ISSN: 2582-7219

# International Journal of Multidisciplinary Research in Science, Engineering and Technology

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*

**Impact Factor: 8.206**

**Volume 8, Issue 10, October 2025**

© 2025 IJMRSET | Volume 8, Issue 10, October 2025| DOI:10.15680/IJMRSET.2025.0810002

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

# The Role of Artificial Intelligence in Cyber Defence

**Kumudha H T[1], Sumukha S[2], Sahana M Savanth[3]**

Assistant Professor, Dept. of CSA, The Oxford College of Science, Bangalore, Karnataka, India[1]

MCA II Student, Dept. of CSA, The Oxford College of Science, Bangalore, Karnataka, India[2-3]

**ABSTRACT:** The rapid growth of cyberspace has intensified the frequency and complexity of cyberthreats, exposing the limitations of traditional security measures. Artificial Intelligence (AI) offers transformative capabilities in cyber defence by enabling real-time detection, proactive mitigation, and adaptive response. Through machine learning, deep learning, and intelligent automation, AI systems analyse vast network data, detect anomalies, and predict breaches with greater accuracy. Key applications include AI-enabled Security Information and Event Management (SIEM) systems, intrusion detection systems (IDS), malware classification, and natural language processing (NLP) for threat intelligence. AI-powered anomaly detection, behavioural analytics, and automated penetration testing further enhance resilience against zero-day and advanced persistent threats. However, challenges such as adversarial attacks, privacy concerns, and overreliance on automation highlight the need for strong integration frameworks and human oversight.

## I. INTRODUCTION

The digital revolution has transformed modern society, enabling seamless communication, online transactions, and interconnected infrastructures. However, this rapid expansion of cyberspace has also resulted in an alarming rise in the frequency, sophistication, and diversity of cyberthreats. Traditional security mechanisms, while effective against conventional attacks, are increasingly inadequate in addressing advanced persistent threats (APTs), zero-day exploits, and large-scale coordinated cyberattacks. The evolving nature of these threats demands innovative approaches that extend beyond rule-based detection and reactive defence. The adoption of AI in cybersecurity presents critical challenges, including adversarial manipulation of AI models, data privacy concerns, and the risk of excessive reliance on automated systems. Addressing these issues requires the development of robust integration frameworks that combine AI's analytical power with human expertise. Consequently, understanding the role, applications, and limitations of AI in cyber defence has become essential for designing resilient digital infrastructures capable of withstanding future cyberthreats.

## II. ARTIFICIAL INTELLIGENCE IN CYBER DEFENCE

The rise of cyberspace has brought immense benefits to communication, business, and daily life, but it has also led to an increase in the frequency and complexity of cyberattacks. Traditional security measures such as firewalls and antivirus tools are no longer sufficient to handle advanced persistent threats, zero-day attacks, and large-scale data breaches. This has created the need for more intelligent and adaptive defence mechanisms.

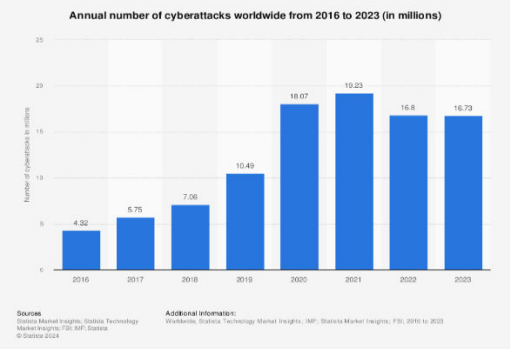
Artificial Intelligence (AI) plays a vital role in modern cyber defence by using machine learning, deep learning, and automation to analyse large amounts of data, detect unusual patterns, and predict possible breaches. AI-powered tools such as intrusion detection systems, anomaly detection platforms, and automated penetration testing help organizations respond quickly and effectively to attacks. However, challenges like data privacy, adversarial attacks, and overreliance on automation make it necessary to combine AI with human expertise for stronger and more reliable cybersecurity.

