UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR IT-SICHERHEIT

Übungsblatt 12

Machine Learning Security

Abgabe bis 11. Juli 2023 um 12:00 Uhr im Moodle

Cybersecurity im Sommersemester 2023

Thomas Eisenbarth, Esfandiar Mohammadi, Jan Wichelmann, Anja Köhl

Aufgabe 1 Verständnisfragen (8 Punkte)

1. Gegeben sei ein Bild X. Was ist eine irreführende Modifikation (*Adversarial Example*) Y von Bild X für einen gegebenen Classifier C?
   i. Eine Modifikation von X, sodass C etwas anderes für X als für Y ausgibt.
   ii. Eine Modifikation von X, sodass C etwas anderes für X als für Y ausgibt, während Y für einen Menschen jedoch wie X aussieht.
   iii. Eine Modifikation von X, sodass C etwas anderes für X als für Y ausgibt, wobei über Y mehr Informationen preisgegeben werden als über X.
2. Warum gibt Federated Learning mehr Information preis als 'Train & Pass'?
3. Wie kann ein Angreifer sich gegen Outlier Detection schützen?

Aufgabe 2 Wiederholung: k-Anonymität (6 Bonuspunkte)

Diese Aufgabe muss nicht bearbeitet werden um den Zettel zu bestehen, liefert jedoch Bonuspunkte, die auf einen anderen nicht bestandenen Zettel angerechnet werden können.

Eine Arbeitsgruppe aus Mitarbeitern und Studenten soll bei der Verbesserung einer Lehrveranstaltung helfen. Hierzu soll ihnen eine Datenbank zur Verfügung gestellt werden, die Klausurergebnisse mit Altersgruppen und Studiengängen verknüpft:

Name	Studiengang	Alter	Klausurnote
Angela Gillie	MML	21	2,7
Bella Eichorn	MML	18	2,7
Carolin Blanchard	MML	19	2,3
Daniel Hickmon	MML	21	3,3
Eva Samuelson	ITS	25	2,0
Fabian Vento	ITS	27	3,3
Gina Jantzen	ITS	22	1,7
Helen Anstett	ITS	21	3,0
Iliana Kiernan	MI	19	3,0
Jasmin Janeway	MI	18	2,0

Um die in der Datenbank enthaltenen Personen zu schützen, wird diese anonymisiert, indem die Namen entfernt und die Altersgruppen zusammengefasst werden:

Name	Studiengang	Alter	Klausurnote
*	MML	18-21	2,7
*	MML	18-21	2,7
*	MML	18-21	2,3
*	MML	18-21	3,3
*	ITS	25-27	2,0
*	ITS	25-27	3,3
*	ITS	21-22	1,7
*	ITS	21-22	3,0
*	MI	18-21	3,0
*	MI	18-21	2,0

1. Erklären Sie kurz den Begriff k-*Anonymität*. Welches k hat die obige anonymisierte Datenbank, wenn Sie annehmen, dass Studiengang und Altersgruppe quasi-identifizierende Attribute sind?

2. Nennen und erklären Sie die Schwächen von k-Anonymität.

   Orientieren Sie sich hierbei an folgendem Beispiel: Die Studentin Sophie ist Mitglied der Arbeitsgruppe und hat damit Zugriff auf die anonymisierte Datenbank. Sie weiß, dass Iliana MI studiert, und möchte mittels dieser Hintergrundinformation mehr über sie herausfinden. Welche Informationen kann Sie über Iliana gewinnen?

3. Welchen Wert hätte k, wenn man bei ITS alle Altersgruppen in einer Klasse zusammenfassen und jemand aus dem MML Studiengang zu MI wechseln würde?