A Comprehensive Survey of Machine Learning: Foundations, Challenges, and Cross-Domain Applications in Cybersecurity, Healthcare, and Education

Mujtaba Ali Khan[1], Juveria Fatima[2]
[1]*College of Computing and Informatics, University of North Carolina, USA*
[2]*Department of Mathematics, College of Science, Jazan University, Jazan, KSA*
doi.org/10.64643/IJIRTV12I12-203142-459

***Abstract*—Machine Learning (ML) has become a key technology that underpins data-driven decision making in almost all fields of industry and society. But, the real-world implementation of ML models faces a number of ongoing issues with data quality, privacy, security, compliance, scale, and interpretability. In this paper, the authors provide an extensive survey of ML methods, modeling paradigms, and their cross-domain applications, focusing on three areas of great impact: cybersecurity, healthcare, and education. We use a structured review protocol that follows the PRISMA guidelines, examining the literature that has been published in peer-reviewed journals from 2015–2025, and synthesize results based on an overall taxonomy, mapping the different ML paradigms (supervised, unsupervised, semi-supervised, reinforcement, deep, online, federated, and generative learning) to representative model families and domain-specific applications. We perform comparisons with the performance of the models, explore the use of generative AI and large language models (LLMs), and address related issues such as data privacy, fairness, explainability, robust models to withstand adversarial attacks, and energy sustainability. We found that while ML has shown to have transformative abilities in threat detection, clinical diagnostics, and adaptive education, it still needs to be implemented in a responsible manner, with interdisciplinary frameworks that combine privacy-preserving computations (including federated learning, differential privacy, etc.), explainable AI (XAI) and strict data governance. Finally, we discuss open research questions and future directions towards trustworthy, scalable, and fair ML systems.**

***Index Terms*—Machine learning, Deep learning, Cybersecurity, Healthcare AI, Education technology, Federated learning, Large Language Models, Explainable AI, Adversarial robustness, Data privacy, Systematic review.**

I. INTRODUCTION

While once a niche academic research field, Machine Learning (ML) is now a ubiquitous engineering field which is responsible for the recommendations systems in e-commerce, self-driving cars, medical diagnostics, fraud detection in financial systems, network security, and personalized learning in educational systems. Three factors – the exponential growth of digital data, the development of scalable optimization algorithms and the proliferation of inexpensive parallel processors (GPUs, TPUs, and neuromorphic processors) – have made ML a general-purpose technology similar in impact to the Internet itself.

Yet, the adoption of ML in real-world contexts is still limited by a myriad of socio-technical issues. Models that do well on curated benchmarks generally fail when dealing with non-stationary data distributions, adversarial inputs, or under-represented sub-populations [2]. Blackbox deep neural networks are not easily interpretable, which makes it difficult to regulate them in critical areas like healthcare [4] [7]. Additional privacy rules like the EU's General Data Protection Regulation (GDPR) and the U.S. Health Insurance Portability and Accountability Act (HIPAA) further limit the possibility of aggregating sensitive data from a centralized repository, which is where approaches like federated learning and differential privacy come into play [8, 10].

The existing surveys usually focus on one specific sector of application, such as ML for cybersecurity [1], healthcare [4] and education [11] respectively,

without comparing algorithmic, ethical and infrastructural patterns common to different sectors of application. We fill this gap in our paper by introducing an integrated cross-domain view that bridges ML paradigms to representative model families and deployable patterns in three key sectors of societal impact.

The results of this paper can be summarized as follows:

We perform a literature review inspired by PRISMA and search for peer-reviewed papers between the years 2015–2025 in IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, PubMed and arXiv.

We envision a common taxonomy that would relate the eight ML paradigms to some of the representative model families and to domain-specific applications (Fig. 1).

We offer comparisons of state-of-the-art ML techniques, datasets and reported performance across cybersecurity, healthcare and education.

We aggregate the overarching lessons on privacy, fairness, explainability, robustness and sustainability; and explain open research challenges and future directions.

Paper organization:
A method for reviewing is described in Section II. Section III places the work into context of previous surveys. Section IV looks at deployment issues. The unified taxonomy is developed in Section V. The Programming Ecosystems are described in Section VI. Applications on the domain are in Section VII. The cross-cutting findings are discussed in section VIII. Open challenges and future directions are presented in Section IX and limitations of the review are discussed in Section X. Section XI concludes.

II. REVIEW METHODOLOGY

We used a structured review protocol that we adapted from the PRISMA 2020 guidelines [12] to ensure transparency and reproducibility. The protocol included four phases: identification, screening, eligibility and inclusion.

IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, PubMed and arXiv were used to search for articles from January 2015 to February 2025. The two Boolean queries were:

(1) a paradigm query (("machine learning" OR "deep learning" OR "neural network" OR "federated learning" OR "reinforcement learning" OR "generative AI" OR "large language model")) and
(2) a domain query (("cybersecurity" OR "intrusion detection" OR "malware" OR "healthcare" OR "medical imaging" OR "electronic health record" OR "education" OR "intelligent tutoring" OR "adaptive learning")).

Inclusion criteria. We incorporated peer-reviewed journal and conference papers, high-citation arXiv preprints (more than 50 citations) and authoritative technical reports. Papers were intended to (i) propose, evaluate or systematically review an ML model; (ii) report quantitative results on a public or clearly described private dataset; and (iii) be written in English.