© 2025 IJMRSET | Volume 8, Issue 10, October 2025| DOI:10.15680/IJMRSET.2025.0810002

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





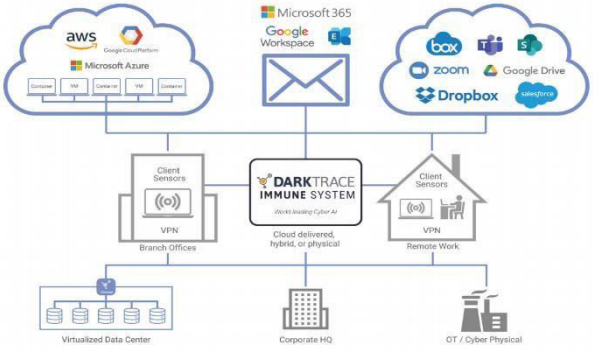
**Darktrace:**

It is an innovator in cybersecurity powered by AI. Its self-learning AI examines typical user, device, and system behaviour to identify minute irregularities instantly. In contrast to conventional defences, it recognizes new and erratic threats, like adversarial tactics and AIdriven polymorphic malware, and reacts automatically to reduce risks. AI that performs hypothesis testing, correlates data across domains, and investigates critical patterns to detect anomalies, thereby identifying novel threats and responding automatically to reduce risks.

## III. HOW DARKTRACE LEVERAGES AI FOR CYBERSECURITY

The Active AI Security Platform from Darktrace uses a multi-layered machine learning technique. AI that is integrated into an organization's infrastructure checks out interactions and data to create a constantly updated baseline of typical behaviour at the person and system levels. This makes it possible to identify unusual activity—including new and unanticipated threats— with high precision.

The platform gives security teams better visibility and control by integrating behavioural prediction, incident investigation, and real-time threat detection and response. Darktrace enables enterprises to react proactively to changing cyberthreats by decreasing dependency on conventional signature-based protection. This strategy demonstrates how AI-driven cybersecurity solutions are becoming more and more crucial for predicting, detecting, and thwarting sophisticated threats in contemporary digital environments.




**Microsoft Defender/Sentinel**

Microsoft offers a multi-layered cybersecurity strategy by combining Microsoft Defender and Microsoft Sentinel. Microsoft Defender protects endpoints, apps, and user identities from malware, phishing, and other assaults with an emphasis on proactive threat prevention. As a cloud-native Security Information and Event Management (SIEM) and Security Orchestration, Automation, and Response (SOAR) solution, Microsoft Sentinel, on the other hand, provides centralized visibility, real-time log analysis, and AI-driven threat detection in on-premises, cloud, and hybrid environments. Together, Defender provides front-line defence by seeing and stopping malicious activity at the application and device layers, while Sentinel guarantees thorough monitoring by facilitating automated incident response and sophisticated threat hunting. These technologies work together to give businesses a cohesive, smart, and expandable security against changing online threats.

© 2025 IJMRSET | Volume 8, Issue 10, October 2025| DOI:10.15680/IJMRSET.2025.0810002

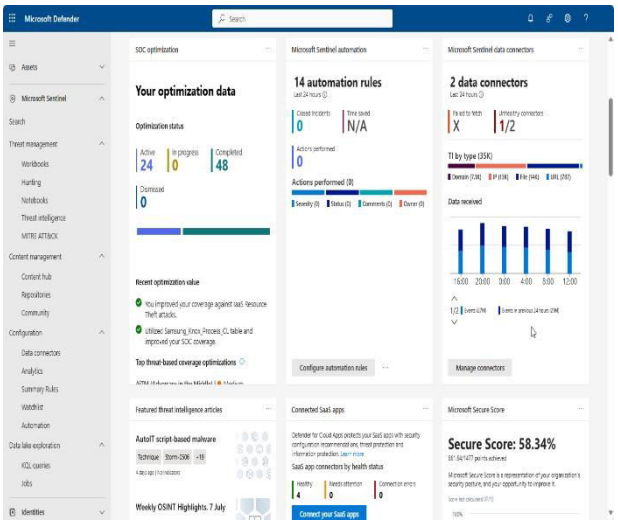
ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

