An Overview of Criminal Exploitation and Abuse of Artificial Intelligence

[1] SRI LAKSHMI Y , [2] SHRUTHI M T

[1]Student,Department of Master of Computer Applications,BIET,Davangere
[2]Assistant professor ,Department of MCA,BIET,Davangere

ABSTRACT

The rapid advancement of Artificial Intelligence (AI) has led to its integration across various sectors of society. However, AI has also become increasingly embedded in criminal and harmful activities, thereby amplifying existing vulnerabilities and introducing new threats. This article examines pertinent literature, reports, and key incidents to develop a typology of the malicious use and abuse of AI-enabled systems. The primary goal is to categorize these activities and assess the associated risks. We begin by identifying the vulnerabilities inherent in AI models and discussing how malicious actors exploit these weaknesses. Following this, we explore AI-enabled and AI-enhanced attacks. While our review is comprehensive, it is not meant to be a definitive or exhaustive classification. Instead, we aim to provide an overview of the risks associated with advanced AI applications, thereby contributing to the ongoing discourse on this topic. Specifically, we identify four categories of AI abuse (integrity attacks, unintended AI outcomes, algorithmic trading, membership inference attacks) and four categories of AI misuse (socialengineering, misinformation/fake news, hacking, autonomous weapon systems). By mapping these threats, we facilitate a deeper consideration of governance strategies, policies, and actions that can be developed or refined to mitigate risks and prevent harmful outcomes. Strengthening collaboration among governments, industries, and civil society is essential to enhance preparedness and resilience against the malicious use and abuse of AI.

Keywords: AI, Crime, Malicious, Attack, Hacking

1. INTRODUCTION

The impact of systems utilizing Artificial Intelligence (AI) is a focal point in many academic studies, political discussions, and reports from civil society organizations. The development of AI has garnered significant praise for its unparalleled technological capabilities, such as the enhanced ability for automated image recognition (e.g., detecting cancer in medical applications). However, AI has also faced criticism and fear due to uncertainties, such as the potential consequences of automation on the labor market, raising concerns about mass unemployment. This dual nature of AI, encompassing both positive and negative aspects, is also evident in the realms ofcybersecurity and cybercrime.

Governments harness AI to bolster their capabilities, yet the same technology can be exploited for attacks against them.Although the recent surge in AI development has been driven by the private sector, particularly in customer-focused applications, sectors like defense may employ similar technologies in their operations. It is becoming increasingly challenging to differentiate between the activities of state and non-state actors. This complexity was highlighted by a series of ransomware attacks targeting public infrastructure in various countries, such as the Colonial Pipeline attack in the United States in May 2021. Moreover, programs and applications initially developed for benign purposes can be repurposed or modified with malicious intent, leading to potential harm. The dual-use nature of technology is not a new concern in the context of cybercrime or cybersecurity. However, the ways in which AI can be exploited for malicious purposes introduce novel vulnerabilities. Continuous assessment of the threat landscape is essential for developing and adapting governance mechanisms, proactive measures, and enhancing cyber resilience.To build upon existing research and deepen the understanding of how AI expands the scope of malicious activities online, this article evaluates the primary categories of AI use and abuse in a criminal context. We provide several pertinent examples to illustrate the challenges involved. Based on these examples, we present a typology that categorizes the main harmful AI-based activities. Developing knowledge and understanding of the potential malicious use and abuse of AI enables cybersecurity organizations and governmental agencies to better anticipate such incidents and enhance their preparedness against attacks. Additionally, this typology is highly valuable for structuring research efforts and identifying areas where further study is needed.

2. LITERATURE SURVEY

Wrongdoing takes many forms and occurs in many places throughout the world. Many academics have proposed an instrument to analyse the relationship between misbehaviour and socioeconomic characteristics such as unemployment, income, and educational attainment. Suhong Kim and Param Joshi [1] introduced two AI models for expectation, K nearest neighbour calculation (KNN), and the decision tree technique. When predicting wrongdoing examples and determining the wrongdoing kind, the accuracy obtained varies from 39 to 44 percent.Benjamin Frederick In order to transmit more facts, David. H [2] compelled an information mining technique that includes examining and reviewing massive previous datasets. New examples are extracted and cross-checked against pre-defined datasets.In order to predict wrongdoings, Shraddha S. Kavathekar [3] used affiliation rule mining. Deep Neural Network (DNN) and Artificial Neural Network (ANN) calculations have been deduced using machine learning. Using the element level dataset, a deep brain network functions more precisely. The expectation model was built with DNN using entirely associated convolution layers, mostly for multi-named information characterization. It was done with Tenserflow, which is an API designed for deep learning methods with dropout layers. These findings suggest that when there is a high number of missing qualities, there is a need for pre-handling because wrongdoings do not occur in a uniform manner but rather cluster around a few distinct areas. The Anti-Counterfeit Neural Network (ANN) is based on pattern recognition.It necessitates a significant amount of handling constituents in order to construct a model. In removing the highlights for information management using cloud figure, Chandy and Abraham