Exclusion criteria. Position papers lacking empirical assessment, duplicates, papers that only talk about hardware without contributing any algorithmic development, and short opinion pieces were excluded. Synthesis. 142 primary studies were identified after de-duplication and screening. We selected the following for our analysis: (i) ML paradigm and model family; (ii) application domain and task; (iii) dataset and evaluation measure; (iv) reported performance; and (v) discussion of privacy, fairness, or explainability. Both narrative and comparative tables (in Sections IV and VII) were used to synthesize findings.

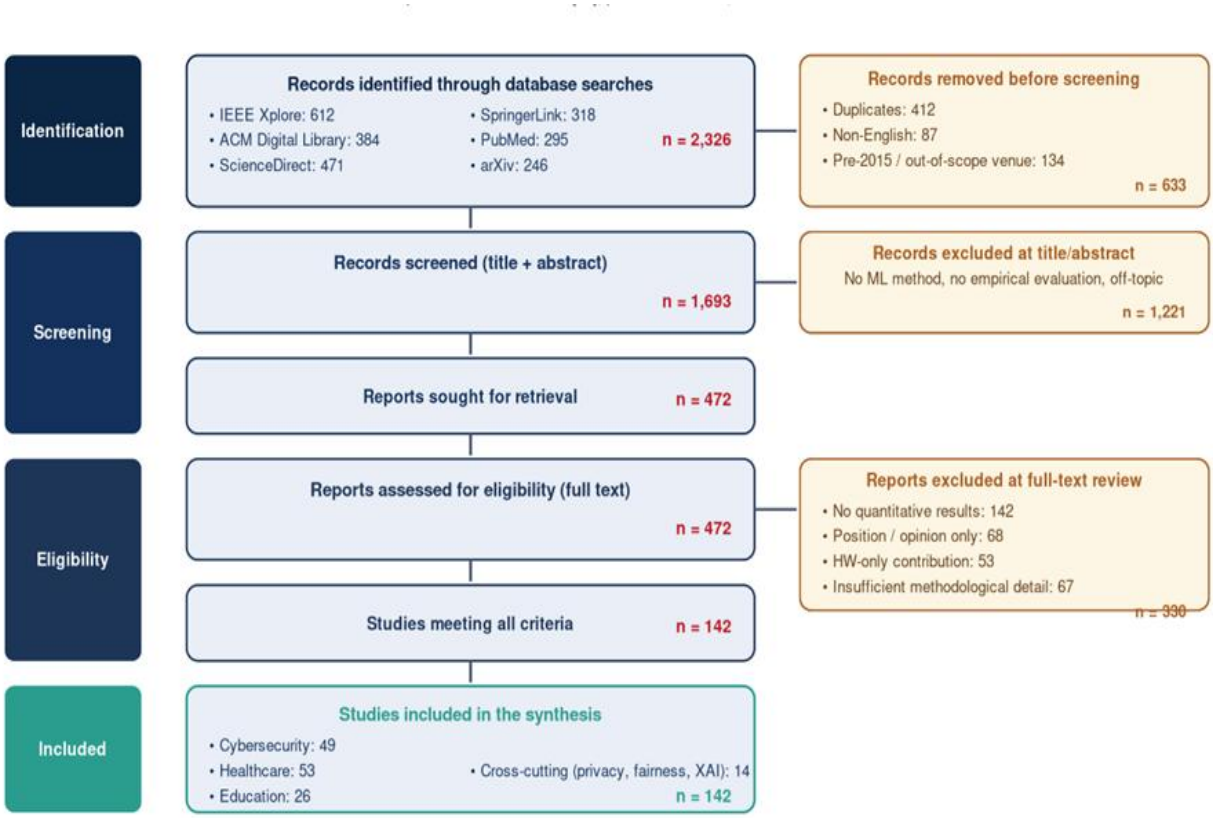

Fig. 2. PRISMA-style study selection flow. From 2,326 records identified across six databases, 142 primary studies were retained for synthesis.

III. BACKGROUND AND RELATED WORK

The literature reviews on Machine Learning (ML) can be categorized under four main categories: (i) general literature on machine learning methodology and challenges; (ii) literature on machine learning in cybersecurity and information security; (iii) literature on machine learning in healthcare and biomedical informatics; and (iv) literature on machine learning in education and learning analytics. We summarize the most influential and recent contributions in each strand, explain what they have contributed, and list the gaps that spur us on to undertake this present survey. A single table for comparison is presented in Table I.

A. General ML Methodology and Challenges

Barbierato et al. [2] critically discussed the challenges of ML adoption, highlighting issues such as data quality, reproducibility, and performance differences between benchmarks and production systems. L'Heureux et al. [3] conducted a survey on ML with big data, which included the following categories of scalability: horizontal partitioning, model parallelism, and approximate inference. Apart from this, the authors of Tufail et al. [1] provided a catalog of model architectures and libraries and Marcinkevičs and Vogt [7] gave an overview on interpretable and explainable ML with focus on methods. Lee and Shin [6] explored enterprise adoption, specifically with regard to algorithm selection. None of these take a comprehensive approach to incorporating privacy-preserving learning or generative AI.

B. ML for Cybersecurity

Buczak and Guven [21] were among the first to present a taxonomy of the ML for cyber-analytics. In the study by Alwahedi et al., [1] they have explored the ML techniques for IoT security and highlighted the emerging role of Generative AI. Wang et al. formalised the fuzzing of vulnerabilities using ML.

