MALICIOUS USES AND ABUSES OF ARTIFICIAL INTELLIGENCE

As of 2020, 37% of businesses and organizations use artificial intelligence (AI) in some form within their systems and processes. Its value is evident — especially to the criminals who aim to misuse and abuse AI and machine learning (ML) technologies.

In our research, a joint effort among Trend Micro, the United Nations Interregional Crime and Justice Research Institute (UNICRI), and Europol, we aim to raise awareness on the importance of **better understanding how AI and ML can be used and be misused by criminals**. Aside from looking into the different malicious scenarios using AI and ML at present and possibly in the future, **our research also looks into the possible countermeasures and recommendations to mitigate AI- and ML-powered risks, threats, and attacks**.

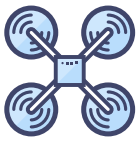
ML, AI Exploitation at Present

One of the more popular and visible malicious uses of AI is **deepfake technology**, which involves the use of AI techniques to craft or manipulate audio and visual content for the purpose of making such content appear authentic.

AI can be abused to thwart CAPTCHA security systems, which are commonly used on websites to hinder malicious activity. XEvil 4.0, a tool that uses neural networks to solve CAPTCHAs, is advertised on Russian underground forums and can be rented out to users.

ML, AI Exploitation in the Future

Cybercriminals can exploit ML-guided image recognition training in autonomous vehicles. In a research conducted by a security company, putting a small piece of black tape on a 35-mile-per-hour speed limit sign duped the image recognition model into mistaking “35” for “85.”



AI-powered facial recognition drones carrying a gram of explosive are now currently being developed. These drones are specifically for micro-targeted or single-person bombings. They are also usually operated via cellular internet and designed to look like insects or small birds. It is safe to assume that this technology will be used by criminals in the near future.

Recommendations

Through our research, stakeholders can **actively anticipate, prevent, mitigate, and properly respond to** the effects of ML- and AI-enabled malicious activities, attacks, threats, and risks in their environments and systems. We also believe that close cooperation between the industry and the academia will help raise awareness on and develop a body of knowledge about the potential use and abuse of AI by criminals. Learning about the capabilities, scenarios, and attack vectors that exploit ML and AI is key to enhancing preparedness, increasing resilience, and ensuring the positive uses of these technologies.

Read more about the exploitation of AI and ML technologies in our research paper at http://bit.ly/ExploitingAI.

TREND MICRO | unicri United Nations Interregional Crime and Justice Research Institute | EUROPOL EC3 European Cybercrime Centre