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

Abstract

Artificial Intelligence (AI) has rapidly evolved and significantly impacted various sectors of society. While this powerful technology offers potential benefits, it has also been increasingly utilized for criminal and harmful purposes, exacerbating existing vulnerabilities and introducing new threats. In this article, an in-depth exploration of this phenomenon is conducted, drawing from relevant sources such as literature, reports, and real-world incidents to create a comprehensive classification of the misuse of AI-powered systems. The primary objective is to illuminate the various AI-related activities and the risks linked to them. The process begins with an examination of the inherent weaknesses in AI models, followed by an explanation of how malicious individuals can exploit these vulnerabilities. Also, delving into AI-driven attacks, which are made possible or more potent due to AI technology. Although the aim is to provide a broad overview, there is no intention to present an exhaustive categorization. Instead, the objective is to offer a comprehensive view of the risks stemming from the increased use of AI, contributing to the growing knowledge on this pressing issue. Specifically, four categories of malicious AI misuse are proposed, encompassing attacks on integrity, unintended AI consequences, algorithmic trading manipulation, and membership inference attacks. Moreover, four types of malicious AI usage are identified, covering social engineering, the dissemination of misinformation and fake news, hacking, and the development and deployment of autonomous weapon systems. This thorough assessment of threats forms a basis for more advanced discussions on governance strategies, policies, and proactive measures that can be developed or refined to mitigate these risks and prevent adverse outcomes. It underscores the urgent need for enhanced cooperation among governments, industries, and civil society stakeholders to strengthen preparedness and resilience against the malicious exploitation of AI.

Guided by
Prof. Maria Yesudas
Assistant Professor
Dept. of CSE

Submitted by
Allwina Anna Soy Jose
SJC20CS021
S7-CSE-B-10