Apruzzese et al. [22] critically reviewed adversarial ML in network security, Ferrag et al. [23] deep learning for cyber threat detection. Sarker enumerated ML algorithms for security tasks. A recurring gap is the lack of cross domain analysis of security insights with other privacy-sensitive areas.

C. ML for Healthcare

Ghassemi et al. [4] highlighted several unique challenges of clinical ML: label noise, distribution shift, and interpretability for clinicians. A review of the achievements of deep learning in medical imaging was carried out by Esteva et al [24]. Topol [25] summarized the translation of high performance medicine in greater context. Rajkomar et al. [26] surveyed ML for eHRs and Rieke et al. [27] proposed the future of digital health in the context of federated learning. The majority of healthcare reviews are from before the advent of clinical foundation models like Med-PaLM and BioGPT in 2023–2025 [18].

D. ML for Education

Romero and Ventura [11] offered a refreshed review of the educational data mining and learning analytics fields. Hwang et al. conducted a literature review on AI in education while focusing on the research trend. Zawacki-Richter et al. [28] conducted a systematic review of AI applications in higher education and Chen, Xie and Hwang charted 20 years of AI-in-education research. Crompton and Burke [29] discussed the pedagogical implications of using AI tools. The focus of these previous reviews is not extensive enough on the coverage of LLM-based tutors and affective computing.

E. Cross-cutting Reviews

There are a few works that extend to one or more areas of privacy [8, 30] or fairness [7] or explainability [7] but mostly from a single methodological point of view, such as differential privacy alone, or SHAP alone. Bommasani et al. conducted a general survey of foundation models. To the best of our knowledge, the only previous review that connects ML paradigms with the cybersecurity, healthcare and education domains focuses on privacy, fairness, explainability, robustness and sustainability aspects, but separately.

Research gap. Within specific sectors there is substantial literature, but a lack of integration between sectors. Those who apply ML in the cross-over areas of these domains (such as a hospital information-security team, or an educational platform that processes sensitive student information) must combine the evidence to apply it. This is addressed by the present survey, which provides a cross domain perspective, based on a structured protocol that is inspired by PRISMA, and includes domain-comparative tables.

TABLE I: Comparative Summary of Representative Prior Reviews on Machine Learning

Ref.	Year	Domain Focus	ML Paradigms / Methods	Key Contribution	Limitation / Gap
[21]	2016	Cybersecurity (general)	Classical ML (SVM, DT, kNN, RF)	Foundational taxonomy of ML and data mining for cyber-analytics.	Pre-deep-learning era; no DL or generative AI.
[3]	2017	Big-data ML	Distributed and parallel learning	Classification of scalability strategies for big-data ML.	No domain-specific applications; pre-LLM.
[4]	2020	Healthcare	Mixed (DL, classical)	Identifies clinical-ML challenges: label noise, drift, interpretability.	Predates clinical foundation models.
[11]	2020	Education /	Mixed	Updated survey of educational	Limited treatment of

		EDM	(regression, clustering, DL)	data mining and learning analytics.	LLM tutors and affective AI.
[25]	2019	Healthcare	Deep learning	High-performance medicine through DL; broad translational view.	Editorial scope; limited methodological depth.
[24]	2019	Medical imaging	CNNs	Reviews DL milestones in radiology, pathology, dermatology.	Imaging-only; little on EHR or genomics.
[26]	2018	EHR / clinical ML	Deep learning on EHR	Survey of DL applied to electronic health records.	Pre-transformer; no foundation models.
[27]	2020	Healthcare (federated)	Federated learning	Articulates federated learning for privacy-preserving health AI.	Single-domain; no education or cyber link.
[22]	2022	Cybersecurity	Adversarial ML	Critical review of adversarial ML in network security.	Adversarial focus; no privacy or LLM coverage.
[23]	2020	Cybersecurity	Deep learning	Comprehensive DL-based cyber threat detection survey.	Limited treatment of evasion at scale.
[2]	2024	General ML	Methodology	Critical review of ML adoption bottlenecks.	No applied benchmarking; sector-agnostic.
[7]	2023	Interpretability (XAI)	Methods-centric XAI	Methods-focused overview of interpretable and explainable ML.	Methodology only; no domain comparison.
[28]	2019	Education (HE)	Mixed	Systematic review of AI applications in higher education.	Predates LLM era and adaptive ITS at scale.
[29]	2023	Education (ChatGPT)	LLMs	Pedagogical implications of generative-AI tools.	Single-tool, narrow scope.
[30]	2021	Fairness	Bias mitigation	Survey of fairness definitions and bias mitigation in ML.	Single-axis (fairness only).
[1]	2024	Cybersecurity (IoT)	ML, generative AI, LLMs	Surveys ML for IoT security with future GenAI vision.	Single-sector (IoT).
This work	2026	Cyber + Health + Education	Eight paradigms; LLM, FL, GenAI	Unified cross-domain taxonomy + PRISMA-style synthesis; treats privacy, fairness, XAI, robustness, sustainability as first-class.	—

IV. CHALLENGES IN MACHINE LEARNING DEPLOYMENT

We categorize the most significant problems related to ML deployment into ten thematic groups that are summarized in Table II. For all three application areas, the three canonical challenges we found in the literature are data quality, interpretability, and privacy.

