European Journal of
Technology
(EJT)

Role of Artificial Intelligence in the Detection of Social Engineering Attacks

Oluwatosin Temitope Ogunlade

European Journal of Technology
ISSN 2520-0712 (online)
Vol.9, Issue 1, pp1 86 - 97, 2025

www.ajpojournals.org

Role of Artificial Intelligence in the Detection of Social Engineering Attacks

Oluwatosin Temitope Ogunlade
University of East London, UK

Article history
Submitted 25.08.2025 Revised Version Received 29.09.2025 Accepted 30.10.2025

Abstract

Purpose: Social engineering attacks are a major concern in cybersecurity, leveraging human psychology to access sensitive information or systems without authorization. Phishing, Chief Executive Officer (CEO) scams, and deep-fake impersonation have resulted in enormous financial and reputational loss to organizations globally. All these have proved conventional security systems to be inadequate in countering the highly developed and targeted methods used by cybercriminals. This paper therefore highlights the potentials of Artificial Intelligence (AI) to improve the detection and prevention of social engineering attacks.

Materials and Methods: Popular real-life cases were subjected to critical analysis together with AI tools such as Predictive analytics, AI powered voice, in addition to Multi-modal detection and Natural Language Processing based (NLP-based) fraud detection.

Findings: AI tools were seen to have prevented and provided complete defense against social engineering attacks. Predictive analytics permits pre-emptive detection of threats, with the potential to anticipate attacks and eliminate them before they are launched. Multi-modal detection systems, including NLP to analyze email phishing and voice forensics to detect synthesized voices by probing several communication avenues together.

Unique Contribution to Theory, Practice and Policy: This paper explores how integrating behavioral science with AI-driven detection systems can help organizations identify psychologically targeted threats, implement adaptive threat detection and strengthen security frameworks through intelligent preventive strategies.

This paper also illustrates how the integration of AI in cybersecurity systems enables organizations adopt more adaptive and proactive security postures, thereby countering social engineering threats and enhancing overall security resilience.

Keywords: *Artificial Intelligence (O33), Social Engineering (D83), Cybersecurity (L86), Machine Learning (C45), Predictive Analytics (C63)*

INTRODUCTION

The world is moving at a fast pace in this digital era of increasing cyber threats, thereby exploiting the inadequacies of humans into acts capable of revealing and sharing confidential information. One of such is social engineering attacks, reported to be amongst the most manipulative threats in contemporary cybersecurity [1]. Social engineering attacks play on the psychology of humans, tricking them into overriding security measures and often leading to the release of sensitive data or unauthorized access to systems.

This in comparison to classical exploits on technical systems, exploits cognitive biases and emotional factors such as; trust, fear, and anxiety, making them inherently hard to detect and safeguard against [1], [2]. Although, considerable research has examined the technical aspects of cybersecurity, the human dimension, particularly how attackers exploit psychological tendencies, remains a persistent weak link. Notable attacks in the form of phishing emails, CEO impersonation, and deepfakes have demonstrated the catastrophic implications of such methods on individuals and organizations [3]. However, despite technological improvements in cybersecurity infrastructure, classical detection systems continue to be insufficient, struggling to cope with the sophistication and adaptability of modern social engineering tactics.

This shortfall has driven interest in Artificial Intelligence (AI) as a transformative approach capable of learning and adapting to emerging threat patterns. Unlike conventional systems that rely on static signatures, AI can offer a proactive means of detecting, predicting, and neutralizing attacks even before execution.

Several studies have explored AI applications in cybersecurity; however, most have focused on generic threat detection or network intrusion, with limited attention to the human-centric manipulation strategies that define social engineering. While understanding human psychology is vital, there is limited empirical research connecting behavioral manipulation techniques to AI-driven detection systems. Few studies have systematically mapped AI detection mechanisms, such as; predictive analytics, NLP, and multimodal analysis, directly to real-world social engineering cases. This gap limits both theoretical understanding and practical implementation of AI for defending against psychologically driven attacks.

