DOI: 10.18372/2225-5036.31.20702

ADAPTIVE AI FOR CYBER DEFENCE: PRACTICAL EXAMPLES TO REMOVE CYBER SECURITY BLIND SPOTS

Dmytro Proskurin, Tetiana Hryniuk, Yuliia Polischuk

State University "Kyiv Aviation Institute", Ukraine

PROSKURIN Dmytro, PhD
Date and place of birth: 1994, Makariv, Kyiv Oblast, Ukraine.
Education: Taras Shevchenko National University of Kyiv, 2018.
Position: Researcher at Research Laboratory of Cyber Threats Counteraction in Aviation, State University "KAI".
Scientific interests: machine learning, information security.
Selected works: 18 scientific publications, including practical research & implementation and scientific articles.
E-mail: dmytro.proskurin@kai.edu.ua
ORCID: 0000-0002-2835-4279

HRYNIUK Tetiana
Date and place of birth: 1996, Kyseliv, Chernivtsi Oblast, Ukraine.
Education: Taras Shevchenko National University of Kyiv, 2023.
Position: Junior Researcher at Research Laboratory of Cyber Threats Counteraction in Aviation, State University "KAI".
Scientific interests: machine learning, information security.
Selected works: 2 scientific publications on AI and cybersecurity.
E-mail: tetiana.hryniuk@kai.edu.ua
ORCID: 0009-0007-8717-6752

POLISCHUK Yuliia
Date and place of birth: 1995, Netishyn, Khmelnytskyi Oblast, Ukraine.
Education: National aviation university, 2018.
Position: Junior Researcher at Research Laboratory of Cyber Threats Counteraction in Aviation, State University "KAI".
Scientific interests: information and psychological security, state information security, information security management, information security audit.
Selected works: more than 20 scientific publications, including collective monographs and scientific articles.
E-mail: polishchuk.yu.ya@gmail.com
ORCID: 0000-0002-0686-2328

Abstract. *Cyber-attacks increasingly evade static, rules-based controls by shifting content, infrastructure, and pace. This article synthesizes practical machine-learning patterns that measurably improve defence across six domains: phishing/social engineering, malware detection, network anomaly detection, insider-risk analytics, vulnerability prioritisation, and incident-response automation. The approach highlights transformer-based NLP that reads messages more like people do (with reported F1 scores of approximately 0.98 on public phishing benchmarks), image-based CNNs that recognise malware "byte-textures," autoencoders and sequence models that baseline network behaviour, federated and explainable methods for privacy-preserving insider detection, EPSS-driven triage that prioritises by exploitation likelihood, and reinforcement learning that adapts response actions under guardrails. Emphasis is on deployable patterns – shadow-mode pilots, precision/recall tracking, false-positive budgets, human-in-the-loop review, and continuous learning from user feedback and honeypot telemetry – so organisations can move from brittle signature races to adaptive systems that improve with every campaign observed. The transition to 5G and emerging 6G architectures compounds these challenges, introducing ultra-low latency requirements, massive device densities, and decentralized, edge-based infrastructures. Adaptive AI must therefore operate not only in traditional enterprise networks but also in heterogeneous, mobile, and resource-constrained 5G/6G environments where security, privacy, and resilience are paramount.*

Keywords: *Cybersecurity, Privacy-preserving AI, Phishing detection, Malware detection, CNN, Autoencoders, Ultra-low latency threat detection.*

Introduction

Defenders are outpaced when controls depend on yesterday's indicators. The sections that follow present a cohesive, evidence-informed blueprint for replacing brittle rules with adaptive learning systems: transformer NLP and explainable lightweight variants for phishing across email/SMS/app channels [1-4]; deep-learning malware classifiers that treat executables as images and model dynamic behaviour [5-7]; anomaly baselines via autoencoders/LSTMs and cross-layer features for IoT-heavy networks [8-10]; insider-risk models that respect privacy through federated learning and provide analyst-friendly explanations [11-13]; vulnerability triage that uses FIRST's EPSS to prioritise by near-term exploit probability rather than impact alone [14]; RL-assisted incident response that chooses context-aware actions within explicit guardrails [15-17]. Throughout, we stress mobile-first reality, drift monitoring, calibration to operational costs (misses vs. noise), and stepwise deployment (shadow, partial automation, constrained autonomy).