Quality and quantity of data. Label noise and class imbalance are some of the key issues faced by supervised learners [5]. Some datasets are imbalanced (such as rare diseases, rare attack types), which can result in high accuracy levels but low minority recall.

Overfitting and underfitting. Models that are too complex for the amount of data that's fed into them produce models that overfit training noise. Regularisation, cross-validation, dropout and early stopping still seem to be the go-to mitigations, but principled capacity-control under distribution shift is still under active research.

Interpretability. Deep models, ensembles and large transformers are model-based with excellent accuracy but not easily explained [7]. Post-hoc explainability methods, such as SHAP, LIME, integrated gradients, attention rollout, and intrinsically interpretable models, such as rule lists, generalised additive models, are complementary tools.

Robustness and security in the adversarial setting. ML models can also be attack surfaces. Evasion attacks generate inputs that are perturbed in an imperceptible way to fool classifiers; data poisoning corrupts the training corpora; model inversion and membership inference expose training data; and model extraction recovers proprietary models via query access [8]. Partial remedies include adversarial training, certified robustness, and differentially private training.

Privacy and regulation. The era of a centralised, all-embracing training of sensitive data is becoming more and more limited by privacy regulation. The three techniques of federated learning, secure multi-party computation and homomorphic encryption allow model training without the exchange of raw data, but at the expense of communication overhead and loss of accuracy [10].

TABLE II: Principal Challenges in Machine Learning Deployment

Challenge	Description	Representative Mitigations
Data quality & quantity	Scarce, biased, or noisy labels yield non-generalisable models [5].	Active learning; data augmentation; semi-supervised learning
Data preprocessing	Cleaning, normalisation, feature engineering are time-intensive .	AutoML pipelines; feature stores
Overfitting / underfitting	Models memorise noise or fail to capture trends .	Regularisation; dropout; early stopping; cross-validation
Algorithm selection	Optimal model depends on data and problem structure [6].	Meta-learning; AutoML; Bayesian optimisation
Interpretability	Deep nets and large ensembles lack transparency [4], [7].	SHAP, LIME, attention analysis; intrinsically interpretable models
Scalability	Training and inference at web scale strain infrastructure.	Distributed training; model compression; quantisation
Ethical & legal concerns	Bias, fairness, accountability in high-stakes decisions .	Fairness-aware learning; bias auditing; algorithmic impact

		assessments
Computational resources	GPUs/TPUs and energy costs limit access .	Efficient architectures; transfer learning; green-AI metrics
Security risks	Adversarial, poisoning, inversion, extraction attacks [8].	Adversarial training; certified defences; differential privacy
Continual learning	Concept drift without catastrophic forgetting .	Replay buffers; EWC; modular / mixture-of-experts architectures

V. THE UNIFIED TAXONOMY OF MACHINE LEARNING.

Figure 1 provides a common taxonomy of the eight ML paradigms presented above and their associated model families, and the three application domains investigated in this paper. The paradigms are not mutually exclusive, e.g., deep learning can be supervised, unsupervised or reinforcement; and modern systems often use several paradigms, e.g., self-supervised pre-training combined with supervised fine-tuning.

