Multimodale KI: Der sechste Sinn der Cyberabwehr

Sophos X-Ops erzielt beste Erkennungsraten beim Einsatz multimodaler KI

Künstliche Intelligenz (KI) ist bei der heutigen Bedrohungslage keine Option in der Cyberabwehr, sondern vielmehr Pflicht. Aber auch hier muss die Entwicklung stetig voranschreiten, um den Cybergangstern den entscheidenden Schritt beim Katz- und Mausspiel voraus zu sein. In diesem Zusammenhang hat Younghoo Lee, Principal Data Scientist bei Sophos X-Ops, den Wirkungsgrad multimodaler KI zur noch besseren Erkennung und Klassifizierung von Spam, Phishing und unsicheren Webinhalten näher untersucht.

Überwachung multipler Datenströme
Die multimodale KI ist ein System, das verschiedene Datentypen in ein einheitliches Analyse-Framework integriert. Sie stellt einen bedeutenden Wandel in der Entwicklung und Nutzung von KI in der Cyberabwehr dar, indem sie anstelle der herkömmlichen Einzelmodus-Analyse mehrere Datenströme gleichzeitig verarbeiten und Daten aus mehreren Eingaben synthetisieren kann. Damit ist es jetzt möglich, sowohl Text- als auch Bildinhalte gleichzeitig zu verarbeiten und die komplexen Zusammenhänge zu antizipieren.

Beispielsweise bei der Phishing-Erkennung untersucht multimodale KI die sprachlichen Muster und den Schreibstil des Textes sowie die visuelle Wiedergabetreue von Logos und Markenelementen. Gleichzeitig analysiert sie die semantische Konsistenz zwischen Text- und Bildkomponenten. Dieser ganzheitliche Ansatz ermöglicht es dem System, komplexe Angriffe zu erkennen, die traditionellen Systemen möglicherweise legitim erscheinen. Darüber hinaus lernt die multimodale KI aus den Zusammenhängen zwischen verschiedenen Datentypen und passt sich an diese automatisch an.

Hoch wirksam bei der Erkennung
Die Wirksamkeit multimodaler KI ist im Gegensatz zu herkömmlichen Machine-Learning-Modellen entscheiden höher. Zum Vergleich führte SophosAI eine Reihe empirischer Experimente mit durchschlagenden Erfolgen durch. Das Ergebnis: traditionelle Modelle zeigten gute Ergebnisse bei der Erkennung bekannter Bedrohungen, hatten jedoch Schwierigkeiten mit neuen, unbekannten Phishing-E-Mails. Ihre F1-Werte (ein Maß für die Präzision und Trefferquote zwischen 0 und 1) lagen bei unbekannten Proben bei nur 0,53 und erreichten einen Höchstwert von 0,66. Die multimodale KI (mit GPT-4o) schnitt bei den Versuchen für die Erkennung neuer Phishing-Versuche sehr viel besser ab und erreichte selbst bei unbekannten Marken F1-Werte von bis zu 0,97.

„KI ist eine wichtige Komponente in der Cyberabwehr und sorgt im Verbund mit der rein technischen Abwehr am Endpoint und der nach wie vor nötigen menschlichen Erkennung für einen sehr guten Schutz“, sagt Michael Veit, Security-Experte bei Sophos. „In Verbindung mit dem Cybersecurity-Ökosystem von Sophos repräsentiert die multimodale KI einen weiteren Meilenstein und wird die Cyberabwehr auf ein deutlich höheres Level bei der Erkennung heben.“

Social Media von Sophos für die Presse
Wir haben speziell für Sie als Journalist*in unsere Social-Media-Kanäle angepasst und aufgebaut. Hier tauschen wir uns gerne mit Ihnen aus. Wir bieten Ihnen Statements, Beiträge und Meinungen zu aktuellen Themen und natürlich den direkten Kontakt zu den Sophos Security-Spezialisten.

Folgen Sie uns auf und

LinkedIn: https://www.linkedin.com/groups/9054356/
X/Twitter: @sophos_info

Pressekontakt:

Sophos
Jörg Schindler, Senior PR-Manager EMEA Central
joerg.schindler@sophos.com, +49-721-25516-263

TC Communications
Arno Lücht, +49-8081-954619
Thilo Christ, +49-8081-954617
Ulrike Masztalerz, +49-30-55248198
Ariane Wendt +49-172-4536839
sophos@tc-communications.de