In 5G and future 6G contexts, these adaptive patterns must contend with unique network characteristic – such as dynamic slicing, edge computing nodes, and dense IoT deployments – that shift both the

attack surface and the performance envelope. Threat detection models must be latency-aware, privacy-preserving, and capable of distributed learning to defend against attacks targeting radio access networks, edge orchestration, and cross-slice data flows.

Phishing and Social Engineering Defence

Phishing still works because it fools both machines and people. On the machine side, many email defences rely on static rules and block lists. Attackers bypass these by making small, constant changes – tweaking a link, swapping a domain, rephrasing a sentence, or moving the hosting site. These fast-changing "zero-day" kits are designed to slip past filters that look for yesterday's patterns. On the human side, even well-trained employees can be caught off guard. A message that looks urgent, familiar, or business-as-usual can nudge someone into clicking before they think, especially on a phone or when they're busy.

Several practical projects have shown significant improvement on preventing the mentioned issues improving the over performance and security:

- **Transformer-based NLP for message understanding**. Fine-tuned BERT-family and RoBERTa models spot linguistic anomalies, intent, and sentiment cues indicative of phishing, often achieving F1 scores above 0.98 on public benchmarks. Comparative testing shows these transformers consistently outperform traditional classifiers on precision and recall, making them more resilient to rapidly evolving, zero-day phishing kits [1]
- **Explainable lightweight models**. DistilBERT-based approaches can match state-of-the-art accuracy while exposing attention heatmaps and token-level rationales. This interpretability helps analysts validate alerts and improve user training without sacrificing speed [3]
- **Mobile-first ensemble defences**. Hybrid "super-learner" ensembles tuned for SMS and in-app phishing reduce false negatives without inundating users with false alarms – critical on constrained mobile interfaces [4].
- **Detection enhancement**. On the PhishTank corpus, the team paired optimal n-gram vectorisation with supervised learning and cut false positives by 27% versus a baseline Random Forest, demonstrating how smarter feature engineering and model choice directly translate to fewer user-facing interruptions. The findings indicate that random forests (RF) outperformed the other classifiers, achieving a greater accuracy rate of 97.52%, followed by 97.50% precision, and an AUC value of 97%. Finally, a more robust and lightweight anti-phishing model was introduced, which can serve as an effective tool for security experts, practitioners, and policymakers to combat phishing attacks [2].

Phishing is not going away, but the cycle of playing catch-up can be broken. By pairing transformer-class language models with URL and infrastructure signals – and by keeping humans in the loop with clear, explainable alerts – can cover the two big blind spots: filters that miss novel tricks and people who get rushed or distracted. The practical path is straightforward: pilot models in shadow mode, measure precision/recall and false-positive cost, fold in "report phish" feedback for continuous learning and extend the same approach to SMS and in-app messages. In 5G/6G environments, phishing vectors expand into device-to-device messaging, augmented reality interfaces, and ultra-low-latency app communications. Defences must integrate with telecom-level threat intelligence feeds and adapt across heterogeneous access technologies without centralizing sensitive user data. Done well, this does not just lift benchmark scores; it reduces real incidents, shortens response times, and lowers noise for users. In short, NN/ML turns phishing defence from a rules race into an adaptive system that improves with every campaign it sees.

