, May). Security challenges and methods for protecting critical infrastructure cyber-physical systems. In *2017 International conference on selected topics in mobile and wireless networking (MoWNeT)* (pp. 1-6). IEEE.
- Tizghadam, A., Khazaei, H., Moghaddam, M. H., & Hassan, Y. (2019). Machine learning in transportation. *Journal of Advanced Transportation*, 2019.
- Wang, S., Gu, X., Luan, S., & Zhao, M. (2021). Resilience analysis of interdependent critical infrastructure systems considering deep learning and network theory. *International Journal of Critical Infrastructure Protection*, 35, 100459.

- Witten, I. H., Frank, E., Hall, M. A., Pal, C. J., & Data, M. (2005, June). Practical machine learning tools and techniques. In *Data mining* (Vol. 2, No. 4, pp. 403-413). Amsterdam, The Netherlands: Elsevier.
- Witten, I. H., Frank, E., Hall, M. A., Pal, C. J., & Data, M. (2005, June). Practical machine learning tools and techniques. In *Data mining* (Vol. 2, No. 4, pp. 403-413). Amsterdam, The Netherlands: Elsevier.
- Xu, Z., & Saleh, J. H. (2021). Machine learning for reliability engineering and safety applications: Review of current status and future opportunities. *Reliability Engineering & System Safety*, 211, 107530.
- Yang, D., Usynin, A., & Hines, J. W. (2006, November). Anomaly-based intrusion detection for SCADA systems. In *5th intl. topical meeting on nuclear plant instrumentation, control and human machine interface technologies (npic&hmit 05)* (pp. 12-16).
- Yang, H., Cheng, L., & Chuah, M. C. (2019, June). Deep-learning-based network intrusion detection for SCADA systems. In *2019 IEEE Conference on Communications and Network Security (CNS)* (pp. 1-7). IEEE.
- Yang, H., Zhang, J., Qiu, J., Zhang, S., Lai, M., & Dong, Z. Y. (2016). A practical pricing approach to smart grid demand response based on load classification. *IEEE Transactions on Smart Grid*, 9(1), 179-190.
- Yu, K., Tan, L., Mumtaz, S., Al-Rubaye, S., Al-Dulaimi, A., Bashir, A. K., & Khan, F. A. (2021). Securing critical infrastructures: deep-learning-based threat detection in IIoT. *IEEE Communications Magazine*, 59(10), 76-82.
- Zhang, P., & Peeta, S. (2011). A generalized modeling framework to analyze interdependencies among infrastructure systems. *Transportation Research Part B: Methodological*, 45(3), 553-579.
- Zio, E. (2016). Challenges in the vulnerability and risk analysis of critical infrastructures. *Reliability Engineering & System Safety*, 152, 137-150.