

Movie Recommendation Application Based on Reviews

Description:

The “Movie Recommendation Application” is an advanced tool designed to analyze and curate movie reviews, offering tailored recommendations based on user preferences and sentiments. As a programmer, your objective is to develop this application with precision, ensuring it accurately assesses reviews and suggests movies that align with users’ tastes and moods.

Key Features:

- First, the application will **check** whether there is “**model.pickle**” file in the application directory or not.
- If the file **exists**, then the application will **read** and **load** the **data training** from the **file**.
- If the file **doesn’t exist**, then the application will **train** the review data with **Naïve Bayes Classifier** from **NLTK data twitter sample** provided in “**imdb-movie-dataset.csv**”. The **data training** will be following these rules:
 - **Preprocess** the dataset by **tokenizing the words, remove stop words, remove symbols and number, stemming, and lemmatizing the words**.
 - Compare the review words with the words in the list of dictionaries.
 - If the review is **greater** than **5**, then set the **review category** to **positive**.
 - If the review is **equal to or less** than **5**, then set the **review category** to **negative**.
 - **Train** the model using **Naïve Bayes**.
 - **Save training model to pickle file** with format name “**model.pickle**”.
- The application menu will **show user reviews** and their respective **classifications** based on the **provided categories**, consisting of **four menus**. **Validate** in the menu, that user can only choose number in the range of menu provided (**1 – 4**).
 1. **Write Your Review**
 2. **View Movie Recommendation**
 3. **View Named Entity Recognition**
 4. **Exit**

```
MOVIE RECOMMENDATION APPLICATION BASED ON REVIEWS
YOUR REVIEW : NO REVIEW
YOUR REVIEW CATEGORY : UNKNOWN
1. WRITE YOUR REVIEW
2. VIEW MOVIE RECOMMENDATION
3. VIEW NAMED ENTITIES RECOGNITION
4. EXIT
>>
```

- If user choose **menu 1** (“**Write Your Review**”), then the application will:
 - Ask the user to input review. **Validate** that the input must **at least contains of 20 words**.
 - Afterward, the application will **save** the **review** and **display** the **review category** in the main menu.
- If user choose **menu 2** (“**View Movie Recommendation**”), then the application will:
 - **Check** whether there is a **review** or not.
 - If there is no **review exists**, then show a **message** to notify the user and redirect user **back to main menu**.
 - The application should **support one text representation or word embedding techniques** for feature extraction, allowing the examiner to choose from TF-IDF, Word2Vec, or language models such as unigram, bigram, or trigram. The selected technique should be applied during the recommendation process to **calculate cosine similarities** between user queries and document text.
 - **Calculate** the **cosine similarities** of the user queries with the documents that have been embedded using any text representation or word embedding techniques.
 - **Display** the **top 2 movies** from the **cosine similarity results** as movie recommendations.

```
TOP 2 MOVIE RECOMMENDATION FOR YOU:
1: I Am Legend
2: Annabelle Comes Home

[>] PRESS ENTER TO CONTINUE...
```

- If user choose **menu 3** (“**View Named Entity Recognition**”), then the application will:
 - Display **entity tags** such as **language** and **location**, along with the corresponding **entity items** present in the document.

CATEGORIZED NAMED ENTITIES:

LANGUAGE: English, French, Spanish

LOC: the Die Hard, Inferno, Weber, Hedwig, North, Jupiter Ascending, her;Don, Southwest, Heigl/Kutcher, Captain America, Dracula, West, Proyas, Watts, Black Sea, Brokeback Mountain, Baywatch, Affleck, the Upper West Side, Bay State, Hudson, the Sea, Hood, Earth, Agnes Kittelsen, Black Adam, Britney, the Doris Day-Rock, Atlantic, earth, Fraggie Rock, Pacific, The Wild Bunch, Nile, South, Worth, Crystal, Spring Break, Hyde Park, Valley, Beach Blanket, SESSION, Africa, Everest, Wild, Skull Island, Ellis Island, Melodramatic, Roman Holiday, Mars, Bay, Silicon Valley, Marina, Yakuza, The Golden Circle, Northern California, Bird, Middle America, world;Full, Fantasy Island, Northern Italy, Routine Dark Castle, the Eastern Bloc

[>] PRESS ENTER TO CONTINUE...

- If user choose **menu 4 (“Exit”)**, then **terminate** the application.