Malware Detection and Analysis

Malware evolves faster than signature rules can be updated. Packers and obfuscators reshape binaries, "fileless" techniques shift the action to memory and living-off-the-land tools, and commodity builders churn out endless minor variants. Machine learning helps by learning stable, family-level patterns instead of brittle indicators. Three complementary views matter most: bytes (what the file looks like), behaviour (what it does when it runs), and network (how it talks). In practice, that means pairing a lightweight static model for instant triage with dynamic analysis and network analytics for confirmation – so you catch more threats without drowning analysts in false alarms.

Several practical projects have shown significant improvement on preventing the mentioned issues improving the over performance and security:

- **Binaries-as-images with CNNs**. Turning executables into compact grayscale "byte-images" (e.g., 64×64) lets a CNN learn texture-like patterns that correlate with families and builder kits. On the Malimg dataset, this approach achieved high-accuracy classification and remained effective against variants designed to slip past signature scanners [5]. It's fast, hardware-friendly, and ideal as a gatekeeper before deeper analysis.
- **Deep learning beats classic feature engineering**. A contemporary review shows CNNs, RNNs, and autoencoders consistently outperform traditional methods by extracting invariant features from packed or obfuscated samples – cutting down manual feature work and improving generalization to new variants [6]. Sequence models on API calls and system

events, for example, can flag malicious routines even when the on-disk file looks benign.

- **SDN honeypot + CNN pipeline**. Combining software-defined networking honeypots (for fresh, real-world malware) with image-based CNN classification enables dynamic, near–real-time detection in large, distributed environments, including IoT. This design continuously refreshes the training set with current attacker tradecraft while keeping inference lightweight at the edge [7].

Malware detection should be treated as a multi-view learning problem rather than a signature chase. CNNs on byte-images, sequence models on behaviour, and lightweight network analytics together raise detection while holding the false-positive budget in check. The payoff is practical: faster triage on endpoints, higher recall on packed/obfuscated families, and fresher models fuelled by honeypot data – yielding fewer misses, fewer noisy quarantines, and clearer, defensible decisions at scale. Edge-deployed malware detectors in 5G/6G must balance detection accuracy with minimal inference latency, enabling on-device triage before offloading for deeper cloud analysis. Federated retraining across distributed edge nodes can ensure signatures and AI models remain fresh without centralizing proprietary or sensitive telemetry.

Network Anomaly Detection

Static IDS rules are great at catching known patterns but weak against the "quiet" stuff – slow data exfiltration, lateral movement at odd hours, and new apps that don't look like yesterday's traffic. Deep learning flips the script: first learn what **normal** looks like on your network, then flag departures from that baseline. Two families dominate: **reconstruction models** (autoencoders trained on clean traffic that alert on high reconstruction error) and **prediction models** (RNN/LSTM/transformers that forecast the next flow or feature vector and alert when reality deviates). A recent survey maps when to use which, based on data shape, label scarcity, and attack type [10].

- **Autoencoder baselines work well when labels are scarce**. Train only on "known good," then watch for spikes in error on features like bytes/sec, burstiness, SNI/JA3 rarity, or destination novelty. Adding a temporal head (e.g., LSTM) helps the model respect seasonality (workdays vs. weekends) and user/device rhythms, which cuts false positives. In IoT-heavy environments, single-layer views miss attacks that hop across the stack; cross-layer feature sets and multi-encoder designs capture interactions between link, network, transport, and application layers – the kind of nuance that betrays coordinated DDoS, MitM, or staged exfil [9].
- **Autoencoder + LSTM baseline**. An optimized AE–LSTM trained on clean traffic established a dynamic baseline and detected subtle anomalies (e.g., slow exfiltration) while reducing false positives relative to static rules [8].
- **Cross-layer IoT with M-LDAE**. A multilayer deep autoencoder tailored for IoT fused features across protocol layers to surface complex, cross-layer attacks that single-layer monitors often miss [9].
- **Method selection guide**. A survey of 180+ deep-learning AD methods organizes the field into reconstruction- vs. prediction-based approaches and offers guidance on aligning models with data and threats [10].

