Projektidee: Analyse und Vorhersage des Erfolgs berühmter Bücher

Ziel des Teils-Projekts

# Erstelle eine interaktive Web-App mit Streamlit, um:

Daten über berühmte Bücher zu analysieren (z. B. Sprache, Genre, Jahr, Verkäufe usw.)

Eine eigene Buchidee einzugeben und mithilfe von Machine Learning vorherzusagen, ob das Buch erfolgreich sein könnte (z. B. Verkaufszahlen oder Bestsellerstatus).

? SQL-Datenbank zur Speicherung und Abfrage der Bücher zu nutzen

Python – Datenverarbeitung und Logik

Pandas / NumPy – Datenanalyse

Plotly /Seaborn – Visualisierungen

Scikit-learn oder XGBoost – ML-Modell zur Vorhersage

SQLite oder PostgreSQL – relationale Datenbank mit SQL-Abfragen

Funktionen der Streamlit-Web-App

1. Datenanalyse & Visualisierung

Filter nach Sprache, Genre, Jahr

Diagramme:

Verkäufe pro Genre

Top-Autoren nach Verkaufszahlen

Entwicklung der Bestseller über Jahrzehnte

Heatmap: "Genre × Jahr"

2. Machine Learning: Erfolgs-Vorhersage

Benutzer gibt Buchdaten ein (Genre, Sprache, Seitenanzahl, Verfilmung usw.)

Modell prognostiziert:

Verkaufszahlen (Regression)

Oder: Bestseller (ja/nein) – Klassifikation

Erklärung der Einflussfaktoren (Feature-Importance)

3. ? SQLDatenbank & Abfragen

Eine Datenbankseite mit:

Filterfunktionen

Manuelle SQL-Abfragen (z. B. alle Bücher nach 2000 mit mehr als 10 Mio Verkäufen)

Beispielabfragen als Buttons

4. Eigene Bücher hinzufügen

Benutzerformular zum Hinzufügen eigener Bücher

Neue Bücher werden in SQL-Datenbank gespeichert

Live-Aktualisierung der Visualisierungen

ML-Ideen im Detail

Regression: Verkäufe (in Millionen) vorhersagen

Klassifikation: Bestseller ja/nein

Clustering: Bücher nach Ähnlichkeit gruppieren (z. B. K-Means)

# Projektstruktur

bash

Code kopieren

buchanalyse-projekt/

│

├── data/

│ └── books.csv

│

├── app/

│ ├── main.py # Streamlit-App

│ ├── ml\_model.py # ML-Modelle & Vorhersage

│ ├── database.py # SQL-Logik

│ ├── visualizations.py # Diagramme & Plots

│ └── utils.py # Hilfsfunktionen

│

├── requirements.txt

└── README.md

??? Mögliche Erweiterungen

Verbindung zu externen Buchdaten (z. B. Google Books API)

Vergleich mit Verfilmungen (Box Office vs. Buchverkäufe)

Nutzerkonten mit Favoriten & Notizen

Dashboard für Verlage zur Manuskript-Einschätzung