This paper addresses this gap by critically examining documented cases of social engineering attacks and demonstrating how AI tools could have detected or mitigated them. It also illustrates how integrating behavioral science with AI-based systems can make cybersecurity defenses more adaptive, anticipatory, and resilient to evolving attack strategies. By linking real-world incidents with AI detection mechanisms, this study contributes new insights into the development of proactive and context-aware security frameworks.

LITERATURE REVIEW

Artificial Intelligence (AI) has become a cornerstone of modern cybersecurity, offering dynamic methods to detect, prevent, and respond to social engineering attacks. Existing research emphasizes that AI systems, through machine learning, predictive analytics, and deep learning can uncover behavioral patterns, recognize anomalies, and identify deception across digital communication channels. However, studies also highlight inconsistencies in performance, ethical dilemmas, and dataset limitations that affect real-world deployment. This section reviews prior literature on AI's role in detecting social engineering attacks, organized under five key themes: behavioral and anomaly detection, predictive analytics, multimodal deepfake detection, AI-powered user training, and current research gaps and limitations.

Behavioral Pattern and Anomaly Detection

One of the most effective applications of Artificial Intelligence (AI) in cybersecurity is its ability to identify and interpret behavioral patterns through machine learning (ML) [4]. Unlike static, signature-based methods, ML systems continuously learn from data to establish behavioral baselines and detect deviations that may indicate social engineering attempts [5]. This adaptability makes AI particularly useful in countering manipulation tactics such as phishing, baiting, and pretexting, where subtle behavioral anomalies can be early indicators of an attack.

However, while several studies report strong detection performance, most models suffer from false positives and limited generalizability. AI systems often flag legitimate but infrequent user actions as suspicious, which can overwhelm security teams and reduce trust in automated alerts. In addition, many behavioral models rely on narrow or organization-specific datasets, leading to bias when applied in broader contexts. These limitations reveal that although behavioral anomaly detection offers significant promise, its effectiveness still depends heavily on dataset quality and model adaptability.

Predictive Analytics and Threat Forecasting

AI-driven predictive analytics extends traditional security measures by identifying attack patterns before they occur [7]. Through continuous analysis of network logs, communication flows, and user histories, AI can recognize potential threats in advance and help organizations prepare proactive countermeasures [8], [9]. Such predictive models are especially valuable in detecting social engineering attacks, which often follow recognizable psychological or communication patterns.

Nonetheless, researchers have reported inconsistencies in predictive performance across different contexts. While laboratory results often indicate high accuracy, real-world environments introduce noise and unpredictability that reduce precision. Furthermore, reliance on historical data can limit responsiveness to new or evolving social engineering techniques. Privacy and ethical issues also arise, as predictive monitoring requires access to sensitive user data. Hence, despite their analytical strength, predictive AI systems must be carefully balanced between accuracy, adaptability, and ethical governance.

Multi-Modal and Deepfake Detection Systems

As social engineering techniques evolve, attackers increasingly exploit multiple communication channels, including audio, video, and social media. AI-based multi-modal detection systems address this complexity by combining text, image, and voice analysis to identify inconsistencies that may indicate impersonation or deepfake manipulation [5], [8]. These systems can detect subtle anomalies, such as; mismatched facial cues or irregular vocal tones that signal synthetic or forged content.

Despite their sophistication, multi-modal models remain computationally expensive and sometimes unreliable outside controlled test settings. Many deepfake detection tools perform well on known datasets but struggle with unseen or adversarial examples, where attackers deliberately modify content to evade recognition. As a result, while multi-modal AI provides an essential layer of defense against modern social engineering threats, scalability, interpretability, and adaptability continue to be major research challenges.

AI-Powered User Training and Human Factors

The human element remains the most exploited vulnerability in cybersecurity. AI has been integrated into adaptive security awareness training, offering dynamic, personalized

simulations that help users recognize and resist manipulation tactics [10], [11]. By analyzing users' reactions to phishing simulations, AI systems can tailor content to individual learning patterns, thereby enhancing the effectiveness of training programs [12].