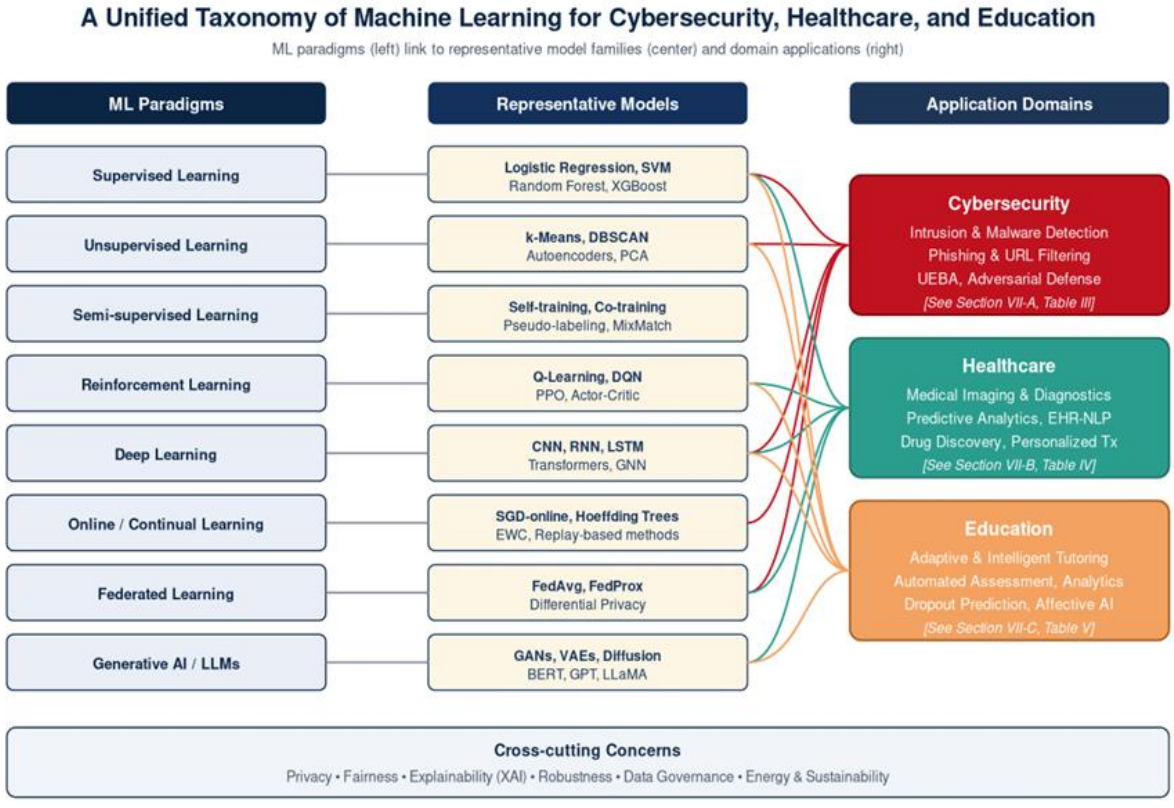

Fig. 1. A unified taxonomy of Machine Learning paradigms, representative model families, and cross-domain applications. Cross-cutting concerns (bottom bar) apply uniformly across all three sectors.

Supervised learning uses labelled examples to approximate a function f: X → Y. Linear models, decision trees, ensembles (Random Forest, XGBoost), and deep networks are the most prevalent models. Unsupervised learning is a method that identifies structure in unlabeled data, such as in clustering, dimensionality reduction, or density estimation. Semi-supervised learning (SSL) is a method that uses a labelled set and an unlabeled set of data, leveraging consistency regularisation or pseudo labelling. The

goal of reinforcement learning (RL) is to optimize a policy π using the goal function $E[\Sigma\ \gamma^t\ r_t]$ to maximize expected discounted return; current deep RL algorithms (e.g., DQN and PPO) support recommender systems and game-playing agents.Deep learning creates differentiable layers to give hierarchical representations, convolutional networks specialize in spatial data, recurrent and transformer models specialize in sequences, and graph neural networks (GNNs) specialize in relational data. Online and continual learning are important when the distribution of data evolves over time, such as in streaming or lifelong learning scenarios. Federated learning [10] is a method that allows training models at decentralised clients, while only exchanging encrypted parameter updates, but not raw data. Generative models (GANs, VAEs, diffusion models, and large language models (LLMs)) generate new data and form the basis for recent developments in conversational AI, drug discovery, and content creation [13].

VI. PROGRAMMING ECOSYSTEMS AND TOOLING

While mature libraries like PyTorch, TensorFlow, JAX, scikit-learn, and Hugging Face Transformers, and an ecosystem for distributed training (Ray, DeepSpeed) and experiment tracking (MLflow, Weights & Biases), have made Python the industry standard language for ML research and production. R has a strong position in statistical modelling and biostatistics. Most performance critical inference engines, such as TensorRT and ONNX Runtime, are built in C++. High-performance numerical computing with Python-like syntax is the focus of Julia. In browser inference is possible with TensorFlow.js and ONNX.js using JavaScript/TypeScript. These are more and more deployment dependent (cloud or edge or browser) than algorithmically preferred.

TABLE III: Programming Ecosystems for Machine Learning

Language	Primary Strength	Representative Libraries / Frameworks
Python	Research and production	PyTorch, TensorFlow, JAX, scikit-learn, Transformers
	ML	
R	Statistical modelling, biostatistics	caret, mlr3, tidymodels
C++	High-performance inference	TensorRT, ONNX Runtime, Eigen
Java	Enterprise integration	Deeplearning4j, Weka
Julia	Scientific computing	Flux.jl, MLJ.jl, SciML
MATLAB / Octave	Engineering pedagogy and prototyping	Statistics & ML Toolbox, Deep Learning Toolbox
JavaScript / TypeScript	Browser and edge inference	TensorFlow.js, ONNX.js, Brain.js

VII. CROSS-DOMAIN APPLICATIONS

A. Cybersecurity

The proliferation of security telemetry is a challenge for modern cyber-defence and is being handled with ML.It is becoming more and more common for cyber-defence to use ML to triage the number of security telemetry, and the volume and velocity of these items. Signature-based defences cannot catch polymorphic malware and zero-day attacks, so it makes sense to develop data-driven anomaly detection.

The network intrusion detection systems (NIDS) utilize supervised classifiers including Random Forest, XGBoost, CNN, and transformer encoders, to classify flow-level features from the NSL-KDD, CIC-IDS2017 and UNSW-NB15 datasets. Closed-world accuracy of over 98% is regularly achieved, but generalisation to novel attack families is an open problem [1], [14].

Malware classification is done with static analysis (opcode n-grams, byte-level CNNs) and dynamic

analysis (API call sequences, sandbox behavioural traces). Learning family-level structural signatures has been demonstrated to detect zero-day variants using deep belief networks and graph neural networks [15].

The Phishing detection uses NLP classifiers for detecting URLs, headers and HTML content in emails. The modest improvements achieved by transformer models trained with web corpora over the hand-crafted features are considerable.

User and Entity Behaviour Analytics (UEBA) uses unsupervised algorithms, such as autoencoders, isolation forests, and one-class SVMs, for detecting patterns that deviate from baselines that have been learned, which can be applied to detecting insider threats and compromised accounts.

Defensive tactics are becoming more and more important. The ML models themselves are attack surfaces and there are partial mitigations like adversarial training, certified robustness through randomised smoothing, and input transformation defences [8].

