International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

The Dark Side of AI: How Criminals Leverage Machine Learning for Illicit Activities in the Context of Assault

[1]Kadapa Chenchi Reddy, [2]Professor. Dr. Mohd. Saleem

Submitted: 15/01/2025 **Revised:** 28/02/2025 **Accepted:** 12/03/2025

Abstract: The rise of artificial intelligence (AI) and machine learning has revolutionized various sectors, but it has also opened avenues for malicious use by criminals, particularly in the context of assault. This article explores the dark side of AI, focusing on how criminals leverage these technologies to carry out both physical and cyber assaults. From weaponizing AI-driven drones for targeted attacks to using machine learning for cyberstalking, harassment, and social engineering, criminals are finding increasingly sophisticated methods to exploit these technologies.The article examines real-world examples of AI-assisted assault, including physical and cyber harassment, and the challenges law enforcement faces in detecting and prosecuting such crimes. Additionally, it discusses the ethical and legal implications of regulating AI to prevent its misuse, highlighting the need for stronger safeguards and collaboration between tech companies, policymakers, and law enforcement. As AI continues to evolve, it is essential to balance innovation with ethical responsibility, ensuring that its potential is harnessed for good while mitigating risks to individuals' safety and privacy. The article calls for increased awareness, regulation, and vigilance to safeguard society from the malicious use of AI in criminal activities.

Keywords: *Artificial Intelligence (AI), Machine Learning, Criminal Misuse, Cyber Assault, AI-driven Harassment, Cyberstalking, Deepfakes, AI in Crime, AI Weaponization. AI Surveillance, Digital HarassmentCriminal Exploitation, Data Mining, Ethical AI*

1. Introduction to Dark Side of AI

Artificial Intelligence (AI) has undoubtedly transformed industries, improving efficiency, driving innovation, and solving complex problems across healthcare, finance, and more. However, as its capabilities expand, so too does the potential for misuse. The dark side of AI refers to the ways in which malicious actors exploit these technologies for criminal purposes. While AI holds immense promise for enhancing human life, it can also be weaponized to harm, deceive, and manipulate individuals, organizations, and even entire societies.

One of the most alarming aspects of AI's dark side is its increasing use in criminal activities, particularly in the context of assault. Criminals are leveraging machine learning algorithms, deep learning models, and autonomous systems to conduct both physical and cyber assaults. From AI-driven drones that target individuals to sophisticated cyberstalking methods powered by deepfake technology, the potential for AI to facilitate harm is vast. The anonymity and scale that AI offers make it a powerful tool for evading law enforcement, creating new challenges in tracking and prosecuting offenders.

Furthermore, as AI systems become more sophisticated, they can be used to manipulate social dynamics, influence behaviors, and exploit vulnerabilities. This includes the creation of AI-driven bots that carry out harassment campaigns, or the use of predictive algorithms to identify potential victims of assault. As criminals gain access to increasingly advanced AI tools, the implications for personal safety, privacy, and security grow more severe. As artificial intelligence (AI) continues to advance, its applications have increasingly been adopted by criminals for both physical and cyber assaults, highlighting a growing concern about its potential misuse. The versatility and power of AI enable criminals to exploit it in ways that were once

1B.E.,M.T.L.,
Research Scholar- School of Law,
Presidency University, Bangalore
2Ph.D.
Associate Professor- School of Law
Presidency University, Bangalore

unimaginable, creating new and sophisticated methods of harming individuals.

This article delves into how AI is being used for illicit activities, specifically in the context of assault, exploring the growing risks and the challenges law enforcement faces in addressing these emerging threats. It also highlights the ethical, legal, and regulatory questions surrounding AI's role in criminal activities and the steps necessary to mitigate its misuse.

1.1. Focus on Physical Assaults:

AI-driven technologies, such as autonomous drones, robots, and surveillance systems, are being weaponized for physical attacks. Criminals may use AI-powered drones to track, surveil, and even target individuals. These drones can be equipped with cameras, sensors, or, in some cases, weapons, allowing for remote, precise, and often anonymous attacks. Similarly, AI-assisted robotics can be used to commit physical assaults, whether by enabling automated vehicles to run over victims or using robots to carry out targeted attacks in a highly coordinated manner. The increasing accessibility of these technologies means that criminals, including organized crime syndicates, can carry out more strategic and less detectable assaults.

