

INTRODUCTION

- The Movie Recommendation System is a content-based filtering system designed to suggest movies based on user preferences (mostly **Hollywood**). Leveraging the power of machine learning, the system analyses various features of movies, including genres, keywords, cast, and director information. The dataset used for this project encompasses a diverse collection of movies, each described by a combination of these features.
- To enhance the user experience, the system employs **Term Frequency-Inverse Document Frequency (TF-IDF)** vectorization and **cosine similarity**. TF-IDF helps convert textual information into feature vectors, while cosine similarity quantifies the similarity between movies based on these vectors. Users can input the name of their favourite movie, and the system intelligently finds the closest match in the dataset. Subsequently, it suggests a list of movies that share similarities with the user's input, providing a personalized and tailored recommendation experience.

DESCRIPTION

1. Data Collection and Preprocessing

- The project starts by loading a dataset (movies_data.csv) containing information about various movies. This dataset includes details such as genres, keywords, original titles, overviews, taglines, cast, and directors and much more.
- Selected features which are used for recommending movies are genres, keywords, cast, directors and many more.

2. Feature Vectorization (Feature Extraction)

- Textual data from the selected features is processed using the Term Frequency-Inverse Document Frequency (TF-IDF) vectorizer. This step converts the textual information into numerical feature vectors.

3. Cosine Similarity Calculation

- Cosine similarity is employed to measure the similarity between the feature vectors of different movies. This metric helps identify movies with similar content characteristics.

4. User Interaction

- Users are prompted to input the name of their favourite movie. The system employs the difflib library to find the closest match to the user's input within the dataset.

5. Recommendation movie

- Once the user's movie is identified, the system compares its features with other movies in the dataset. Movies with higher cosine similarity scores are recommended to the user.
- The system displays a list of recommended movies to the user. These suggestions are tailored to the user's input, ensuring a personalized movie-watching experience.

6. Handling Exceptions

- If the system cannot find a close match to the user's input, a prompt is displayed, indicating that it is not possible to recommend a movie based on the provided information. Users are encouraged to try again with another movie.

RESULTS

- The results of the Movie Recommendation System would be in the form of a list of recommended movies based on the user's input.
- The recommended movies for the favourite movie: **The Shawshank Redemption** is (successful)

```
... The suggested movie for you are listed below

1 -- The Shawshank Redemption
2 -- Prison
3 -- Civil Brand
4 -- Standard Operating Procedure
5 -- Fortress
6 -- Double Jeopardy
7 -- Escape Plan
8 -- The Longest Yard
9 -- Penitentiary
10 -- Get Hard
11 -- The Naked Gun 33½: The Final Insult
12 -- Let's Go to Prison
13 -- Lucky Break
14 -- Goodbye Bafana
15 -- American History X
16 -- Bound
17 -- 25th Hour
18 -- 102 Dalmatians
19 -- Alien³
20 -- Air Force One
21 -- Hurricane Streets
22 -- Bronson
23 -- Kiss of Death
...
27 -- Out of the Furnace
28 -- Buffalo '66
29 -- White Oleander
30 -- Mean Machine

Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
```

- The recommended movies for the favourite movie: **3 idiots** is (unsuccessful)

```
Sorry not possible to recommend any of the movie

You can try again with another movie
```

CONCLUSION

- The Movie Recommendation System stands as a valuable tool for movie enthusiasts seeking tailored suggestions. Its implementation of content-based filtering and cosine similarity contributes to a more enjoyable and personalized movie-watching journey. As technology and datasets evolve, the system can be further refined to provide even more accurate and diverse recommendations, making it a dynamic and evolving solution for movie enthusiasts.

REFERENCES

- Dataset: <https://www.kaggle.com>
- <https://www.youtube.com/@Siddhardhan>