However, while these approaches show short-term improvements in user vigilance, their long-term impact on behavior is less conclusive. Many users revert to unsafe habits over time without consistent reinforcement. Moreover, AI-based monitoring raises ethical concerns related to privacy and autonomy, as continuous behavioral tracking can blur the boundary between legitimate education and intrusive surveillance. Hence, despite promising results, AI-powered training solutions must evolve toward more sustainable, transparent, and ethically aligned frameworks.

Summary and Research Gaps

The reviewed literature demonstrates the diverse applications of AI in detecting and mitigating social engineering attacks from behavioral anomaly detection to predictive analytics, multimodal recognition, and adaptive user training.

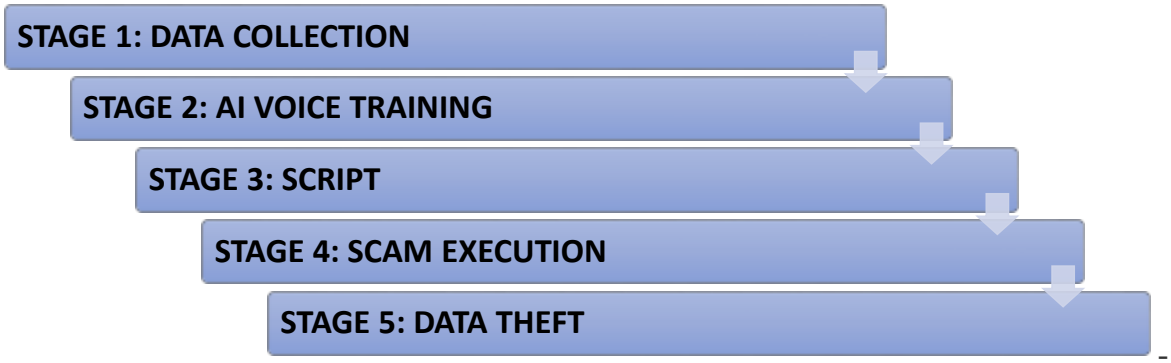
However, existing studies reveal clear research gaps. Many AI models still struggle with real-world adaptability due to limitations in dataset and contextual variability. There is also limited consensus on how to optimally integrate diverse AI methods into unified detection systems. Ethical and privacy concerns surrounding AI-driven behavioral monitoring remain underexplored.

These gaps highlight the need for continued research focused on improving model robustness, ethical compliance, and real-world applicability of AI systems for detecting social engineering attacks.

Voice Phishing

Voice phishing/vishing has been one of the highly convincing tools of social engineers, used by exploiting the trust people place in human conversation. Using deep learning models, cyber criminals can replicate the vocal tone, accent and speech patterns of specific individuals with groundbreaking accuracy. These synthetic voices can be generated from only a few seconds of audio, enabling cybercriminals to impersonate CEOs, colleagues or family members.

AI voice phishing involves 5 stages which are listed below:



Data Collection: In the first stage, cyber criminals collect voice recordings of their target from public phone calls or social media.

- AI Voice Training: The cyber criminals then go on to use AI to study recordings and mimic them, allowing them to say whatever they want sounding like the target.
- Script: Cyber criminals will prepare scripts that will sound convincing and legitimate to victims, making use of the trained AI voice. The voice is typically of someone the victim knows and trusts.
- Scam Execution: The AI voice and script is used to execute a phone call to ask the victim to share confidential or sensitive information.
- Data Theft: The victim deceived by the familiar voice passes the information to the cybercriminal thinking they are a person they know and trust.

Unlike traditional phishing methods, AI-driven voice phishing leverages psychological manipulation in real time, allowing cyber criminals to adjust their approach based on the target's responses. The result of this is a powerful blend of technical sophistication and social engineering capable of bypassing many existing authentication and verification measures.

Techniques used in Detecting Social Engineering Attacks

Artificial intelligence adopts a variety of computational techniques to identify, analyze and reduce social engineering attacks across different communication channels. Listed below are the techniques used by AI in detecting social engineering attacks:

Natural Language Processing (NLP) for Phishing Detection

NLP enables machines to understand and process human language, making it an important tool for detecting phishing attempts. By analyzing text pattern and language irregularities, NLP algorithms can flag suspicious emails, chat messages or social media interactions. Advanced models such as transformer-based architectures (e.g., BERT, GPT) can detect subtle manipulative language patterns, identify urgency or fear-inducing phrases, and differentiate between legitimate corporate communication and fraudulent messages.

