Innovationswettbewerb KI & Cybersicherheit Baden-Württemberg (2022)

Projektsteckbrief

AntiDot – Online-Werkzeuge zur Bekämpfung von Data Poisoning in KI



Baden-Württemberg
MINISTERIUM FÜR WIRTSCHAFT, ARBEIT UND TOURISMUS

Worum geht es: Im Projekt *AntiDot* werden effektive Methoden und Werkzeuge entwickelt, um KI-Systeme gegen Cyberangriffe mit manipulierten Daten, das sog. Data Poisoning, zu schützen. Die Bereitstellung dieser Werkzeuge als „Software-as-a-Service“ soll die Einführung von Absicherungs- und Abwehrmaßnahmen für Unternehmen erheblich vereinfachen. Im Fokus des Projekts steht die Absicherung von KI-Anwendungen in der Automobilindustrie und im Anlagen- und Maschinenbau.

Durchgeführt von: Asvin GmbH, Stuttgart

Innovationswettbewerb KI & Cybersicherheit Baden-Württemberg

In künstlicher Intelligenz (KI) steckt viel Potenzial für innovative Produkte, Dienstleistungen und Geschäftsmodelle – und zwar quer durch alle Branchen. Das eröffnet Firmen aus Baden-Württemberg neue Chancen für Wertschöpfung und Wachstum. Wettbewerbsvorteile entstehen insbesondere dann, wenn KI-Knowhow gezielt mit Branchenwissen kombiniert wird, um neuartige Lösungen zu schaffen.

Zugleich wird in einer zunehmend digital vernetzten Welt der wirksame Schutz vor Cyberangriffen immer wichtiger. Deshalb hat das Wirtschaftsministerium Baden-Württemberg einen Innovationswettbewerb ausgeschrieben, mit dem Unternehmen bei der Entwicklung von neuartigen Produkten und Dienstleistungen zur Abwehr von Cyberangriffen gefördert werden. Im Fokus der geförderten Projekte stehen Innovationen, bei denen KI-Technologien zum Einsatz kommen oder die dazu dienen, KI-Systeme sicherer zu machen.

Der Innovationswettbewerb KI & Cybersicherheit ist eine Maßnahme im Rahmen des Aktionsprogramms KI für den Mittelstand des Ministeriums für Wirtschaft, Arbeit und Tourismus Baden-Württemberg.

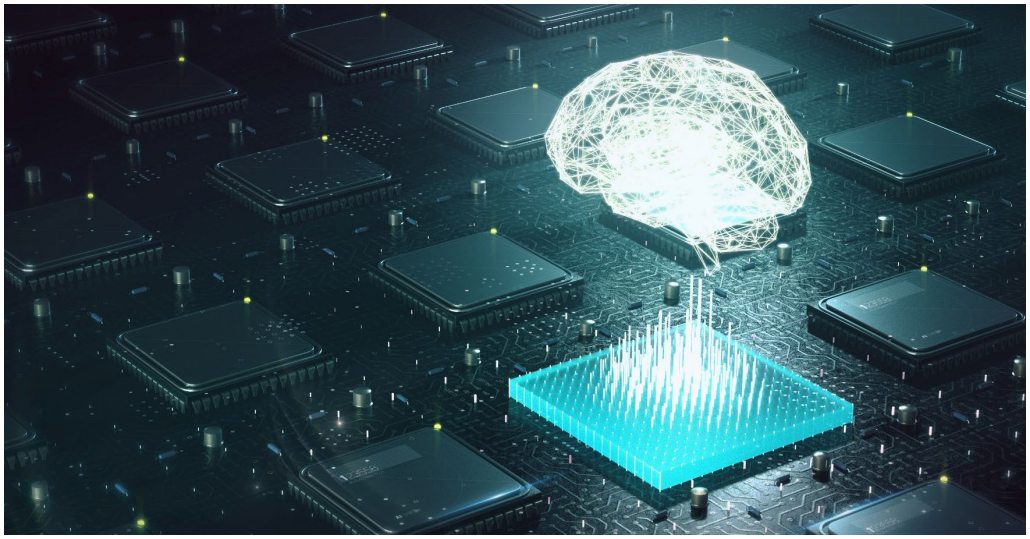
KI-Systeme wirksam gegen Cyberangriffe mit manipulierten Trainingsdaten schützen

Mit der zunehmenden Verbreitung von KI-Systemen wächst auch der akute Bedarf zu deren Absicherung. Cyberangriffe durch Data Poisoning, also das böswillige Einschleusen von manipulierten Daten in KI-Systeme, stellen eine potentielle Gefährdung für eine Vielzahl von KI-Anwendungen dar.

