

Search-Movies

Search Engine for Movies

Problem Statement

Many of us face problems in remembering the names of the movies, but we always remember the basic plot and cast.

Solution?

This search engine helps you find the name of the movie and also recommends movies that have similar plot and genre by using concepts of **Information Retrieval**.

Functionalities

- Reads the .csv file of all movies with plots.
- Removes all the characters that are not letters or numbers, converts the sentence to lowercase and removes the stopwords occurring in the document.
- Removes all the documents that does not have plots and applies the necessary function to the documents that have plots.
- Tokenizes all the words in a document's plot and stores them in a list, then it stores all the lists of individual plots into a corpus.
- Creates a document vector by taking the mean of word vectors of the words in the document.
- Asks for the query from the user.
- Performs preprocessing(above mentioned operations) to the query.
- Displays top ten movie titles that matched with the query given and displays its corresponding document.

Tech Stack

Technologies to be used.

- NLTK(Natural Language Toolkit Library)
- Gensim Library(word2vec model)
- Numpy and Pandas libraries
- Regex Library

Team Members

- [Nikhil Munigela](#)
- [Jathin Badam](#)
- [Venkata Sai Karthik Jagini](#)
- [Nikhil Kandukuri](#)