Network defence should be viewed as a living baseline, not a list of rules. Autoencoders give you that evolving picture of "normal," sequence models expose timing and flow irregularities, and cross-layer features bring IoT noise into focus. Together – and with routine calibration and drift checks – you cut false positives while surfacing the slow exfiltration and lateral movement that matter. The payoff is practical: fewer noisy alerts, earlier detection, and a system that adapts as the network changes instead of falling behind it. In 5G/6G networks, anomaly detection must adapt to dynamic network slicing and virtualized functions, where "normal" can vary by slice type, tenant, or service tier. Cross-layer, cross-slice baselines – potentially maintained by distributed autoencoders – can detect coordinated attacks that span physical, transport, and application domains.

Insider Threat Detection

Insider risk is tricky because the activity often looks "normal": valid logins, approved tools, and access to legitimate data. Labels are scarce, privacy limits what you can centralize, and false positives burn trust with employees and security teams. Machine learning helps by baselining how people work (rhythms, resources, and routes) and then scoring departures – while newer approaches train across organizations without moving raw logs.

- **Federated Learning + CNN**. A personalized FL system trained CNNs on encoded behaviour logs across multiple organizations – without sharing raw data – and reached ap. 92% precision in detecting insider misuse. Strong signal that cross-org learning can raise accuracy while respecting privacy [11].
- **Federated + Explainable in IoT**. Extending FL with explainable AI produced justifications for alerts in IoT settings, improving analyst trust and auditability – key where device behaviour is noisy [12].
- **Hybrid Outlier Scoring**. Blending unsupervised anomaly detection with supervised classification reduced false positives by telling apart unusual-but-legitimate work

from malicious insider actions in enterprise trials [13].

If insider threat is treated as behavioural drift under privacy constraints. Train locally (and federate globally), fuse multiple activity views, and pair anomaly scores with a small and supervised layer for precision. Wrap it with clear explanations and guardrails (data minimization, approved-use policies). The result: fewer noisy escalations, stronger privacy posture, and faster, defensible decisions when someone's access turns risky. As network functions move to the edge in 5G/6G, insider risk extends to administrators of edge clusters, orchestrators, and virtualized network functions. Federated learning across operators can flag anomalous administrative behaviours without exposing customer or tenant-specific logs.

Vulnerability Detection & Patch Prioritization

Most organizations face a flood of new CVEs each week, but only a small share ever see real-world exploitation. Prioritizing by exploitability instead of severity is the key shift. That's the idea behind FIRST.org "Exploit Prediction Scoring System (EPSS)" – a machine-learning model that uses current threat signals and observed attack telemetry to estimate the probability a CVE will be exploited in the next 30 days, with scores refreshed daily [14].

ML can enhance the existing defences by tapping into the following:

- **Exploit likelihood, not just impact**. EPSS augments CVSS by modelling real-world exploitation and producing calibrated probabilities defenders can act on, improving the efficiency of patch queues. Independent evaluations and the inaugural EPSS performance study compare efficiency/coverage trade-offs versus CVSS and KEV lists, showing EPSS enables more targeted remediation [14].
- **Context that raises or lowers risk**. Recent surveys of ML-driven prioritization show that adding contextual features – package/dependency graphs, exploit code availability, threat chatter, and asset exposure – improves triage accuracy beyond CVSS alone [14].

The scoring system also shows:

- **EPSS predictive model (FIRST.org).** Community-driven research behind EPSS (EuroS&PW 2023) details a data-driven model trained on exploitation telemetry; FIRST's docs specify EPSS as a 30-day probability updated daily for every CVE. Teams use thresholds (e.g., remediate CVEs above a chosen EPSS value) to balance coverage vs. effort [14].
- **Vulnerability prioritization survey**. An ACM Computing Surveys review synthesizes ML approaches for exploitability assessment, including graph-based risk models and market/exploit-signal features, offering guidance on selecting techniques for different environments [14].
- **Independent performance analyses**. The Cyentia Institute's inaugural EPSS study examines model performance and operationalization strategies, helping practitioners pick thresholds that maximize risk reduction with minimal patch workload [14].
- **Policy alignment**. NIST's 2025 work references EPSS as a widely used probability for near-term exploitation and discusses how to combine EPSS with KEV-style evidence for composite prioritization [14].