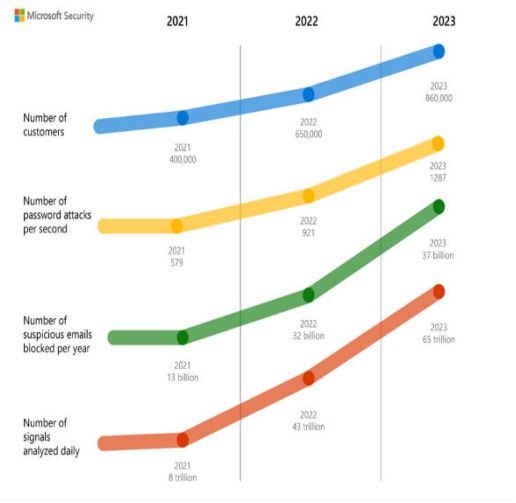
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

## IV. OPERATIONAL WORKFLOW OF MICROSOFT SENTINEL FOR THREAT DETECTION AND RESPONSE

Microsoft Sentinel's centralized visibility, sophisticated analytics, and automatic reaction capabilities make it an essential component of contemporary cyber protection. To provide enterprise-wide situational awareness, it gathers and aggregates log data from hundreds of data sources, such as the Microsoft Defender suite, other Microsoft services, and outside security solutions. Sentinel uses artificial intelligence, machine learning, and integrated correlation rules to identify abnormalities and complicated threats that conventional systems could miss. In addition to detection, it gives security analysts powerful tools for research and hunting, including Jupyter notebooks, visual analytics, and the Kusto Query Language (KQL), which allow for the proactive discovery of possible threats. Through Security Orchestration, Automation, and Response (SOAR) features, where playbooks based on Azure Logic are used, Sentinel further improves efficiency.




**Vastav AI**

India's first deepfake detection system, Vastav AI, was created by Zero Defend Security in 2025. By detecting AI-generated audio, video, and image with 99% claimed accuracy, it confirms the legitimacy of digital media. The platform assists law enforcement, media, cybersecurity companies, and individuals in protecting the integrity of digital information against threats posed by synthetic media through the use of machine learning, forensic analysis, and metadata inspection.

## V. MECHANISMS OF VASTAV AI FOR DETECTING AND CLASSIFYING SYNTHETIC DIGITAL CONTENT

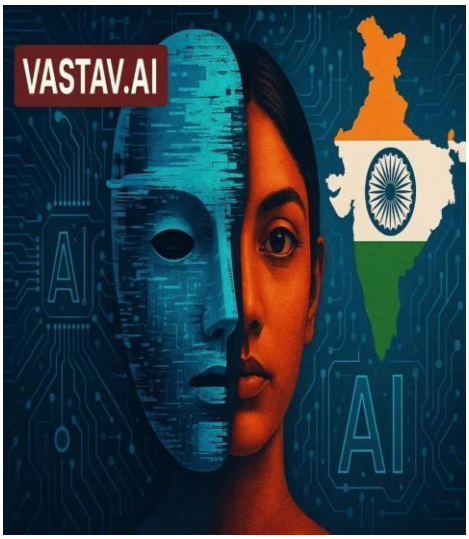
Users can upload digital materials, such as images, videos, or audio samples, to Vastav AI, which then verifies their authenticity. The system detects subtle irregularities often overlooked by the human eye, including hidden metadata, unusual sound patterns, and uneven pixels. By utilizing machine learning models trained on large datasets of both authentic and manipulated media, Vastav AI compares input files with recognized patterns of real and synthetic content. The analysis generates a confidence score, highlights suspicious regions with heatmaps, and produces a detailed report classifying the material as Real, Fake, or Suspicious. Beyond its technical capabilities, Vastav AI carries significant implications for digital trust and security. It assists journalists in ensuring news authenticity, supports law enforcement in validating digital evidence, and combats disinformation by preventing the spread of manipulated media. Furthermore, it protects individuals and organizations against fraud, false endorsements, and misleading advertisements, while promoting the responsible use of artificial intelligence to strengthen the integrity of digital ecosystems.