Speech Recognition and Analysis for Voice Scams

AI-powered speech recognition systems can transcribe voice communications and assess them for signs of deception or impersonation. By examining tone, pitch, rhythm, and speech pauses, AI models can detect unusual vocal patterns that deviate from a known speaker's baseline. These systems can identify fraudulent calls, including those generated by AI voice cloning (vishing).

Computer Vision for Detecting Image/Video Deepfakes

The spread of deepfake technology has introduced a new threat vector in social engineering, enabling attackers to create realistic but fabricated video or image content. AI-based computer vision models can detect such manipulations by analyzing facial inconsistencies, unnatural blinking patterns and inconsistencies in lighting or shadow. Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) have been successfully applied in detecting fake media, which is critical for preventing deepfake-enabled impersonation scams.

Anomaly Detection Models

Anomaly detection in AI focuses on identifying deviations from a user's normal behavior, which could indicate a compromised account or ongoing manipulation. This includes unusual login times, atypical device usage, abnormal transaction patterns or sudden changes in communication style. Machine learning models such as; clustering algorithms and auto-encoders are trained on human's behavioral profiles and can trigger alerts when irregularities occur.

Multi-Modal AI Systems

Given the multi-channel nature of modern social engineering attacks, multi-modal AI systems integrate multiple detection layers for improved accuracy. For example, an AI platform may simultaneously analyze the linguistic content of an email (text), the metadata of an attached image (visual) and a follow-up call's audio patterns (voice) to assess risk extensively. By correlating information from different sources like; text, voice and images, these systems can spot threats more accurately, make fewer mistakes and give stronger protection against tricky multi-step attacks.

Challenges and Limitations

Artificial Intelligence (AI) has demonstrated significant potential in detecting and curbing social engineering attacks yet, several challenges continue to hinder its smooth operation. These challenges are not only technical but also ethical and operational, influencing how AI systems are designed, trained, and deployed to safeguard organizations.

Deceptive Attacks and Model Manipulation

Attackers can deliberately manipulate AI models through carefully crafted data that causes the system to wrongly classify or overlook malicious activity. This is often referred to as adversarial input manipulation, where subtle alterations in data deceive the model into producing false outcomes. For example, during data collection, AI-driven voice recognition systems can be misled by synthetic voices designed to mimic legitimate users, while NLP-based phishing detectors may be confused by linguistic alterations intentionally crafted to bypass filters.

To counter these tactics, robust AI models now incorporate adversarial training and data validation layers, enabling systems to recognize suspicious perturbations and resist manipulation attempts.

False Positives, False Negatives and Alert Fatigue

AI systems can generate both false positives (flagging safe content as dangerous) and false negatives (failing to detect actual threats). This can overwhelm cybersecurity teams and lead to "alert fatigue," where legitimate warnings are ignored.

Modern AI frameworks address this using ensemble models and reinforcement learning, in which multiple detection algorithms such as; NLP for message content, sentiment analysis for

tone, and behavioral analytics for timing, work together to cross-verify alerts. This multi-layered verification helps reduce classification errors and ensures that genuine threats receive the right level of attention.

Evolving Threat Landscape

Social engineering tactics evolve rapidly, and attackers increasingly weaponize AI to create deepfakes, voice clones, and highly personalized phishing campaigns.

AI counters this escalation through multi-modal detection systems that integrate computer vision, speech analysis, and natural language processing (NLP). For example, NLP models analyze linguistic cues in phishing emails, such as urgency or authority bias, while computer vision algorithms detect inconsistencies in facial motion or lighting within deepfake videos. This integration of multiple AI modalities allows detection systems to identify deception even when attackers exploit several communication channels simultaneously.

Data Privacy and Ethical Concerns

Training AI detection models requires access to large datasets that often contain sensitive user information, voice samples, and message logs. Without strict privacy measures, this can lead to data breaches, compliance issues, and user mistrust.

