

AI-Driven Security Techniques in Social Media

Madhavi Dhingra

Associate Professor, Department of Computer Science and Engineering, Amity School of Engineering and Technology, Amity University Madhya Pradesh, Gwalior, MP, India.

madhavi.dhingra@gmail.com

ABSTRACT

The use of social media has created a complex environment for communication and data sharing. However, these platforms are susceptible to security threats, including privacy breaches, misinformation, account hijacking, cyberbullying, and other harmful activities. Artificial Intelligence (AI) plays a crucial role in enhancing security on social media by detecting, preventing, and mitigating these threats. AI-driven security techniques use machine learning (ML), natural language processing (NLP), and deep learning algorithms to identify patterns, anomalies, and suspicious behaviors in real time. This enables the detection of fake accounts, automated bots, phishing attempts, and harmful content like hate speech and misinformation. With AI systems, inappropriate content can be identified and flagged, reducing the need for human moderation while improving response speed and accuracy. Additionally, AI models can detect changes in user interactions, signaling potential compromises or unauthorized access to accounts. Advanced AI tools also enhance user privacy by using encryption and anonymization techniques to protect personal data shared on social platforms. Despite these advancements, challenges persist in ensuring AI-driven security solutions' fairness, transparency, and accountability. Ongoing research and refinement of AI models are needed to address issues like algorithmic bias and privacy concerns. Nonetheless, AI revolutionizes social media security, providing scalable and adaptive defenses against evolving threats. This paper examines the critical role of AI-driven techniques in bolstering security on social media platforms, highlighting both the benefits and challenges of their deployment in today's digital environment.

Keywords: Artificial Intelligence, Social Media Security, Security, AI Tools for Social Media