Cybersecurity & Vertrauen wir der KI zu viel

Cybersecurity für den normalen Nutzer (Evaluation/Interview)

Ideen:

Passwort-Manager mit KI-Integration

- **Idee**: Ein Passwort-Manager, der KI nutzt, um die Sicherheit von Passwörtern zu überprüfen und Vorschläge für stärkere Passwörter zu machen. Dabei könnte die KI auch versuchen, Muster zu erkennen bzw die Passwöter zu erraten, die oft bei schwachen Passwörtern auftreten, und den Benutzer darauf hinweisen.
- **Technologien**: Machine Learning für Mustererkennung, einfache Benutzeroberfläche zur Verwaltung von Passwörtern.

Passwort-Sicherheit-Checker

- **Idee**: Ein Tool, das überprüft, ob ein Passwort sicher ist. Der Nutzer gibt sein Passwort ein, und das Tool bewertet es basierend auf Kriterien wie Länge, Verwendung von Zahlen, Sonderzeichen, Groß-/Kleinschreibung und ob es häufig verwendete oder unsichere Passwörter enthält.
- **Technologie**: Python, einfache Benutzeroberfläche mit Tkinter oder als Webanwendung mit Flask. ML-Modell nutzen, das bekannte unsichere Passwörter klassifiziert

Passwort-Knacker vs. Passwort-Manager (KI vs. KI)

- **Angreifer-KI**: Ein einfaches KI-Modell, das versucht, Passwörter zu erraten, indem es gängige Passwort-Strategien wie Brute Force oder Wörterbuchangriffe verwendet.
- **Verteidigungs-KI**: Ein Passwort-Manager, der mithilfe eines einfachen KI-Systems die Stärke von Passwörtern überprüft und sofort Sicherheitstipps gibt. Diese KI könnte auch versuchen, ungewöhnliche Login-Versuche zu erkennen und den Benutzer zu warnen.
- **Technologie**: Python, cryptography für Verschlüsselung, einfache Heuristiken zur Passwortsicherheit.
- ein KI-Modell zur Passwortbewertung verwenden und den Angreifer mit einer Liste häufiger Passwörter und Brute-Force-Algorithmen konfrontieren.

KI-gestütztes Sicherheits-Tutorial für Anfänger

- I**dee**: Ein interaktives Tool, das den Nutzer durch grundlegende Sicherheitspraktiken führt, z. B. wie man sichere Passwörter erstellt, wie man Phishing erkennt, wie man Software sicher installiert und aktualisiert usw. KI könnte personalisierte Empfehlungen basierend auf den Eingaben des Nutzers machen.

- **Technologien**: Interaktive Benutzeroberfläche, basierend auf einfachem Chatbot- oder Dialogsystem, das durch die wichtigsten Themen führt.

Sicherheits-Tipps-Chatbot für Anfänger

- **Idee**: Erstelle einen einfachen Chatbot, der den Nutzer durch grundlegende Sicherheitspraktiken führt. Der Chatbot könnte dem Benutzer Fragen zu seinem Verhalten in Bezug auf digitale Sicherheit stellen und ihm einfache Tipps und Best Practices geben. Zum Beispiel: "Hast du ein sicheres Passwort?" oder "Hast du deine Software auf dem neuesten Stand?"
- **Technologie**: Python mit einer Bibliothek wie NLTK oder ChatterBot für einfache Chatbot-Funktionen. einfache natürliche Sprachverarbeitung (NLP)
- mit einfachen Regeln und Antworten ohne komplexe ML-Modelle?

Spam-Filter vs. Spam-Versender (KI vs. KI)

- **Angreifer-KI**: Diese KI simuliert einen „Spam-Versender“, der versucht, Spam-E-Mails zu generieren, um sie in die Postfächer der Benutzer zu schleusen.
- **Verteidigungs-KI**: Eine Spam-Filter-KI, die den Eingang von E-Mails überwacht und diese auf typische Spam-Muster überprüft (z.B. unseriöse Absender, verdächtige Wörter, große Anhänge).
- **Technologie**: Python, sklearn, NLTK für Textverarbeitung.

Verhaltensbasierte Anomalie-Erkennung auf dem Computer

- **Idee**: Ein Tool, das das Verhalten eines Computers überwacht (z. B. unübliche Dateioperationen, verdächtige Netzwerkaktivitäten) und mithilfe von KI Anomalien erkennt, die auf einen möglichen Sicherheitsvorfall hinweisen könnten.
- **Technologien**: Anomaly Detection-Algorithmen, z. B. Isolation Forest oder Autoencoders.

Verhaltensbasierte Anomalie-Erkennung im Verhalten des Users

- **Idee**: Ein Tool, das das Verhalten des Nutzers speichert und bei auffälligen Formulierungen und Aktionen, die auf einen möglichen Sicherheitsvorfall hinweisen könnten. -> Verhaltenspasswort
- **Technologien**: Anomaly Detection-Algorithmen, z. B. Isolation Forest oder Autoencoders.

Fake-News-Generierung vs. Fake-News-Erkennung (KI vs. KI)

- **Angreifer-KI**: Entwickle eine KI, die absichtlich „Fake News“ generiert, indem sie Aussagen erstellt, die typisch für Desinformation oder Falschmeldungen sind.

- **Verteidigungs-KI**: Eine andere KI prüft die Texte auf Unstimmigkeiten, Falschinformationen oder Verlinkungen zu falschen Quellen. Du könntest ein Modell für Textklassifikation verwenden, das gefälschte Informationen erkennt.
- **Technologie**: Python, transformers (z.B. BERT oder GPT-Modelle), einfache Textklassifikation.
- vortrainierte Modelle für die Fake-News-Erkennung