[4] created an irregular woodland classifier. The demand number, customer ID, expiry time, season of appearance, and other variables are separated.Memory is a must. The expectation of duty is completed after highlight extraction by employing the prepared information from the learning stage.that enables you to become acquainted with the intricacies of the extricatedhighlights from the client's request

Rohit Patil, Muzamil Kacchi, Pranali Gavali, and Komal Pimparia [5] offer an Apriori calculation for ceaseless examples, which is based on the results of K-implies. Because of the rise in crime rates in recent years, the framework must cope with a massive amount of data, necessitating a bigger expenditure in physically inspecting them. As a result, advanced machine learning techniques such as K means bunching were used.
existing oblivious schemes**.**

To build on previous work and expand the understanding of how AI broadens the potential for malicious activities online, this article evaluates the main categories of use and abuse of AI in a criminal context.We provide several salient examples that allow us to illustrate the challenges at hand. Based on these examples, we present a typology that catalogs the main harmful AI-based activities. Developing knowledge and understanding about the potential malicious use and abuse of AI enables cybersecurity organizations and governmental agencies to anticipate such incidents and increase their preparedness against attacks. Furthermore, a typology is greatly useful in structuring research efforts and identifying gaps in knowledge in areaswhere more research is warranted.

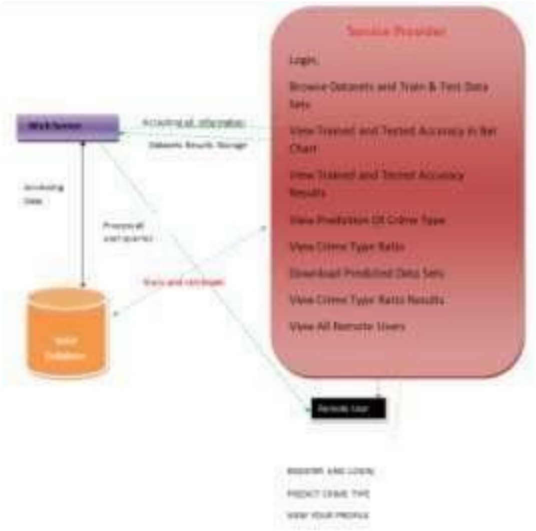

Service Provider
Login
Browse Datasets and Train & Test Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of Crime Type
View Crime Type Ratio
Download Predicted Data Sets
View Crime Type Ratio Results
View All Remote Users
Web Server
Datasets Results Storage
Accessing Data
Process all user queries
Remote User
REGISTER AND LOGIN
PREDICT CRIME TYPE
VIEW YOUR PROFILE

Fig 1. Architecture

3.METHODOLOGY

With the typology presented in this paper, we hope to make the following contributions: a. Add to the emerging body of knowledge that maps types of malicious use and abuse of AI systems. To understand the main concepts, threat scenarios, and possibilities is necessary to develop much-needed preventive measures and proactive responses to such attacks.

b. Help in establishing a shared language among and across different disciplines, especially between STEM disciplines and legal practitioners, as well as policymakers. Interdisciplinary research on the topic can reduce confusion caused by excessively technical or monodisciplinary language and aid in bridging existing gaps.

c. Propose mitigation strategies, as well as demonstrating that a collective effort among government, academia, and industry is needed.

The methodology is based on an analysis of the available literature on cybercrime and the potential malicious use and abuse of AI systems. A literature review informs this study and findings using the following databases: IEEE Xplore, Science Direct, Wiley Online Library, and Google Scholar. We used keywords, titles, and screened abstracts. The search terms included are (Artificial Intelligence OR AI OR Machine Learning OR ML) AND (malicious OR crime OR harmful OR cyber attack). Additionally, we examined lists of references obtained from reviewed papers and reports, as well as news sources describing past AI incidents. We only reviewed papers/reports/web pages available in English and Portuguese. After analyzing these sources, we were able to identify the different types of malicious use and abuse of AI

systems.Machine learning (ML) has become more prevalent in recent years. This has created incentives for attackers to manipulate models (e.g., the software itself) or the underlying data, making ML models prone to integrity attacks. In integrity attacks, hackers attempt to inject false information into a system to corrupt the data, undermining their trustworthiness.

4.IMPLEMENTATION

Data Owners

In this module, the data provider uploads their encrypted **Owners** data in the Cloud server. For the security purpose the user encrypts the data file and then store in the server. The User can have capable of manipulating the encrypted data file and performs the following operations Register and Login, Upload Blocks, Verify Block (Data Auditing), Update Block, Delete File, View Uploaded Blocks.