If patching is treated as a probabilistic triage problem and EPSS is used to rank by likelihood of exploitation, then business context (asset criticality, internet exposure, reachability) and ground truth (KEV evidence) are layered in to finalize action lists. This shifts remediation from "fix all high-CVSS" to "fix what's most likely to bite us soon," cutting noise while protecting what matters. In telecom and critical infrastructure contexts, EPSS-driven prioritization must account for vulnerabilities in virtualized network functions, RAN components, and IoT devices connected via 5G/6G links, balancing patch urgency against service continuity in latency-sensitive applications.

Incident Response Automation

Modern SOCs drown in alerts, and static playbooks can't keep up with incident variety or speed. Machine learning – especially reinforcement learning (RL) – treats response as a sequential decision problem: given the live state of an incident, choose the next best action (isolate a host, block a hash, roll back changes, escalate) to minimize impact and operator load.

ML can enhance the existing defences by tapping into the following:

- **Adaptive actions, not rigid steps**. RL learns policies that pick actions based on current signals (confidence, blast radius, asset criticality), outperforming one-size-fits-all playbooks in simulations and controlled pilots.
- **Faster correlation and triage**. ML models cluster related alerts, rank likely root causes, and route work to the right analyst or automation lane – shrinking mean time to respond.
- **Guardrails and explainability**. Policies run inside constraints (e.g., "never isolate domain controllers without human OK") and record why an action was chosen, so teams can review, roll back, and improve safely.
- **Train in safe sandboxes**. Use replay/simulation from historical incidents to learn policies offline before enabling "auto-approve" for low-risk actions.

Several practical projects have shown significant improvement on preventing the mentioned issues improving the over performance and security:

- **ARCS RL Framework**. An adaptive RL framework for SOC playbooks reported 89% effective responses and cut average handling to ~23.5 seconds in simulation, showing how policy learning can compress decision time [15].
- **Adaptive RL for Mitigation**. A live-parameter RL system dynamically selected mitigations, reducing false positives and improving recovery speed versus static playbooks [16].
- **ML-integrated IR overview**. A recent review details practical gains from ML in IR – faster alert correlation, automated root-cause hints, and predictive workload balancing for on-call teams [17].

Incident response should be treated as sense + decide + act, with RL/ML driving the "decide" step under clear guardrails. Start in shadow mode, measure time-to-contain and roll-back rates, and allow automation to handle repetitive, low-risk remediations while analysts focus on complex cases. The result is quicker containment, fewer handoffs, and playbooks that improve as they learn. In 5G/6G, RL-driven response orchestration can extend to automated isolation of compromised network slices, edge nodes, or device clusters while maintaining service for unaffected tenants. Policies must operate within strict SLA and latency bounds, ensuring resilience even during targeted DDoS or signalling storms.

Conclusion

Modernising cyber defence means institutionalising learning loops, not just adding models. Pair language-understanding detectors with URL/infrastructure signals and clear analyst explanations to cut phishing risk without user fatigue; use CNN "byte-images" for fast malware triage and enrich with dynamic/network views where stakes are higher; establish living network baselines with autoencoders/sequence heads to surface slow exfiltration and lateral movement; detect insider drift with privacy-preserving federation and rationale outputs to maintain trust; rank patches by exploitation likelihood via EPSS, tempered by asset/business context; and let RL optimise response decisions inside hard guardrails, starting in simulation and shadow mode. Operationally, measure precision/recall and false-positive costs, feedback "report phish" and honeypot captures for continuous learning, watch for model/data drift, and document governance (explanations, rollback, and safeties). Done this way, organisations shift from reactive rule-tuning to adaptive defence that reduces real incidents, shortens time-to-contain, and keeps noise within budge. The same adaptive AI patterns described here – transformer NLP, image-based CNNs, federated insider detection, EPSS triage, and RL automation – are directly applicable to 5G/6G cyber defence. Future research should focus on lightweight, latency-aware models for edge deployment, cross-operator federated learning for shared situational awareness, and AI governance frameworks that address privacy, safety, and resilience in hyper-connected mobile ecosystems.