1.2. Cyber Assaults:

On the cyber side, AI is being used to execute a wide range of harmful activities, such as harassment, stalking, and manipulation. AI algorithms are used to generate deepfake videos, where a person’s likeness is digitally altered to create harmful or false content. This can be used to attack someone's reputation or to intimidate them. Additionally, AI tools can automate cyberstalking, enabling criminals to track individuals' online activities, gather sensitive information, and exploit it for blackmail or psychological abuse. Social engineering attacks are also increasingly sophisticated, with AI-powered systems crafting highly personalized phishing scams, designed to deceive individuals into revealing private information or falling victim to financial exploitation.

AI's ability to analyze vast amounts of data and detect patterns makes it easier for criminals to predict victims' behaviors, vulnerabilities, and routines. These capabilities allow them to carry out highly targeted and precise assaults, whether online or in the physical world. As AI technologies continue to evolve, the threat of their exploitation by criminals for malicious purposes grows, raising concerns about safety, privacy, and the need for stronger regulation and oversight to prevent such misuse.

2. The Ethical Use vs. Criminal Use

Machine learning (ML) is a powerful tool that has revolutionized industries, from healthcare to finance, helping organizations make better decisions, improve efficiencies, and innovate in ways that were once unimaginable. At its core, machine learning is designed to analyze large datasets, detect patterns, and make predictions based on past data, all with the goal of creating positive outcomes. In ethical hands, machine learning can help doctors diagnose diseases more accurately, optimize logistics for companies, enhance customer experiences, and even combat climate change by analyzing environmental patterns. Its potential to drive positive societal change is immense, and it has been used to create technologies that save lives, improve productivity, and foster innovation.

However, as with any powerful technology, machine learning can be manipulated and misused by bad actors for harmful purposes. Criminals have found ways to exploit its capabilities, often using it in ways that are directly harmful to individuals, organizations, and society. One of the key concerns with machine learning is its ability to automate processes at scale, which can be weaponized for cybercrime, harassment, fraud, or even physical assault.

For instance, malicious actors can use ML algorithms to develop sophisticated phishing attacks, where AI tools automatically generate personalized messages to deceive individuals into disclosing sensitive information, such as login credentials or financial details. In the realm of cybercrime, machine learning can be used to automate the process of hacking, where algorithms identify vulnerabilities in systems and exploit them faster than human hackers can. Furthermore, criminals have used machine learning to create deepfakes, fabricating realistic but fake videos and audio recordings to deceive, manipulate, or defame individuals.

Another significant concern is the use of machine learning to carry out targeted assaults, such as using AI-powered surveillance to track and harass victims, or employing autonomous drones or robots in physical attacks. These technologies, while initially developed

for peaceful or productive purposes, can be easily repurposed by criminals to harm others.

The ethical use of machine learning relies on proper safeguards, regulation, and oversight. When used correctly, machine learning is a force for good. However, when abused, it becomes a tool for manipulation and harm. This dual nature of machine learning underscores the importance of establishing clear ethical guidelines, regulations, and responsible development practices to ensure that its potential for harm is minimized and that it continues to be used for positive societal benefit.

Advancements in artificial intelligence (AI) have led to the development of technologies such as drones, robotics, and autonomous vehicles, which, while offering numerous benefits, are increasingly being exploited by criminals for physical assaults. These AI-driven systems provide unprecedented capabilities that malicious actors can manipulate to execute sophisticated and targeted attacks.

3.1.AI-Driven Drones in Criminal Activities

Criminal organizations have begun integrating drones into their operations, leveraging AI for surveillance, trafficking illicit goods, and potentially conducting physical assaults. For instance, Mexican drug cartels have adopted unmanned aerial vehicles (UAVs) to monitor rival territories and transport narcotics, enhancing their operational efficiency and evading law enforcement detection. Moreover, researchers have demonstrated that AI-powered drones can be manipulated into performing harmful actions. By exploiting vulnerabilities in large language models, criminals could potentially instruct drones to engage in unauthorized surveillance or deliver contraband to restricted areas, thereby facilitating criminal activities.

