Personalized News Retrieval

Readme

Overview

This project implements a Personalized News Recommendation Engine using Python, Pandas, Scikit Learn, and WordCloud. The system trains a Naive Bayes classifier on a labeled news dataset and recommends articles based on the user's preferred category. It also generates a topic word cloud for visualization.

<u>Features</u>

- Upload CSV dataset dynamically using Google Colab file uploader.

- Train a TF-IDF Vectorizer + Multinomial Naive Bayes model.

- Predict news categories based on titles.

- Calculate a custom recommendation score using:

- AI confidence

- Category match

- Recency (newness) of article

- Display top 5 personalized news recommendations.

- Generate a word cloud for key topics in the selected category.

<u>How It Works</u>

- Load dataset (news_dataset.csv).

- Extract 'title' and 'category'.

- Train Naive Bayes classification pipeline.

- User inputs preferred category (sports, politics, health, tech).

- System predicts categories & computes confidence.

- System applies a scoring formula:

- Sort results and display top 5.

- Generate word cloud.

Technologies Used

- Python

- Pandas, NumPy

- Scikit-Learn (TfidfVectorizer, MultinomialNB)

- Matplotlib

- WordCloud

- Google Colab

How to Run

- Open the notebook in Google Colab.

- Upload news_dataset.csv when prompted.

- Run all cells.

- Enter your preferred category.

- View recommended news and auto-generated word cloud.