Acknowledgement

This work was carried out within the research project "Cryptographic Data Protection System Based on Post-Quantum Encryption Algorithms and Artificial Intelligence." (No. 0125U001588), funded by the Ministry of Education and Science of Ukraine during 2025-2027.

References

[1] Meléndez R. Comparative Investigation of Traditional Machine-Learning Models and Transformer Models for Phishing Email Detection / R. Meléndez, M. Ptaszynski, F. Masui // Electronics. – 2024. – V. 13, No. 24. – Art. 4877. – doi: 10.3390/electronics13244877.

[2] Tamal M.A. Unveiling Suspicious Phishing Attacks: Enhancing Detection with an Optimal Feature Vectorisation Algorithm and Supervised Machine Learning / M.A. Tamal, M.K. Islam, T. Bhuiyan, et al. // Frontiers in Computer Science. – 2024. – V. 6. – Art. 1428013.

[3] Uddin M.A. An Explainable Transformer-Based Model for Phishing Email Detection: A Large Language Model Approach / M.A. Uddin, I.H. Sarker // SSRN Electronic Journal. – 2024. – Available at: SSRN 4785953. – doi: 10.2139/ssrn.4785953.

[4] Rao R.S. A hybrid super learner ensemble for phishing detection on mobile devices / R.S. Rao, C. Kondaiah, A.R. Pais, et al. // Scientific Reports. – 2025. – V. 15. – Art. 16839. – doi: 10.1038/s41598-025-02009-8.

[5] Musa H. Image-Based Malware Detection Using Deep CNN Models / H. Musa, M. Younis // Iraqi Journal for Computers and Informatics. – 2025. – V. 51, No. 1. – P. 64–74. – doi: 10.25195/ijci.v51i1.542.

[6] Redhu A. Deep learning-powered malware detection in cyberspace: a contemporary review / A. Redhu, P. Choudhary, K. Srinivasan, T.K. Das // Frontiers in Physics. – 2024. – V. 12. – Art. 1349463. – doi: 10.3389/fphy.2024.1349463.

[7] Kumar S. Image-based malware detection based on convolution neural network with autoencoder in Industrial Internet of Things using Software Defined Networking Honeypot / S. Kumar, A. Kumar // Engineering Applications of Artificial Intelligence. – 2024. – V. 133, Part D. – Art. 108374. – doi: 10.1016/j.engappai.2024.108374.

[8] Narmadha S. Improved network anomaly detection system using optimized autoencoder − LSTM / S. Narmadha, N.V. Balaji // Expert Systems with Applications. – 2025. – V. 273. – Art. 126854. – doi: 10.1016/j.eswa.2025.126854.

[9] Saranya K. A multilayer deep autoencoder approach for cross layer IoT attack detection using deep learning algorithms / K. Saranya, A. Valarmathi // Scientific Reports. – 2025. – V. 15. – Art. 10246. – doi: 10.1038/s41598-025-93473-9.

[10] Huang H. Deep Learning Advancements in Anomaly Detection: A Comprehensive Survey / H. Huang, P. Wang, J. Pei, J. Wang, S. Alexanian, D. Niyato // IEEE Internet of Things Journal. – 2025. – doi: 10.1109/JIOT.2025.3585884.