To balance effectiveness with ethics, federated learning and privacy-preserving AI methods are increasingly adopted. These approaches enable models to train across distributed data sources without allowing organizations to benefit from sensitive & collective information while maintaining data confidentiality.

High Costs and Implementation Barriers

Deploying advanced AI detection systems demands substantial computational resources, specialized expertise, and continuous maintenance. This often poses difficulties for small and medium-sized enterprises that may lack the infrastructure for frequent model retraining or multi-modal detection pipelines.

However, the emergence of cloud-based AI and AI-as-a-service solutions now offers scalable, cost-effective alternatives. These services provide access to powerful detection tools such as; NLP-based phishing filters, voice anomaly detectors, and deepfake recognition APIs, without requiring organizations to invest heavily in infrastructure.

The Need for Robust Multi-Modal Integration

Multi-modal AI extends detection capability by analyzing inputs from text, audio, and video in a unified framework. Recent developments highlight two primary levels of integration:

- Feature-level fusion, where data from multiple channels (emails, voice calls, facial cues) are combined before analysis to enhance early detection accuracy.
- Decision-level fusion, where independent subsystems (such as NLP classifiers and computer vision models) share verdicts, and a higher-level model consolidates the results for a final, more reliable decision.

This fusion architecture strengthens precision against sophisticated attacks like deepfake-enabled CEO frauds that blend fake voice and video evidence. Nonetheless, such integration increases computational complexity and data synchronization requirements, emphasizing the trade-off between detection accuracy and operational efficiency.

Case Studies and Examples

Some Case Studies and Examples Include:

Case study 1: AI and Preventing the $100 Million Phishing Scam against Google and Facebook

The phishing scam orchestrated by Evaldas Rimasauskas targeted Facebook and Google with emails impersonating a legitimate business partner, Quanta Computer. The attackers used forged invoices and fake email accounts to deceive employees into wiring over $100 million. AI could have played a critical role in preventing this scam through behavioral pattern recognition and NLP-based fraud detection.

AI systems using machine learning can analyze email metadata, identify unusual sender addresses, and detect inconsistencies in the language of communications [5]. For instance, NLP-based models trained on phishing datasets have demonstrated measurable success in reducing detection time and financial losses in large-scale enterprises [5]. AI-driven NLP could have flagged the urgent tone and deviations in the sender's email address red flags inconsistent with legitimate correspondence.

Additionally, behavioral analytics could have identified anomalies in the requests made to employees, such as large, sudden fund transfers or unusual authorization patterns. Machine learning algorithms have proven effective in identifying transaction anomalies in high-volume organizations [5], and could have flagged this deviation before transfers occurred.

Case study 2: AI in Preventing the "Fake President" Scam at FACC

The FACC attack involved criminals impersonating the CEO to request a multimillion-euro transfer. The attack succeeded because no detection system recognized the abnormal request. AI-driven predictive analytics and behavioral profiling could have prevented this by identifying communication patterns inconsistent with executive behavior [9].

AI models can analyze transaction histories and flag anomalies such as large transfers originating from new or spoofed email domains. Empirical studies confirm that predictive models leveraging historical transaction data improve fraud detection accuracy in enterprise environments [9]. Multi-modal AI systems could also have analyzed any supporting audio or video messages to verify authenticity. AI-powered voice forensics can detect deepfakes or synthetic speech inconsistencies in real time [5], making impersonation much harder to execute.

Case study 3: AI in Detecting Phony Tech Support Scams Targeting Remote Workers

During the COVID-19 pandemic, scammers posed as tech support to steal credentials and financial information. AI-driven voice analysis and multi-modal detection systems could have helped identify these vishing attacks. Machine learning models can analyze caller tone, cadence, and anomalies indicative of impersonation [5]. Recent applications of speech-recognition AI have shown strong potential in identifying social engineering cues in real-time calls [8].

Moreover, AI-driven adaptive user training can simulate phishing and malvertising scenarios, helping employees recognize and avoid manipulation [12]. Organizations that implemented AI-based awareness platforms reported measurable reductions in phishing click rates [12].

