



Aufgabe 3: Automatisierung einer Machine-Learning-Pipeline

In dieser Aufgabe sollen Sie ein grundlegendes Verständnis für die Automatisierung von Machine-Learning-Workflows erlangen. Dazu sollen Sie eine Pipeline erstellen, die Ihren Machine-Learning-Prozess von der Datenaufbereitung bis zur Evaluierung automatisiert. Sie können dabei auf Ihre vorherigen Lösungen (CNN, RNN) zurückgreifen oder eine neue Implementierung wählen. Ziel ist es, die Reproduzierbarkeit der Trainingsprozesse zu gewährleisten und die Effizienz durch Automatisierung zu steigern.

1. Konzeption einer ML-Pipeline

- Erläutern Sie das Konzept einer Machine-Learning-Pipeline: Welche Schritte sind wichtig und wie hängen diese zusammen?
- Stellen Sie eine Übersicht der geplanten Schritte Ihrer Pipeline vor (z. B. Datenaufbereitung, Modelltraining, Evaluation).

2. Technologie-Stack und Tools

- Wählen Sie geeignete Tools und Technologien zur Automatisierung Ihrer Pipeline aus (z. B. GitHub Actions, Jenkins, Apache Airflow).

3. Automatisierung der Pipeline

- Erstellen Sie eine Pipeline, die Ihren Machine-Learning-Prozess automatisiert.
- Implementieren Sie die Schritte zur Datenvorbereitung, zum Modelltraining und zur Evaluierung in Ihrer Pipeline.

4. Evaluierung und Monitoring

- Evaluieren Sie die Funktionsweise Ihrer automatisierten Pipeline: Was sind die Vorteile und wo lagen die Herausforderungen?

5. Visualisierung der Ergebnisse

- Visualisieren Sie die wichtigsten Ergebnisse Ihrer Pipeline, z. B. Trainingsergebnisse oder Vergleich von Vorhersagewerten.
- Dokumentieren Sie die Schritte der Automatisierung und die Auswirkungen auf die Modellleistung.