3.2.Weaponize of AI-Powered Robot

The weaponize of AI-driven robots presents significant security concerns. Studies haveZshown that robots powered by large language models can be easily manipulated into performing dangerous and unethical actions. For example, researchers at the University of Pennsylvania demonstrated that a simulated self-driving car could be tricked into driving off a bridge, highlighting the potential for AI systems to be coerced into harmful behaviors.

Additionally, the development of robot dogs equipped with weaponry has raised alarms. These robots, designed for military and law enforcement applications, can be armed with rifles, rocket launchers, and flamethrowers. Their deployment in conflict zones, such as Ukraine, underscores their potential for both tactical support and the execution of targeted assaults, raising ethical and security concerns.

3.3.Exploitation of Autonomous Vehicles

The rise of autonomous vehicles introduces new avenues for criminal exploitation. Criminals could potentially hack into the AI systems of self-driving cars to hijack them for smuggling operations or to facilitate physical assaults by manipulating the vehicle's navigation and control systems. While specific incidents may be limited, the theoretical risks are significant, prompting discussions about the need for robust cybersecurity measures in autonomous vehicle technologies.

3.4.Emerging Threats and Security Challenges

The integration of AI into physical systems presents emerging threats that require urgent attention. Experts warn about the risks of autonomous AI systems acting unpredictably due to their reliance on large datasets and closed-loop decision-making processes. For instance, there is concern that AI systems in healthcare or judiciary could prioritize operational goals over ethical considerations, potentially causing harm or perpetuating biases.

Addressing these challenges necessitates a comprehensive approach, including robust cybersecurity measures, ethical guidelines for AI development, and international collaboration to prevent the misuse of AI technologies. As AI continues to evolve, staying ahead of potential threats is crucial to ensure these technologies serve the public good without compromising safety and security.

4.Exploiting AI in Public Spaces

As AI technologies continue to evolve, they are increasingly integrated into public systems, creating significant benefits such as enhanced security, efficiency, and convenience. However, this also opens the door for criminals to exploit vulnerabilities in AI-enabled systems to carry out targeted assaults, both physically and digitally. These systems, designed to optimize services, security, and functionality, can be manipulated by bad actors who understand how to exploit weaknesses in the underlying technology.

AI is increasingly being deployed in public spaces to improve security, surveillance, and overall urban management. For example, AI-driven facial

recognition systems are used in airports, shopping malls, and other public areas to monitor individuals for security purposes. While these technologies are designed to identify potential threats, they also present a significant vulnerability when criminals can manipulate them.

4.1.Facial Recognition Manipulation:

Criminals could exploit flaws in facial recognition algorithms to evade detection or target individuals. For instance, using high-resolution images or 3D modeling, bad actors can bypass these systems by impersonating specific individuals or falsely identifying targets. This creates opportunities for identity theft, stalking, or even orchestrating physical assaults on unsuspecting victims. Additionally, facial recognition systems in public spaces may be vulnerable to "spoofing," where criminals use digital tools to trick the system into misidentifying them or recognizing fake faces.

4.3.AI-Driven Surveillance Systems:

AI-powered surveillance cameras in cities or stores are designed to monitor activities and alert authorities to suspicious behavior. However, criminals can exploit these systems by disrupting their operation or by using deepfake technology to mask their identities. Malicious actors could exploit these AI systems to surveil their targets over extended periods, identify vulnerabilities, and then carry out assaults, either physically or digitally.

4,4.Autonomous Attacks on Critical Infrastructure:

Criminals can also exploit AI systems to launch large-scale attacks on critical infrastructure. For example, by manipulating AI-enabled systems that control power grids, water supplies, or transportation networks, attackers can disrupt services, create chaos, and indirectly cause harm to individuals. This type of exploitation can be used as a smokescreen for more direct physical assaults or be an end in itself, destabilizing environments to facilitate more targeted violence.

4.5.AI in Autonomous Weapons and Drones

Criminals can also exploit AI-enabled autonomous weapons, such as drones, to launch physical assaults in both public and private settings. With advancements in AI technology, autonomous drones can be used to carry out surveillance, drop payloads, or even deliver explosives. By hacking into the control systems of these drones, criminals can gain access to sophisticated weaponry and use it for targeted violence or intimidation. This type of weaponization poses a significant threat to both personal safety and public security, as criminals can carry out attacks remotely, making it difficult for law enforcement to prevent or respond in real time.

