Principles of ethical use of AI systems in cybersecurity

kaspersky bring on the future

kaspersky.com

Kaspersky's approach to Machine Learning:

https://www.kaspersky.com/enterprise-security/wiki-section/products/machine-learning-in-cybersecurity

Our world is rapidly changing as advanced technologies play an increasingly crucial role. In particular, we are witnessing the active development of artificial intelligence (AI) which is already bringing many benefits to the world, including improved cybersecurity. With the number of new threats arising every day, it is impossible to detect all of them manually. For years, AI algorithms have been applied in cybersecurity to automate and speed up the process of threat detection, recognize anomalies and enhance the accuracy of malware identification. Notably, Kaspersky has been using machine learning (ML), which can be considered a subset of AI, in its solutions.

At the same time, the use of AI is not risk-free and thus requires a responsible approach from all parties involved. That is why, in order to lead innovation to the benefit of all stakeholders, Kaspersky sets the following ethical principles for the development and use of AI/ML in cybersecurity. We would also like to invite other cybersecurity companies to join and follow these principles.

#1 Transparency

We hold the belief that customers are entitled to be informed about the use of AI/ML technologies by a company in its products and services. Therefore, **we are committed to explaining principles of the way our solutions operate and utilize AI/MI technologies**. Within the Global Transparency Initiative, we operate a growing number of Transparency Centers across the globe, where our customers and other stakeholders can review the Kaspersky development processes, including those leveraging AI/ML technologies, and examine integrity and trustworthiness of Kaspersky's products and solutions. Following the principle of transparency, **we are committed to developing AI/ML systems interpretable to the maximum extent possible and to introduce necessary safeguards to ensure the validity of outcomes provided by these systems**.

#2 Safety

How to Secure Machine Learning in Security Systems

https://content.kaspersky-labs.com/se/media/en/business-security/enterprise/machine-learning-cybersecurity-whitepaper.pdf

How to confuse antimalware neural networks. Adversarial attacks and protection:

https://securelist.com/how-to-confuse-antimalware-neural-networks-adversarial-attacks-and-protection/102949/

Once introduced into the real world, AI/ML algorithms could be vulnerable to many forms of attacks designed to force these systems to make deliberate errors. In cybersecurity, the cost of mistakes in threat detection is high, therefore it is crucial to focus on safety and resilience when it comes to potential threats. For our AI/ML systems, **we are committed to prioritizing safety in the development and use of AI/ML systems**. This is being done by implementing rigorous measures to ensure the quality of all AI/ML systems. Key pillars of these measures include conducting security audits specific to ML/AI and 'red-teaming'; minimizing dependence on third-party datasets in the training process; an ensemble operation based on multi-layered protection design; favoring cloud-based ML technologies with the necessary safeguards instead of the models installed on clients' machines.

#3

Human control

As malware mutates through advanced tricks such as code obfuscation, packaging, encryption, dynamic code generation and etc., an expert opinion is needed especially for the analysis of Advanced Persistent Threats (APTs) and other complex threats. In order to provide the best protection, **we are committed to maintaining human control as an essential element** of all our AI/ML systems. Although our AI/ML systems are designed to operate in a self-contained and autonomous mode, their performance is being monitored continuously by specialists. With real-time access to information regarding live and incoming security threats, the experts are capable of using their knowledge to correct the work of our AI/ML systems if necessary, and adapt them to counter newly emerging cyberthreats. To provide holistic protection against ever-evolving cyberthreats Kaspersky combines ML algorithms with human expertise, backed by the big data of threat intelligence.

#4

Privacy

Kaspersky's approach toward data processing

https://www.kaspersky.com/about/data-protection

Big data plays a vital role in the implementation of AI/ML systems, where part of the data that may be used, can qualify as personal information. Accordingly, an ethical approach to the use of such data must take the privacy of individuals into account comprehensively. Therefore, **we are committed to respecting the rights of individuals to privacy**. In more concrete terms from a cybersecurity perspective, this can range broadly from limiting processing, reducing data composition, pseudonymizing or anonymizing wherever possible, ensuring data integrity, and applying other technical and organizational measures to protect data and systems, and ensure the meaningful exercise of rights – all with the aim of protecting the privacy of individuals

#5

Developed for Cybersecurity

Building and maintaining trust within the cybersecurity community and among users is paramount. Aligned with Kaspersky's core values centered around protecting individuals, organizations, and building a safe world, **we are committed to utilizing AI/ML systems solely for defensive purposes**. For us, a company's reputation and integrity are vital assets. By focusing exclusively on defensive technologies, we follow our mission and demonstrate our commitment to protecting users and their data, thereby enhancing our reputation as a responsible cybersecurity provider. We believe in a tomorrow where technology improves all of our lives. This is why we secure it, so everyone, everywhere, can benefit from the endless opportunities it brings.

#6

Open for Dialogue

Contribution to the informal dialogue under the aegis of the Open-Ended Working Group on security of and in the use of ICT:

https://docs-library.unoda.org/Open-Ended_Working_Group_on_Information_and_Communication_Technologies_-_(2021)/Kaspersky_SUBMISSION_OEWG_MAY_22.pdf

We are committed to promoting dialogue with all stakeholders in order to share best practice in the ethical use of AI. In this regard, Kaspersky stands ready for discussions with all interested parties, including within the UN (Global Digital Compact, Open-ended Working Group, Internet Governance Forum etc.) and other leading global platforms. Our stance is that it is only through ongoing collaboration among all stakeholders that we can overcome obstacles, drive innovation and open new horizons.

Kaspersky is looking forward to hearing your position on the use of AI in cybersecurity. In case of any further questions or feedback regarding the above-mentioned principles, please do not hesitate to contact us via e-mail:

TransparencyCenter@kaspersky.com.

Cyber Threats News: securelist.com
IT Security News: business.kaspersky.com
IT Security for SMB: kaspersky.com/business
IT Security for Enterprise: kaspersky.com/enterprise

www.kaspersky.com