Project Report: Movie Recommendation System

**Introduction:**

The project aimed to build a movie recommendation system using Natural Language Processing (NLP) techniques and machine learning algorithms. The system suggests movies to users based on the similarity of movie features such as genres, keywords, cast, and director.

**Technologies Used:**

* Python programming language
* Libraries: NumPy, Pandas, scikit-learn
* Algorithms: TF-IDF Vectorization, Cosine Similarity

Steps Taken:

1)Data Preparation:

* Loaded movie data from a CSV file (movies.csv) using Pandas.
* Explored the dataset to understand its structure and contents.

2)Feature Selection and Preprocessing:

* Selected relevant features for similarity calculation: genres, keywords, tagline, cast, and director.
* Handled missing values by replacing them with empty strings.

3)Feature Combination:

* Combined selected features into a single feature vector for each movie using string concatenation.

4)Vectorization:

* Utilized TF-IDF Vectorization to convert text data into numerical feature vectors.

5)Similarity Calculation:

* Calculated cosine similarity between feature vectors to measure the similarity between movies.

6)User Interaction:

* Enabled user input to specify their favorite movie.
* Utilized the difflib library to find close matches to the input movie name in the dataset.

7)Recommendation Generation:

* Based on the selected movie, retrieved the similarity scores of all movies.
* Sorted the movies based on similarity scores in descending order to recommend the most similar movies.
* Displayed the top recommended movies to the user.

**Results:**

The movie recommendation system successfully recommends movies based on user input. It provides a list of top similar movies, allowing users to discover new movies based on their preferences.