Artificial Intelligence Crime: An Overview of Malicious Use and Abuse of AI

ABSTRACT

The capabilities of Artificial Intelligence (AI) evolve rapidly and affect almost all sectors of society. AI has been increasingly integrated into criminal and harmful activities, expanding existing vulnerabilities, and introducing new threats. This article reviews the relevant literature, reports, and representative incidents which allows to construct a typology of the malicious use and abuse of systems with AI capabilities. The main objective is to clarify the types of activities and corresponding risks. Our starting point is to identify the vulnerabilities of AI models and outline how malicious actors can abuse them. Subsequently, we explore AI-enabled and AI-enhanced attacks. While we present a comprehensive overview, we do not aim for a conclusive and exhaustive classification. Rather, we provide an overview of the risks of enhanced AI application, that contributes to the growing body of knowledge on the issue. Specifically, we suggest four types of malicious abuse of AI (integrity attacks, unintended AI outcomes, algorithmic trading, membership inference attacks) and four types of malicious use of AI (social engineering, misinformation/fake news,

hacking, autonomous weapon systems). Mapping these threats enables advanced reflection of governance strategies, policies, and activities that can be developed or improved to minimize risks and avoid harmful consequences. Enhanced collaboration among governments, industries, and civil society actors is vital to increase preparedness and resilience against malicious use and abuse of AI.

Machine learning is an important component of the growing field of data science. Through the use of statistical methods, different type of algorithms is trained to make classifications or predictions, and to uncover key insights in this project. These insights subsequently drive decision making within applications and businesses, ideally impacting key growth metrics.

Machine learning algorithms build a model based on this project data, known as training data, in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of datasets, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks.

**EXISTING SYSTEM**

To build on previous work [14]\_[16] 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 areas where more research is warranted.

**Disadvantages**

* An existing methodology not proposed the term ``AI-Crime'' to describe the situation in which AI technologies are re-oriented to facilitate criminal activity.
* An existing system doesn't implement for MALICIOUS ABUSE OF AI and VULNERABILITIES OF AI MODELS.

Proposed System

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.

**Advantages**

* The system aims to propose a typology of the malicious use and abuse of AI based on empirical evidence and contemporary discourse, analyzing how AI systems are used to compromise confidentiality, integrity, and data availability.
* Objectives are limited to identifying essential elements of the malicious use and abuse of AI, and to collect evidence of their use in practice. The compiled data enable further analysis of the possible ways in which AI systems can be exploited for criminal activities.

**SYSTEM REQUIREMENTS**

➢ **H/W System Configuration:-**

➢ Processor - Pentium –IV

➢ RAM - 4 GB (min)

➢ Hard Disk - 20 GB

➢ Key Board - Standard Windows Keyboard

➢ Mouse - Two or Three Button Mouse

➢ Monitor - SVGA

**SOFTWARE REQUIREMENTS:**

* **Operating system :** Windows 7 Ultimate.
* **Coding Language :** Python.
* **Front-End :** Python.
* **Back-End :** Django-ORM
* **Designing :** Html, css, javascript.
* **Data Base :** MySQL (WAMP Server).