TABLE IV: Selected ML Applications in Cybersecurity

Application	ML Technique	Reported Outcome / Benchmark
Network IDS	Random Forest / CNN / Transformer	98–99.5% accuracy on CIC-IDS2017 [14]
Malware classification	Byte-CNN, Graph NN, DBN	Detects zero-day variants; F1 > 0.95 on EMBER [15]
Phishing URL detection	BERT + Logistic Regression	Real-time blocking; AUC > 0.99
Insider threat detection	Autoencoder anomaly detection	Identifies rare events; precision varies by base rate
Adversarial robustness	Adversarial training, randomised smoothing	Certified robustness against ℓ_2 perturbations [8]

B. Healthcare

From population screening to bedside decision support, ML is moving fast. They have also beaten the accuracy of dermatologists in the diagnosis of melanoma and the accuracy of radiologists in the classification of pneumonia, and now are central to FDA-cleared diagnostic devices [4], [16].

Recurrent and transformer architectures are used in predictive analytics of EHR to predict onset of sepsis, 30-day readmission and ICU mortality. Long Short-Term Memory (LSTM) and temporal-convolutional models have been used to create vital sign streams-based EWSs with hours of lead-time.

Graph neural networks have revolutionized drug discovery by revolutionizing molecular property prediction and generative diffusion models by revolutionizing de novo molecule design [17]. New protein-structure-prediction algorithms like AlphaFold have brought near-experimental accuracy, boosting the field of structural biology and rational drug design.

In personalised medicine, the reinforcement learning is used to optimise the treatment trajectory, taking into account the efficacy and toxicity of the treatment along multiple decision points, especially in oncology and chronic disease management.

Clinical NLP. Biomedical models trained on biomedical corpora (ClinicalBERT, BioGPT, Med-PaLM) capture diagnoses, drugs, and adverse events from unstructured physician notes, which alleviate documentation overload and aid in the process of cohort discovery [18].

Privacy-preserving ML. Federated learning across hospitals, along with the guarantee of differential privacy, is becoming more necessary than ever because of the sensitivity of clinical data [10].

TABLE V: Selected ML Applications in Healthcare

Clinical Task	ML Approach	Reported Outcome / Source
Pneumonia detection (chest X-ray)	Pretrained CNN (ResNet, DenseNet)	>90% accuracy; radiologist-level [16]
Sepsis prediction in ICU	LSTM / temporal CNN on vitals	Hours-ahead early warning

Clinical Task	ML Approach	Reported Outcome / Source
Drug-target interaction	Graph neural networks	Accelerated screening [17]
De novo molecule generation	Diffusion / VAE generative models	Novel scaffolds with target affinity [17]
Clinical note summarisation	ClinicalBERT, BioGPT, Med-PaLM	Reduced documentation time [18]
Multi-hospital training	Federated learning + DP	Privacy-preserving collaboration [10]

C. Education

Educational technology has gone from the delivery of a single set of content to learning technology systems that rely on data to adapt content to the needs of individual learners [11]. There are four main categories of system powered by ML: knowledge tracing, automated assessment, dropout prediction and intelligent tutoring.

Knowledge tracing models describe a learner's growth in a deep understanding of small concepts. Bayesian Knowledge Tracing (BKT) and its deep variant, Deep Knowledge Tracing (DKT) with LSTMs, estimate the likelihood of getting a right answer to the next problem and personalised content recommending.

Automated assessment relies on transformer models to reliably score short answer responses and essay answers to questions in a way similar to humans, and LLM can provide formative assessment. Demographic and dialectal bias needs to be checked in these systems.

Predicting dropouts and at-risk students. Students at risk of withdrawal are identified using gradient-boosted trees on student information from the learning-management-system, such as logins, time-on-task, and assignment patterns, allowing for timely intervention. Usually reported AUCs range from 0.80 to 0.92.

Intelligent Tutoring Systems (ITS). Pedagogical actions (hints, problem selection, dialogue moves) are optimized using reinforcement learning with learning-gain signals. In recent examples, LLM prototypes are integrated as ‘Socratic tutors,' with a focus on raising questions about the factual reliability and educational equity of LLM-based teaching tools, which require careful assessment.

TABLE VI: Selected ML Applications in Education

Application	Technique	Reported Outcome / Impact
Knowledge tracing	BKT / Deep Knowledge Tracing (LSTM)	Improved next-problem prediction
Personalised learning paths	Collaborative filtering, contextual bandits	Higher engagement, completion rates
Automated essay scoring	BERT, RoBERTa, GPT-class LLMs	Human-comparable inter-rater agreement
Early warning systems	XGBoost, Random Forest	AUC 0.80–0.92; improved retention
Conversational tutoring	Sequence-to-sequence + RLHF	24/7 student support; bias auditing required
Affective computing	CNN on facial cues; multimodal fusion	Detects engagement and frustration

VIII. DISCUSSION: CROSS-CUTTING FINDINGS

The four themes that are cross-cutting across the three sectors come up with high frequency.

Now privacy is infrastructure and not an afterthought. The most valuable applications, in each of the sectors, are those involving data that are neither legally nor ethically transferable from the institution that produced it. Federated learning, differential privacy,

secure aggregation are thus not something extra to add but part and parcel of architecture decisions.

Foundation models skew the cost curve. Large pre-trained models (vision transformer in radiology, code-pretrained transformer in malware analysis, language model in educational dialogue) provide very good baselines through fine-tuning and/or prompt-tuning on very small tasks-specific datasets. This is moving the competition from volume of raw data to curation and evaluation of high quality tasks.

