ANONYMISIERUNG

KI-Systeme sicherer machen

Im Kompetenzcluster AnoMed entwickelt ein Konsortium neuartige anonymitätsbewahrende Datenverarbeitungsverfahren sowie eine öffentlich zugängliche Plattform, mit der Entwickler von KI-Systemen ihre verwendeten Anonymisierungsverfahren testen können.

Ob Prophylaxe, Diagnosestellung, Therapie oder Nachsorge – in allen Bereichen der Medizin fallen inzwischen riesige Mengen von Patientendaten an, die nicht nur gespeichert und verwaltet, sondern – das Einverständnis der Patienten vorausgesetzt – auch für die Forschung verwendet werden. Falls die Patienten ihr Einverständnis unter der Voraussetzung geben, dass eine anonymitätsbewahrende Datenverarbeitung erfolgt, müssen die Forscherinnen und Forscher sicherstellen, dass die veröffentlichte Studie keine Rückschlüsse auf personenbezogene Patientendaten zulässt. Eine klassische Anonymisierung der Patientendaten aber lässt solche Rückschlüsse zu, zum Beispiel durch Querreferenzierung mit anderen Daten, was nicht mehr dem Stand der Forschung entspricht. Deshalb entwickelt ein Konsortium im Kompetenzcluster AnoMed neuartige anonymitätsbewahrende Datenverarbeitungsverfahren sowie eine öffentlich zugängliche Plattform mit einer Reihe an medinischen Testaufgaben und De-Anonymisierungsangriffen, um Anonymisierungsverfahren für medizinische Anwendungen zu testen.

AnoMed ist eines von fünf Kompetenzclustern des Forschungsnetzwerks Anonymisierung für eine sichere Datennutzung, das seit Ende 2022 vom Bundesministerium für Bildung und Forschung (BMBF) gefördert wird (siehe Kasten). Die riesigen Mengen an anonymisierten Patientendaten werden vor allem für die Entwicklung von Verfahren der künstlichen Intelligenz (KI) benötigt. Beim Maschinellen Lernen zum Beispiel trainieren Wissenschaftler Algorithmen mit Röntgenbildern und weiteren medizinischen Informationen, damit sie Tumore auf den Aufnahmen

DREI FRAGEN AN

… Prof. Dr. Esfandiar Mohammadi

Wie können Unbefugte über einen Algorithmus an personenbezogene Daten gelangen?

Die jüngste Forschung hat eine Reihe neuartiger Angriffe auf KI-Techniken gefunden. Bei einem maschinellen Lernalgorithmus zum Beispiel versuchen Angreifer, Informationen über die Trainingsdaten aus dem Maschinenlernmodell zu extrahieren. Falls die Trainingsdaten Patientendaten beinhalten, können sie dadurch Rückschlüsse auf Patientendaten ziehen.

© Raphael Reischuk

Prof. Dr. Esfandiar Mohammadi, Institut für IT-Sicherheit der Universität zu Lübeck, ist wissenschaftlicher Leiter des Kompetenzclusters AnoMed.

© Raphael Reischuk

Gibt es Techniken, um diese Angriffe zu verhindern?

Ja, die gibt es. Neuartige anonymitätsbewahrende Lernalgorithmen lernen nur sehr vorsichtig aus einzelnen Datenpunkten und lernen daher nur Effekte, für die es nicht nur Evidenz von einzelnen Personen gibt, und von Ausreißern wird nur mäßig gelernt. Dazu müssen wir die Lernalgorithmen entsprechend anpassen. Bei kleinen Datensätzen kann dieses Vorgehen problematisch sein. Deswegen entwickeln wir komplementär dazu sichere verteilte Lernverfahren, die sicherstellen, dass verschiedene medizinische Institutionen, die gemeinsam Studien durchführen, gemeinsam sicher ein Modell lernen können: Jede Partei lernt dabei nur das finale Modell, nichts über die Daten der anderen Parteien.

Wie realistisch ist es überhaupt, dass ein Angreifer Daten aus einem Algorithmus extrahiert?

Wir sind aktuell noch in einer Phase der IT-Sicherheit, in der Angriffe zwar bekannt sind, es aber noch keine organisierten kriminellen Strukturen gibt, die das ausnutzt. In dieser Phase war das Internet vor 20 Jahren. Heutzutage gibt es organisierte Kriminalität, die die Schwachstellen im Internet ausnutzt, um zum Beispiel Krankenhäuser anzugreifen und zu erpressen. Wir sollten darauf vorbereitet sein, dass sich die organisierte Kriminalität früher oder später den medizinischen Algorithmen widmet. ▪

Das Interview führte Dr. Michael Lang, freier Journalist.

> **„Heute gibt es organisierte Kriminalität, die die Schwachstellen im Internet ausnutzt, um Krankenhäuser anzugreifen und zu erpressen. Wir sollten darauf vorbereitet sein, dass sich die organisierte Kriminalität früher oder später den medizinischen Algorithmen widmet.**
>
> Prof. Dr. Esfandiar Mohammadi, Institut für IT-Sicherheit der Universität zu Lübeck

mittels Mustererkennung identifizieren können. Da es theoretisch möglich ist, aus dem Algorithmus Informationen über die Röntgenbilder zu extrahieren und Rückschlüsse auf individuelle Patienten zu ziehen, sollen Entwickler die Plattform nutzen können, um ihre Algorithmen gegen Datenlecks abzusichern.