5.AI in Private Settings and Home Security Systems

In private settings, AI-enabled systems, including smart home security, virtual assistants, and surveillance cameras, are often connected to the internet and control vital aspects of people's lives. While they offer convenience and safety, they also provide criminals with a way to exploit vulnerabilities for malicious purposes.

5.1. Hacking AI-Driven Security Systems:

Home security systems powered by AI, such as smart cameras or smart locks, are susceptible to hacking. Criminals could exploit weaknesses in the network or software to disable these systems, allowing them to break into homes undetected. In some cases, AI systems might be tricked into misidentifying the criminal as a legitimate occupant or owner, facilitating the intrusion. Additionally, vulnerabilities in smart devices can be used for surveillance, allowing criminals to track a person's movements within their home or property, making it easier for them to plan and execute a targeted assault.

5.2.Virtual Assistants and Social Engineering:

AI-powered virtual assistants, like Amazon’s Alexa or Google Assistant, can be exploited by criminals to gather personal information about victims. Criminals could hack into these systems to eavesdrop on conversations, learn about a person’s schedule, and use this information to stage attacks. By leveraging social engineering techniques, criminals could manipulate AI-powered assistants to trick users into providing confidential details, thereby facilitating more targeted and personalized assaults, whether through financial theft, blackmail, or even physical harm.

6. Real-World Examples AI’s Exploitation

As AI technologies become more integrated into our daily lives, their potential for both positive and

negative outcomes grows. While AI has made significant contributions to sectors such as healthcare, education, and finance, it has also been exploited by criminals for malicious purposes. This section explores real-world examples and case studies that highlight how AI is being weaponized for physical and cyber assaults.

6,1. AI-Powered Drones Used for Smuggling and Surveillance

One of the most concerning examples of AI-driven technology used for criminal activities is the use of autonomous drones in drug trafficking. Criminal organizations, particularly drug cartels, have increasingly turned to AI-powered drones to evade law enforcement and smuggle illicit goods across borders.

In Mexico, drug cartels have employed drones equipped with AI to fly under the radar of authorities, transporting drugs over borders, into cities, and to remote areas. These drones are capable of autonomously navigating complex terrains, avoiding detection from traditional surveillance systems. They also use AI to analyze their surroundings in real-time, allowing them to detect obstacles and adjust flight paths accordingly. In some cases, these drones have been used to transport drugs into prisons, delivering contraband directly to inmates, significantly increasing the efficiency of smuggling operations.

The criminal groups can exploit AI for real-time decision-making, enabling them to perform coordinated operations that would otherwise be impossible to execute with manual control. The integration of AI into these operations makes it more challenging for law enforcement to predict or intercept the drones before they complete their illicit tasks.

6.2. Weaponized AI Robots in Military and Civilian Applications

AI-powered robots are another tool that has been weaponized for both military and criminal use. While these robots were initially developed for military operations, their potential for misuse by non-state actors or criminal organizations has raised significant ethical and security concerns.

One of the most widely known examples of AI robots being used for military and security purposes is Boston Dynamics' "robot dogs," also known as "Spot." These robots are equipped with advanced AI, which allows them to navigate complex environments and perform tasks such as surveillance, bomb detection, and delivery. While originally intended for peaceful applications in industries like logistics and construction, these robots have been adapted for military and law enforcement roles.

However, concerns have been raised about the weaponization of these robots. In conflict zones, such as Ukraine, there have been reports of robots being deployed with weaponry, including rifles and flamethrowers. While these robots are typically deployed for tactical support, criminals could exploit these same systems to carry out attacks, surveillance, or sabotage. The ability of these robots to autonomously move and make decisions could be exploited by malicious actors to target individuals or infrastructure with high precision, making them highly effective tools for violence.

6.3. Autonomous Vehicles Hacked for Smuggling and Violent Acts

Autonomous vehicles (AVs) are increasingly being deployed on roads and in logistics systems, promising to revolutionize transportation. However, AVs, like any AI-powered system, are vulnerable to hacking and exploitation, which could lead to violent acts and targeted assaults.