Evaluation has outstripped deployment, and is now catching up. It is necessary, but not sufficient, that the accuracy of the test set be that of a closed world. Distribution-shift, sub-population, robustness, fairness and calibration analyses are now commonplace additions to benchmarks. Also, concept drift monitoring and safe-rollback functionalities are essential for operational deployment.

The energy and sustainability are important. Serving and training large models is not energy or water efficient. In addition to accuracy, metrics of Green-AI, such as FLOPs, energy per query, and parameter efficiency, are increasingly reported [19].

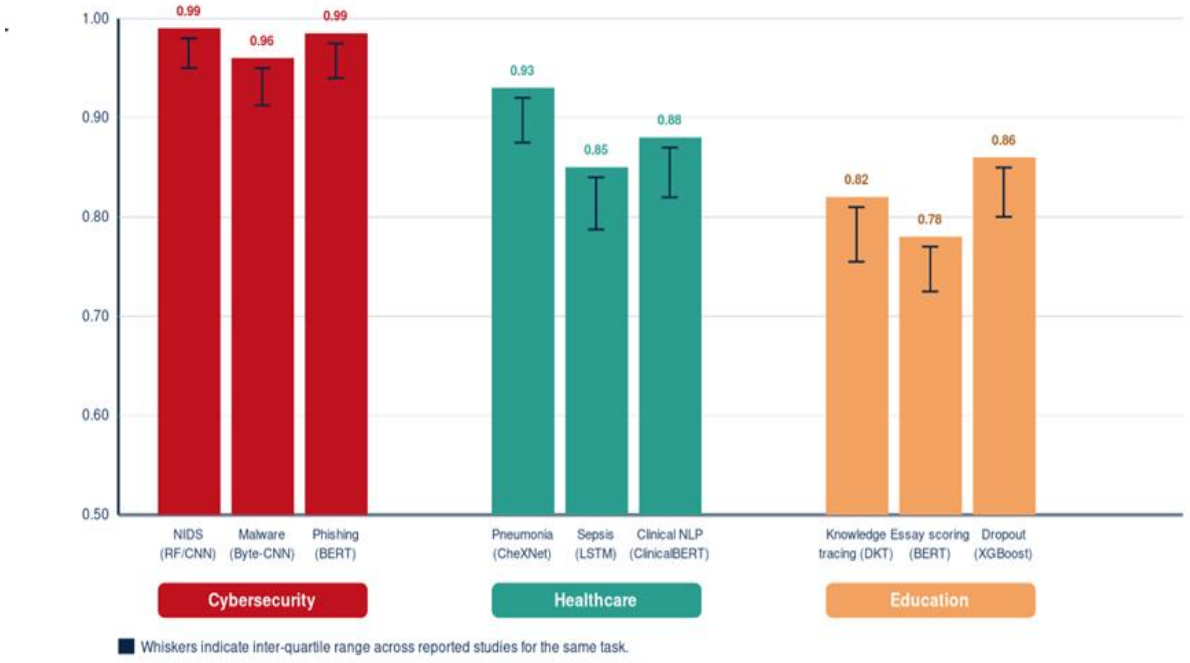

Fig. 3. Reported performance of representative ML models across cybersecurity, healthcare, and education tasks. Bars show best-reported metric (Accuracy / F1 / AUC); whiskers show inter-quartile range across studies addressing the same task in our review (n = 142).

IX. OPEN RESEARCH CHALLENGES AND FUTURE DIRECTIONS

Based on our consideration we derive a number of research directions.

At scale privacy preserving ML. It is still a challenging problem to reduce the accuracy and communication overhead for federated learning under realistic non-IID client distributions. There are systems combining the above technologies, namely Federated learning, secure aggregation, and trusted-execution environments, which are promising.

Basic models for science and clinical practice. In the short term, it is likely that the future of applied ML for a decade will be shaped by domain-specialised foundation models (clinical, security, educational), trained on well-curated corpora with strict safety evaluations.

Strong and qualified ML. Proven guarantees against adversarial perturbations, distribution shift, and data poisoning are making progress but currently only

exist for small models. There is an open challenge to scale certified defences to foundation-model regimes.
Discuss: Explainable and Trustworthy AI. Evaluation frameworks for intrinsic interpretation models and post hoc explanations are richer and need to be developed in order to relate the quality of explanation to user-relevant outcomes.
Sustainability. The realisation of large-scale ML on an environmentally scalable basis is key to continued progress in the following areas: parameter-efficient fine-tuning (LoRA, adapters); quantisation; and architecture and hardware optimisation.
Governance and regulation. A set of EU requirements, the NIST AI Risk Management Framework and the new sector-specific guidance, impose audit, transparency, and risk-management requirements on high impact ML systems. Applicability of these frameworks in engineering workflows is still a research challenge [20].

X. LIMITATIONS OF THIS REVIEW

There are three major limitations to this review. Firstly, we conducted our search in English language publications only in six databases, so publications in other languages and databases may not have been found. Second, given the speed of generative AI advancements, some of the outcomes that are reported in 2024-2025 may be outdated by the time of this report's publication. Third, because of the heterogeneity of tasks, datasets, and metrics across studies, it was not possible to do a full meta-analysis with statistical pooling of effect size using our PRISMA-inspired protocol to improve transparency.