[11] Ye X. Research on insider threat detection based on personalized federated learning and behavior log analysis / X. Ye, F. Luo, H. Cui, et al. // Scientific Reports. – 2025. – V. 15. – Art. 19214. – doi: 10.1038/s41598-025-04029-w.

[12] Amiri-Zarandi M. A federated and explainable approach for insider threat detection in IoT / M. Amiri-Zarandi, H. Karimipour, R. Dara // Internet of Things. – 2023. – V. 24. – Art. 100965. – doi: 10.1016/j.iot.2023.100965.

[13] Yi J. Insider Threat Detection Model Enhancement Using Hybrid Algorithms between Unsupervised and Supervised Learning / J. Yi, Y. Tian // Electronics. – 2024. – V. 13, No. 5. – Art. 973. – doi: 10.3390/electronics13050973.

[14] Jacobs J. Enhancing Vulnerability Prioritization: Data-Driven Exploit Predictions with Community-Driven Insights / J. Jacobs, S. Romanosky, O. Suciu, B. Edwards, A. Sarabi // 2023 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). Delft, Netherlands. – 2023. – P. 194–206. – doi: 10.1109/EuroSPW59978.2023.00027.

[15] Ren S. ARCS: Adaptive Reinforcement Learning Framework for Automated Cybersecurity Incident Response Strategy Optimization / S. Ren, J. Jin, G. Niu, Y. Liu // Applied Sciences. – 2025. – V. 15, No. 2. – Art. 951. – doi: 10.3390/app15020951.

[16] Klein T. Optimizing Cybersecurity Incident Response via Adaptive Reinforcement Learning / T. Klein, G. Romano // Journal of Advances in Engineering and Technology. – 2025. – V. 2, No. 1. – doi: 10.62177/jaet.v2i1.212.

[17] Joseph J.E. Intelligent Incident Response Systems Using Machine Learning / J.E. Joseph, N.T. Aleke, O.P. Onyeanisi // Mikailalsys Journal of Advanced Engineering International. – 2025. – V. 2, No. 1. – P. 33–54. – doi: 10.58578/mjaei.v2i1.4540.

Проскурін Д.П., Гринюк Т.В., Поліщук Ю.Я. Адаптивний ШІ для кібербезпеки: практичні приклади усунення сліпих зон

Анотація. *Кібератаки все частіше уникають статичних, заснованих на правилах контролю, змінюючи контент, інфраструктуру та темп. У цій статті синтезовано практичні моделі машинного навчання, які помітно покращують захист у шести сферах: фішинг/соціальна інженерія, виявлення шкідливого програмного забезпечення, виявлення мережевих аномалій, аналітика внутрішніх ризиків, пріоритизація вразливостей та автоматизація реагування на інциденти.*

Ключові слова*: Кібербезпека, ШІ із збереженням конфіденційності, виявлення фішингу, виявлення шкідливого програмного забезпечення, CNN, автокодери, виявлення загроз з наднизькою затримкою.*

Проскурін Дмитро Петрович, к.т.н., науковий співробітник лабораторії протидії кіберзагрозам в авіаційній галузі, Державний університет «Київський авіаційний інститут».
Proskurin Dmytro, PhD, Researcher at Research Laboratory of Cyber Threats Counteraction in Aviation, State University «Kyiv Aviation Institute».

Гринюк Тетяна Василівна, молодший науковий співробітник лабораторії протидії кіберзагрозам в авіаційній галузі, Державний університет «Київський авіаційний інститут».
HRYNIUK Tetiana, Junior Researcher at Research Laboratory of Cyber Threats Counteraction in Aviation, State University «Kyiv Aviation Institute».

Поліщук Юлія Ярославівна, молодший науковий співробітник дослідник лабораторії протидії кіберзагрозам в авіаційній галузі, Державний університет «Київський авіаційний інститут».
Polischuk Yuliia, Junior Researcher at Research Laboratory of Cyber Threats Counteraction in Aviation, State University «Kyiv Aviation Institute».