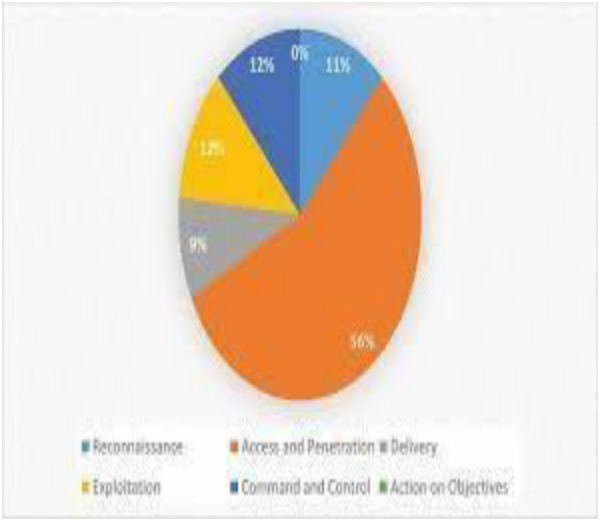
© 2025 IJMRSET | Volume 8, Issue 10, October 2025| DOI:10.15680/IJMRSET.2025.0810002

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)




## VI. CONCLUSION

Artificial Intelligence has emerged as a transformative force in cybersecurity, addressing the growing complexity and frequency of modern cyberthreats. By enabling real-time detection, predictive analytics, and automated responses, AI significantly enhances resilience against advanced persistent threats, zero-day exploits, and synthetic media manipulation. Case studies such as Darktrace, Microsoft Defender/Sentinel, and Vastav AI demonstrate how AI-driven platforms provide adaptive, scalable, and intelligent defence mechanisms. However, the challenges of adversarial attacks, data privacy, and overdependence on automation highlight the need for a balanced approach that combines AI capabilities with human expertise. Moving forward, the integration of AI into cyber defence must prioritize transparency, accountability, and robust frameworks to ensure trust and reliability. Ultimately, AI is not a replacement for human oversight but a critical enabler of secure, adaptive, and future-ready digital ecosystems.

## REFERENCES

1. [Microsoft Security] - https://www.microsoft.com/en-us/security/blog/wpcontent/uploads/2023/01/Picture1-1024x791.png
2. [Vastav.AI] - https://blogger.googleusercontent.com/img/b/R29vZ2xl/AVvXsEh1kvPWwJatQHnprqUee7Bp3vFlhy8ty9AngPyzpHn0NegvisW2ououl1v9y3UMsEAl9XQfv8a7U4eckUQmbHjnSDuRXbc-V13XArVmuaoocv_KnGHtQeS64g7PD8gatbjPuCz5vRmOAOBxZv5rvnBTLbUjBpQZMIcmOOIhkKy509954_s2ksaCh79ao4/s1200/VastavAi-deepfake.webp
3. [ResearchGate] - https://www.researchgate.net/publication/359038562/figure/fig5/AS:11431281256779688@1719571325128/Identified-AI-Driven-Cyberattack-Techniques.jpg
4. [Smart Dev] - https://smartdev.com/wpcontent/uploads/2024/09/statistic_id1485031_global-number-of-cyberattacks-2016-2023.png
5. [Microsoft Defender] - https://learn.microsoft.com/enus/azure/sentinel/media/microsoft-sentinel-defender-portal/navigation-defenderportal.png
6. [ResearchGate] - https://www.researchgate.net/profile/Meenakshi-Dwivedi3/publication/376375202/figure/fig1/AS:11431281210740019@1702142556755/Curr_ent-AI-Applications-in-Cybersecurity.jpg
7. [Cyber AI Works] - https://cyberaiworks.com/why-darktrace.asp

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

## IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com