In a notable example, criminals have been exploring the potential of hijacking autonomous vehicles to smuggle drugs, weapons, and even people. Autonomous trucks and delivery vans equipped with AI navigation systems are vulnerable to remote hacking, where criminals could gain control of the vehicle and reroute it for illicit purposes. This could be used for drug smuggling, human trafficking, or transporting illegal goods across borders without detection by authorities.

While there have been no widely publicized cases of AVs being directly used for physical assaults, the possibility remains that malicious actors could take control of a self-driving vehicle and use it as a weapon. AI systems that control the vehicles' navigation, braking, and acceleration could be manipulated to cause accidents or target specific individuals, creating significant risks to public safety.

6.4. Deepfake Technology Used for Defamation and Harassment

AI-driven deepfake technology has emerged as one of the most concerning applications of machine learning, particularly in the realm of cyber-assaults. Deepfakes allow individuals to manipulate videos, audio, and images to create realistic, yet entirely fabricated

content. Criminals can use deepfakes to harass, defame, and exploit their victims in harmful ways.

One real-world case involved a prominent individual who became the target of a deepfake blackmail scheme. Criminals used AI to create a fabricated video of the individual in a compromising position, then threatened to release the video unless a large sum of money was paid. The deepfake was so realistic that it was difficult to distinguish from real footage, making the victim's situation even more harrowing.

In another case, deepfakes were used to create false images of a woman, portraying her in highly inappropriate scenarios to damage her reputation. The perpetrators used AI to alter her likeness, resulting in widespread harassment and emotional distress. Deepfakes like these have been used to ruin careers, destabilize relationships, and even ruin lives, as the victims struggle to clear their names and deal with the psychological trauma inflicted by these malicious AI tools.

The misuse of AI technologies for criminal activities, ranging from cyber assaults to physical violence, represents a significant and growing threat. As AI continues to advance and integrate into various sectors, it is essential to develop and enforce robust security measures, legal frameworks, and ethical guidelines to prevent its exploitation for malicious purposes. The real-world examples and case studies discussed here demonstrate the range of criminal activities that can be facilitated by AI, underscoring the urgent need for vigilance and proactive measures to mitigate these risks.

AI technologies such as drones, autonomous vehicles, robotics, and AI-powered surveillance systems provide immense potential for innovation and societal progress. However, their vulnerability to exploitation by bad actors has raised serious concerns about the safety and security of both individuals and societies. From weaponized drones used in smuggling and trafficking to AI-driven deepfakes designed for defamation and blackmail, the misuse of AI poses significant challenges to public and private safety

The rapid development and integration of artificial intelligence (AI) into various sectors has undoubtedly brought about numerous advancements and benefits, from improving efficiency in industries to enhancing security and healthcare systems. However, as explored throughout this article, the dark side of AI is becoming increasingly evident, with criminals exploiting its capabilities for malicious purposes, including physical and cyber assaults.

Acknowledgments

I thank our Research Guide and Supervisor Associate Prof. Dr.Mohd. Saleem from School of Law, Presidency University, Bangaore who is coauthor and provided insight and expertise that greatly assisted the research, although they may not agree with all of the interpretations/conclusions of this paper. I thank Prof. Saroj Sharma, Dean of School of Law for assistance in this article, that greatly improved the manuscript.

AuthorS contributions

Kadapa Chenchi Reddy.B,E.,M.T,L.,

Associate Professor Dr. Mohd. Saleem. Ph.D.,

School of Law, Presidency University, Bangalore.

Conceptualization, Methodology, AI Laws, Indian contexts, Writing-Original draft preparation, Software, Validation. And Field study

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Buiten, Miriam, Alexandre d. Streel, and Martin Peitz. 2023. "The law and economics of AI liability." *Computer Law & Security Review* 48 (April): 1-20. https://doi.org/10.1016/j.clsr.2023.105794

[2] De Conca, S. 2022. "Bridging the liability gaps: Why AI challenges the existing rules on liability and how to design human-empowering solutions." https://doi.org/10.1007/978-94-6265-523-2_13

[3] Čerka, Paulius, Jurgita Grigienė, and Gintarė Sirbikytė. 2015. "Liability for damages caused by artificial intelligence." *Computer Law & Security Review* 31, no. 3 (June): 376-389. https://www.sciencedirect.com/science/article/abs/pii/S026736491500062X