XI. CONCLUSION

We have tried to give a cross-domain view of ML, showing its main paradigms and model families, and their application areas in cyber, health, and educ. Our analysis shows that the use of ML has been crucial to transformative impacts in these areas, but responsible application requires multi-faceted solutions to data quality, privacy, fairness, explainability, robustness and sustainability. The promise of ML is most likely to be realised in the real world in the future if we are able to combine the three key ingredients of privacy-preserving training, the use of domain-specialised foundation models, certified robustness, and good governance.

REFERENCES

[1] F. Alwahedi, A. Aldhaheri, M. A. Ferrag et al., "Machine learning techniques for IoT security: Current research and future vision with generative AI and large language models," Internet of Things and Cyber-Physical Systems, vol. 4, pp. 167-185, 2024.
[2] E. Barbierato and A. Gatti, "The challenges of machine learning: A critical review," Electronics, vol. 13, no. 2, p. 416, 2024.
[3] A. L'Heureux, K. Grolinger, H. F. Elyamany, and M. A. M. Capretz, "Machine learning with big data: Challenges and approaches," IEEE Access, vol. 5, pp. 7776-7797, 2017.
[4] M. Ghassemi, T. Naumann, P. Schulam et al., "A review of challenges and opportunities in machine learning for health," AMIA Jt. Summits Transl. Sci. Proc., vol. 2020, pp. 191-200, 2020.
[5] A. Jain, H. Patel, L. Nagalapatti et al., "Overview and importance of data quality for machine learning tasks," in Proc. 26th ACM SIGKDD (KDD), 2020, pp. 3561-3562.
[6] Lee and Y. J. Shin, "Machine learning for enterprises: Applications, algorithm selection, and challenges," Business Horizons, vol. 63, no. 2, pp. 157-170, 2020.
[7] R. Marcinkevics and J. E. Vogt, "Interpretable and explainable machine learning: A methods-centric overview with concrete examples," WIREs Data Mining and Knowledge Discovery, vol. 13, no. 3, p. e1493, 2023.
[8] L. Song and P. Mittal, "Systematic evaluation of privacy risks of machine learning models," in Proc. USENIX Security Symposium, 2021, pp. 2615-2632.
[9] J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," in Proc. ICLR, 2015.
[10] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Proc. AISTATS, 2017, pp. 1273-1282.
[11] C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated

survey," WIREs Data Mining and Knowledge Discovery, vol. 10, no. 3, p. e1355, 2020.

[12] M. J. Page, J. E. McKenzie, P. M. Bossuyt et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," BMJ, vol. 372, p. n71, 2021.

[13] T. Brown, B. Mann, N. Ryder et al., "Language models are few-shot learners," in Proc. NeurIPS, 2020.

[14] Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in Proc. ICISSP, 2018, pp. 108-116.

[15] H. S. Anderson and P. Roth, "EMBER: An open dataset for training static PE malware machine learning models," arXiv:1804.04637, 2018.

[16] P. Rajpurkar, J. Irvin, K. Zhu et al., "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," arXiv:1711.05225, 2017.

[17] Schwaller, T. Laino, T. Gaudin et al., "Molecular transformer: A model for uncertainty-calibrated chemical reaction prediction," ACS Central Science, vol. 5, no. 9, pp. 1572-1583, 2019.

[18] Singhal, T. Tu, J. Gottweis et al., "Towards expert-level medical question answering with large language models," Nature, vol. 620, pp. 172-180, 2023.

[19] R. Schwartz, J. Dodge, N. A. Smith, and O. Etzioni, "Green AI," Communications of the ACM, vol. 63, no. 12, pp. 54-63, 2020.

[20] European Parliament and Council, "Regulation on harmonised rules on artificial intelligence (AI Act)," Official Journal of the European Union, 2024.

[21] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153-1176, 2016.

[22] G. Apruzzese, M. Andreolini, L. Ferretti, M. Marchetti, and M. Colajanni, "Modeling realistic adversarial attacks against network intrusion detection systems," Digital Threats: Research and Practice, vol. 3, no. 3, pp. 1-19, 2022.

[23] M. A. Ferrag, L. Maglaras, S. Moschoyiannis, and H. Janicke, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," Journal of Information Security and Applications, vol. 50, p. 102419, 2020.

[24] A. Esteva, A. Robicquet, B. Ramsundar et al., "A guide to deep learning in healthcare," Nature Medicine, vol. 25, pp. 24-29, 2019.

[25] E. J. Topol, "High-performance medicine: The convergence of human and artificial intelligence," Nature Medicine, vol. 25, no. 1, pp. 44-56, 2019.

[26] A. Rajkomar, E. Oren, K. Chen et al., "Scalable and accurate deep learning with electronic health records," NPJ Digital Medicine, vol. 1, p. 18, 2018.

[27] Rieke, J. Hancox, W. Li et al., "The future of digital health with federated learning," NPJ Digital Medicine, vol. 3, p. 119, 2020.

[28] Zawacki-Richter, V. I. Marin, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education-where are the educators?" International Journal of Educational Technology in Higher Education, vol. 16, no. 1, p. 39, 2019.

[29] H. Crompton and D. Burke, "Artificial intelligence in higher education: The state of the field," International Journal of Educational Technology in Higher Education, vol. 20, p. 22, 2023.

[30] Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A survey on bias and fairness in machine learning," ACM Computing Surveys, vol. 54, no. 6, pp. 1-35, 2021.