Data Poisoning erlaubt es Angreifern, die Entscheidungen von KI-Anwendungen gezielt zu beeinflussen. Dadurch können beispielsweise industrielle Produktionsanlagen, die mittels KI-Technologien über Digitale Zwillinge gesteuert werden, sabotiert werden. Und falsche Entscheidungsfindungen beim autonomen Fahren können Gesundheit und Leben von Verkehrsteilnehmern gefährden.

AnitDot verfolgt das Ziel, KI-Systeme gegen Cyberangriffe mit Data-Poisoning-Methoden zu schützen und den Einsatz von innovativen Online-Werkzeugen zur Absicherung zu erproben und schrittweise zu verbessern. Das Projekt liefert Ergebnisse zu folgenden wissenschaftlich-technischen Innovationszielen:

- die automatisierte Konfiguration von Sicherheitsmechanismen zur Erkennung und Abwehr von Data-Poisoning-Angriffen im maschinellen Lernen;



- Prognosen zur Qualitätsbeeinflussung der KI-Aussagen bei identifizierten Verschmutzungen der Trainingsdaten;
- die Entwicklung von praxistauglichen und leicht zu bedienenden Online-Werkzeugen, mit denen KI-Anwendungen davor geschützt werden können, dass mittels eingeschleuster Daten die Aussagen im Supervised Learning manipuliert werden;
- die Entwicklung von Methoden zur „Neutralisation“ von sog. Adversarial Attacks, bei denen durch Rauschmuster kompromittierte Bilddaten beim Deep Learning von neuronalen Netzen eingeschleust wurden.

Die Projektergebnisse werden im Rahmen von Open Science und Open Source bereitgestellt. Die entwickelten Lösungen und Methoden stehen so für eine breite Verwendung in Wissenschaft und Wirtschaft zu Verfügung.

Leicht zu bedienende Online-Werkzeuge senken Einstiegshürden für KI-Anwender bei KI-Sicherheit

asvin wird im Rahmen von *AntiDot* den Einsatz der experimentellen Abwehrwerkzeuge als leicht zu bedienender „Software-as-a-Service“ erproben und online anbieten. Damit können KI-Trainingsdaten auf Abruf („on demand“) auf Data Poisoning überprüft und Angriffe entschärft werden. Dadurch sollen die technischen Hürden bei der Einführung von Cybersicherheits-Werkzeugen für KI gegenüber komplizierten „on premise“-Expertensystemen drastisch gesenkt und ein Alleinstellungsmerkmal geschaffen werden.

Die entwickelten Lösungen zur Sicherung digitaler Zwillinge und zum Schutz des autonomen Fahrens sind u.a. für Anwendungen in den Bereichen Industrie 4.0, Mobilität, Gesundheitswesen und Energiewirtschaft interessant.

Kontakt

Asvin GmbH
Mirko Ross (CEO)
Konrad Buck (Pressesprecher)
Christian Billmann (Marketing)
Schulze-Delitzsch-Straße 16
70565 Stuttgart

https://asvin.io
m.ross@asvin.io
k.buck@asvin.io
c.billmann@asvin.io

Gefördert durch

Ministerium für Wirtschaft, Arbeit und Tourismus Baden-Württemberg
Postfach 10 01 41
Schlossplatz 4 (Neues Schloss)
70001 Stuttgart
Tel: 0711 123-2869
Fax: 0711 123-2871
pressestelle@wm.bwl.de
www.wm.baden-wuerttemberg.de

Projektwebsite und weitere Informationen

https://antidotlab.com

Quellenhinweis

Weitere Informationen zum Innovationswettbewerb finden Sie auf der

Website der Initiative Wirtschaft 4.0 Baden-Württemberg

Initiative Wirtschaft 4.0 BW

Baden-Württemberg
MINISTERIUM FÜR WIRTSCHAFT, ARBEIT UND TOURISMUS