[4] Park, Sangchul. 2024. "Bridging the Global Divide in AI Regulation: A Proposal for a Contextual, Coherent, and Commensurable Framework." *Washington International Law Journal* 33, no. 2 (August): 1-58. https://arxiv.org/pdf/2303.11196

[5] Jaiman, Ashish. 2020. "Debating the ethics of deepfakes." ORF. https://www.orfonline.org/expert-speak/debating-the-

ethics-of-deepfakes

[6] MacDonald, Abby. 2024. “The Uses and Abuses of Deepfake Technology.” *Canadian Global Affairs Institute*. https://www.cgai.ca/the_uses_and_abuses_of_deepfake_technology#Good

[7] Reuters. 2024. “‘Cheapfakes’, not deepfakes, spread election lies in India.” *The Hindu*, May 31, 2024. https://www.thehindu.com/sci-tech/technology/cheapfakes-not-deepfakes-spread-election-lies-in-india/article68235040.ece

[8] Bond, Shannon. 2023. “People are arguing in court that real images are deepfakes.” NPR, May 8, 2023. https://www.npr.org/2023/05/08/1174132413/people-are-trying-to-claim-real-videos-are-deepfakes-the-courts-are-not-amused

[9] Sadaf, Fahim, & G. S. Bajpai. 2020. “AI and Criminal Liability.” *Indian Journal of Artificial Intelligence and Law*, 1(1). https://www.academia.edu/86155216/AI_and_Criminal_Liability

[10] Hallevy, Gabriel. 2010. “The Criminal Liability of Artificial Intelligence Entities by Prof. Gabriel Hallevy :: SSRN.” https://ssrn.com/abstract=1564096

[11] MacDonald, Abby. 2024. “The Uses and Abuses of Deepfake Technology.” *Canadian Global Affairs Institute*. https://www.cgai.ca/the_uses_and_abuses_of_deepfake_technology#Good

[12] Sala, Alessandra. 2024. “AI watermarking: A watershed for multimedia authenticity.” *ITU*. https://www.itu.int/hub/2024/05/ai-watermarking-a-watershed-for-multimedia-authenticity/

[13] Stricklin, Kasey. 2021. “Social Media Bots and Section 230 Reform with Unintended Consequences.” CNA. https://www.cna.org/our-media/indepth/2021/04/social-media-bots-and-section-230

[14] Khlaaf, Heidy. 2023. “Toward Comprehensive Risk Assessments and Assurance of AI-Based Systems.” *Trail of Bits*. https://www.trailofbits.com/documents/Toward_comprehensive_risk_assessments.pdf

[15] Widder, David G., Dawn Nafus, Laura Dabbish, and James Herbsleb. 2022. “Limits and Possibilities for “Ethical AI” in Open Source: A Study of Deepfakes.” *FAccT '22: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. https://doi.org/10.1145/3531146.3533779

[16] Xiang, Chloe, Janus Rose, Magdalene Taylor, Jordan Pearson, Matthew Gault, Samantha Cole, and Ryan S. Gladwin. 2023. “‘He Would Still Be Here’: Man Dies by Suicide After Talking with AI Chatbot, Widow Says.” *VICE*. https://www.vice.com/en/article/man-dies-by-suicide-after-talking-with-ai-chatbot-widow-says/

[17] Indurkhya, Bipin. 2023. “Ethical Aspects of Faking Emotions in Chatbots and Social Robots.” *32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 1719. https://arxiv.org/pdf/2310.12775

[18] Natale, Simone. 2021. “Chapter 3 The ELIZA Effect: Joseph Weizenbaum and the Emergence of Chatbots.” In *Deceitful Media: Artificial Intelligence and Social Life After the Turing Test*, 50-67. Oxford University Press. https://doi.org/10.1093/oso/9780190080365.003.0004

[19] Ryan, William A., Allen Garrett, Kilpatrick Townsend, and Brad Sears. 2023. “Practical Lessons from the Attorney AI Missteps in Mata v. Avianca.” *Association of Corporate Counsel*. https://www.acc.com/resource-library/practical-lessons-attorney-ai-missteps-mata-v-avianca