Cloud Server

The **Cloud** server manages which is to provide data storage service for the Data Owners.Data owners encrypt their data files and store them in the Server for sharing with data consumers and performs the following operations such as Login, View Data Owners, View End Users, View Hash Table, View File Request, View Transactions, View Attackers, View Results, View File Time Delay Results, View File Throughput Results.

End User

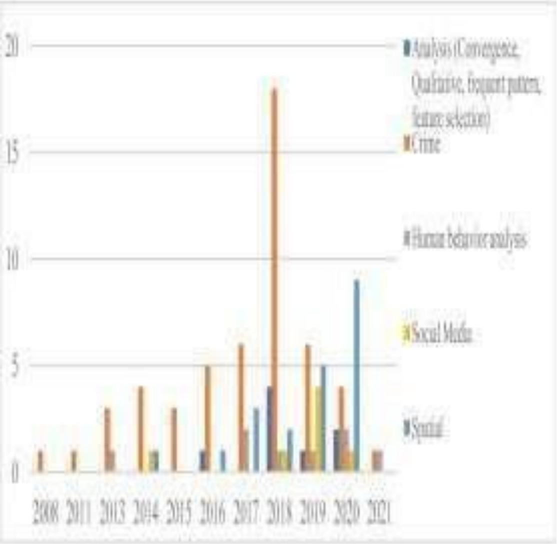
In this module, the user can only access the data file with the secret key. The user can search the file for a specified keyword and end user and can do the following operations like Register and Login, View All Data Owner Files, Request File, View File Response, Download File.

Auditor

In this module, the key issuer performs the following operations Login, View Hash Table, View Attackers, View File Updatedor Deleted, View Results.

5.RESULT

Artificial Intelligence (AI) has revolutionized many sectors, but its rapid advancement has also created opportunities for criminal exploitation and abuse. Criminals can use AI to commit various offenses, such as fraud, identity theft, and cyberattacks, by generating realistic fake identities, creating sophisticated malware, and automating hacking processes. AI-driven tools can scan for system vulnerabilities and adapt to bypass security measures, making them potent instruments for cybercriminals. Furthermore, AI can be misused to conduct unauthorized surveillance, track individuals, and even weaponize autonomous systems for smuggling or physical attacks. The technology also poses significant risks in the form of deepfakes, which can produce convincing fake videos and audio, leading to reputational damage, extortion, and fraud. Criminal organizations can exploit AI for illegal activities such as drug and human trafficking, using AI to optimize operations and evade law enforcement. Moreover, AI can be used to manipulate financial markets, spread misinformation, and influence public opinion through fake news and propaganda. The rapid pace of AI development often outstrips regulatory frameworks, creating challenges for law enforcement agencies in keeping up with new threats. As AI technology continues to advance, it is essential to develop robust measures to mitigate these risks and ensure ethical and responsible use of AI to prevent its exploitation for criminalactivities.



6. CONCLUSION

In this paper, we introduce an efficient, secure, and privacy-preserving mobile cloud storage (MCS) solution. Our proposed scheme effectively protects both data and access patterns simultaneously. Compared to existing approaches, our solution features smaller item sizes, lightweight client-side computation, and a constant communication overhead. Additionally, we account for temporal locality to enhance the scheme's efficiency further. By incorporating an additional method, our scheme is verifiable, providing resistance against malicious cloud service providers.As a fundamental component of our MCS scheme, we also present an oblivious selection and update protocol. This protocol allows a client to obliviously select and update one of its encrypted data items stored in the cloud using a small vector, requiring minimal client computation and communication. We believe this protocol holds independent value for other secure multi- party computation scenarios.Security and privacy analyses, along with formal proofs, demonstrate that our scheme ensures data confidentiality and maintains a high level of privacy preservation. Finally, we compare our scheme with two other oblivious storage solutions and thoroughly evaluate our construction in a simulated environment. The results indicate that our scheme is significantly more efficient and delivers superior performance.

7. REFERENCES

1 Suhong Kim, Param Joshi, Parminder Singh Kalsi, and Pooya Taheri, "Crime Analysis Through Machine Learning," IEEE Transactions, November 2018.

2 Benjamin Fredrick David. H and A. Suruliandi, "Survey on Crime Analysis and Prediction Using Data Mining Techniques," ICTACT Journal on Soft Computing, April 2012.

3 Shruti S. Gosavi and Shraddha S. Kavathekar, "A Survey on Crime Occurrence Detection and Prediction Techniques," International Journal of Management, Technology and Engineering, Volume 8, Issue XII, December 2018.

4 Chandy, Abraham, "Smart Resource Usage Prediction Using Cloud Computing for Massive Data Processing Systems," Journal of Information Technology, Vol. 1, No. 02.

Rohit Patil, Muzamil Kacchi, Pranali Gavali, and Komal Pimparia, "Crime Pattern Detection, Analysis & Prediction Using Machine Learning," International Research Journal of Engineering and Technology (IRJET), e-ISSN: 2395-0056.