Testplattform für Algorithmen

Aktuell bauen die beteiligten Forscherinnen und Forscher die Infrastruktur für die Testplattform mit einem rechnungsstarken GPU-Cluster auf – einem auf KI-Verfahren ausgelegter Rechnerverbund; im Projekt sollen auch Algorithmen untersucht werden, die medizinische Bilder analysieren. Außerdem implementieren die Sicherheitsexperten von AnoMed Anonymitäts- und Nützlichkeitstests. Geplant ist, dass Entwickler aus aller Welt ihre Anonymisierungslösung in Form von Algorithmen auf die Plattform hochladen. Die Experten von AnoMed untersuchen dann im Rahmen von Studien mit simulierten Angriffen, ob der zur Prüfung eingereichte Algorithmus den De-Anonymisierungsangriffen standhält und falls nicht, wo die Schwachstellen liegen. Damit die De-Anonymisierungstests einmal vollautomatisch ablaufen können, werden die Experten von AnoMed auch untersuchen, welche indirekten Rückschlüsse sie aus den resultierenden Studien ziehen können, und wie potenzielle Angreifer sogenannte Seitenkanalinformation dabei nutzen könnten, um

sensible Daten zu extrahieren. Bei einer Seitenkanalattacke wird der Algorithmus nicht direkt angegriffen. Stattdessen nutzt der Angreifer Erkenntnisse, die er aus beobachtbaren physikalischen Effekten bei der Verarbeitung der Daten gewinnt. Darüber hinaus überprüfen die Forscherinnen und Forscher, ob der Algorithmus nach der Anonymisierung noch brauchbare Ergebnisse liefert.

Geplant sind auch Tests mit den Daten von MMIC-III, einer weltweit frei zugänglichen Datenbank mit standardisierten medizinischen Informationen von über 40 000 Patienten, die zwischen 2001 und 2012 auf der Intensivstation des Beth Israel Deaconess Medical Center in Boston, Massachusetts, behandelt wurden. MMIC-III enthält unter anderem Daten zu Demografie, Vitalparametern, Laborergebnissen, Medikamenten und zur Sterblichkeit. Alle Daten wurden vor Aufnahme in die Datenbank de-identifiziert, indem beispielsweise personenbezogene Daten gelöscht und Datumsangaben in die Zukunft verschoben wurden.

Bessere Anonymisierungslösungen

Eine zweite Säule von AnoMed befasst sich mit der Verbesserung von Anonymisierungslösungen. Dabei wollen die Forscherinnen und Forscher insgesamt sieben Klassen von medizinischen Algorithmen so verändern, dass sie Daten nur noch spärlich verwenden und so den Schutz der Privatsphäre nach dem Stand der Technik bewahren. Die zu entwickelnden Lösungen reichen von anonymitätsbewahrenden Lernverfahren, über spezielle Hardware-gestützte Berechnungsverfahren auf verschlüsselten Daten und die Härtung des Systems bis hin zur automatischen Informationsflussanalyse. Anonymitätsbewahrende Lernverfahren zum Beispiel lernen nur sehr wenig von Ausreißern und mehr von Effekten, für die es nicht nur Evidenz von einzelnen Personen gibt.

Für Medizinprodukte ist nicht nur die Entwicklung neuartiger Algorithmen, sondern auch die Entwicklung neuer Medizintechnik, inklusive neuer Hardware-Lösungen, relevant. Deshalb beschäftigt sich AnoMed auch mit medizintechnischen Prüfungen. In diesem Rahmen stellt AnoMed FPGA-Hardware (Field Programmable Gate Array) zur Verfügung. Hierbei handelt es sich um programmierbare integrierte Schaltkreise, die zum Beispiel für die Echtzeitverarbeitung von Algorithmen verwendet werden. KI-Entwickler sollen so die Effizienz ihrer Hardware beurteilen können – ob zum Beispiel ein Medizingerät eine gute Batterielaufzeit hat und trotzdem noch eine gute Nützlichkeit und Anonymisierung hinbekommt.

Komplementär zu den technischen Betrachtungen wird sich AnoMed mit den rechtlichen Fragestellungen beschäftigen, die im Zusammenhang mit den anonymitätsbewahrenden Lernverfahren auftreten. Die am Projekt Beteiligten wollen so ein tieferes Verständnis der rechtlichen Auswirkungen der neuartigen technischen Lösungen erarbeiten. Zum Beispiel wollen die Fachleute das Restrisiko im Kontext der aktuellen Gesetzte diskutieren. ▬

Dr. Michael Lang
Freier Journalist

Projekt AnoMed

Koordinator: Universität zu Lübeck

Projektpartner:
Unabhängiges Landeszentrum für Datenschutz, Kiel
Eppdata, Hamburg
Ingrano Solutions, Berlin
Perfood, Lübeck
UniTransferKlinik Lübeck
DFKI, Kaiserslautern
Universität Hamburg
Universitätsklinikum Schleswig-Holstein, Lübeck
Fraunhofer IMTE, Lübeck

Volumen: rund 10 Millionen Euro

Laufzeit: November 2022 bis November 2025