Case study 4: AI in Detecting Twitter Phishing Campaigns Targeting UK Banks

A phishing campaign targeting UK bank customers used fake Twitter support accounts to steal credentials. AI could counter such threats using social media monitoring powered by

NLP and pattern recognition. These systems can automatically scan messages for linguistic deception patterns and detect impostor accounts mimicking legitimate profiles [8].

Predictive models analyzing user interaction patterns have proven effective in detecting fraudulent behavior across social media platforms [8]. This approach would enable early detection of phishing attempts, alerting victims and institutions before sensitive data is exposed.

Across these four cases, several similarities and differences emerge, as presented in Table 1:

Table 1: Comparative Summary of Case Studies 1-4

Case Study	Attack Type	Channel	Scale	Ai Application Type
1	Phishing	Email	Large enterprises	NLP-based fraud detection, Behavioral analytics
2	CEO Fraud	Email + Voice	Corporate finance	Predictive analytics, Multimodal AI (voice + text)
3	Tech Support Scams	Voice + Web Ads	Individual workers	Voice forensics, Predictive analytics, Adaptive training
4	Twitter Phishing Campaigns	Social Media	Public social media users	NLP, Pattern recognition, Social media AI monitoring

The comparative summary highlights a consistent pattern: **AI's effectiveness against social engineering depends on contextual alignment between the attack vector and the AI modality.** NLP-based systems excel in textual deception such as phishing, whereas voice forensics and behavioral analytics prove more effective against impersonation and vishing attacks. Multi-modal AI (combining these different streams at the decision level) offers the most resilient defense, enabling cross-verification of data sources and reducing false positives. However, the cases also reveal that AI's success hinges on continuous model training and access to high-quality, domain-specific datasets. This summary underscores the need for integrated, adaptive frameworks that fuse linguistic, behavioral, and visual intelligence to address the full spectrum of social engineering threats.

CONCLUSION AND RECOMMENDATION

Conclusion

Social engineering attacks are perhaps the greatest challenge in contemporary cybersecurity. In contrast to attacks based on technical weaknesses, where trust, anxiety, and time pressure are exploited by the attacker, it is more challenging to discover and protect oneself from such attacks. The examples discussed in the form of the $100 million Google and Facebook phishing attack, the "fake president" attack on FACC, and the tech support scams on home workers illustrate the catastrophic effects of such attacks. But they also underscore the potential of artificial intelligence in revolutionizing cybersecurity defenses and limiting the damage of such attacks. AI provides several types of tools that can be used to detect and prevent social engineering attacks. Machine learning and behavioral analytics enable systems to detect abnormal behavior, such as strange transactions or access requests that might be signs of social engineering. Predictive analytics permits pre-emptive detection of threats, with the potential to anticipate attacks and eliminate them before they are launched. Multi-modal detection systems, including NLP to analyze email phishing and voice forensics to detect

synthesized voices, provide a complete defense by probing several communication avenues together. By leveraging these AI-based methods, organizations are able to transcend the weaknesses of conventional detection systems that tend to overlook the sophisticated and ever-changing methods of social engineers. With the capability to work with huge volumes of data in real time and learn from them, AI is able to discover new threats, pinpoint changing patterns of attacks, and make new adjustments.

Recommendations

For the maximization of the effectiveness of AI in detecting social engineering attacks, the following strategies are recommended:

- A Human-AI Hybrid Approach should be adopted: AI should complement and not replace, human judgment. Security teams should use AI for initial detection and prioritizing of threats, while human analysts perform a manual review for accuracy.
- Investment in Continual AI Models Training: AI detection models must be updated frequently with fresh threat intelligence to keep up with the evolving attack methods, especially AI-generated threats.
- Integration of Multi-Modal Detection: The combination of Natural Language Processing (NLP), voice biometrics, computer vision and behavioral analytics can drastically improve accuracy by cross-verifying threats across multiple channels.
- Implementation of Robust Privacy Measures: Organizations should adopt privacy-focused AI methods like federated learning and differential privacy to safeguard sensitive information during data training.
- Improve User Awareness with AI Simulations: AI can be used to generate realistic phishing simulations, deepfake scenarios, and voice-cloning examples to train employees on recognizing modern social engineering attacks.

Regulatory and Industry Collaboration: Governments, cybersecurity experts, and organizations should work together to establish guidelines, standards, and shared threat intelligence networks to fight against social engineering attacks on a global scale.

REFERENCES

[1] Wang, Z., Sun, L., & Zhu, H. (2020). Defining social engineering in cybersecurity. *IEEe Access, 8*, 85094–85115.s

[2] Dalmiere, A., Nicomette, V., Auriol, G., & Marchand, P. (2025). *A classification of manipulation technique used in social engineering attacks and underlying cognitive biases, needs, norms, and emotions*. https://hal.science/hal-05027416/

[3] Hadnagy, C. (2010). *Social engineering: The art of human hacking*. John Wiley & Sons. https://books.google.com/books?hl=en&lr=&id=9LpawpklYogC&oi=fnd&pg=PR13&dq=Hadnagy,+C.+(2018).+Social+engineering:+The+science+of+human+hacking.+Wiley.&ots=vdiBFUl4PM&sig=gbfPBishpB7nHQvpBCcjWXnvzrU

[4] Manoharan, A., & Sarker, M. (2023). Revolutionizing Cybersecurity: Unleashing the Power of Artificial Intelligence and Machine Learning for Next-Generation Threat Detection. *DOI: Https://Www. Doi. Org/10.56726/IRJMETS32644*, *1*. https://www.academia.edu/download/112737594/REVOLUTIONIZING_CYBERSECURITY.pdf

[5] Pakina, A. K., Kejriwal, D., & Pujari, T. D. (2025). Adversarial AI in Social Engineering Attacks: Large-Scale Detection and Automated Counter measures. *International Journal Science and Technology*, *4*(1), 1–11.

[6] Kolluri, V. (2024). Revolutionary research on the ai sentry: An approach to overcome social engineering attacks using machine intelligence. *International Journal of Advanced Research and Interdisciplinary Scientific Endeavours*, *1*(1), 53–60.

[7] Fakhouri, H. N., Alhadidi, B., Omar, K., Makhadmeh, S. N., Hamad, F., & Halalsheh, N. Z. (2024). Ai-driven solutions for social engineering attacks: Detection, prevention, and response. *2024 2nd International Conference on Cyber Resilience (ICCR)*, 1–8. https://ieeexplore.ieee.org/abstract/document/10533010/

[8] Schmitt, M., & Flechais, I. (2024). Digital deception: Generative artificial intelligence in social engineering and phishing. *Artificial Intelligence Review*, *57*(12), 324. https://doi.org/10.1007/s10462-024-10973-2

[9] Manyam, S. (2022). *Artificial intelligence's impact on social engineering attacks*. https://opus.govst.edu/capstones/561/

[10] Shanthi, D., Ashok, G., Biswal, C., Udharika, S., Varshini, S., & Sindhu, G. (2025). Ai-Driven Adaptive It Training: A Personalized Learning Framework For Enhanced Knowledge Retention And Engagement. *Metallurgical and Materials Engineering*, 136–145.

[11] Vadivel, S., Banupriya, R., Nivodhini, M. K., Surendhar, N. D., Subashree, N., & Murali, M. S. (2025). AI-Powered Personalization in Online Learning Systems for Enhanced Engagement and Effective Learning using Collaborative and Content-Based filtering algorithms. *International Conference on Sustainability Innovation in Computing and Engineering (ICSICE 2024)*, 1114–1139. https://www.atlantis-press.com/proceedings/icsice-24/126011423

[12] Ali, S. (2024). *The Role of AI in Social Engineering Attack Prevention: NLP-Based Solutions for Phishing and Scams*. https://www.researchgate.net/profile/Sajid-Ali-178/publication/388525951_The_Role_of_AI_in_Social_Engineering_Attack_Prevention_NLP-Based_Solutions_for_Phishing_and_Scams/links/679bcf2f52b58d39f25da252/The-Role-of-AI-in-Social-Engineering-Attack-Prevention-NLP-Based-Solutions-for-Phishing-and-Scams.pdf

License