WiDS '22 - '23 Final Documentation

<25 - Movie Recommendation System>

<Mentor: Amar>

Team Member Name	Roll Number	Email-Id
Dheeraj Kumar Solanki	210020041	dheerajsolanki203@gmail.com

Introduction to Problem Statement

To Develop a movie recommendation from raw data of 5000+ movies.

Existing Resources

https://www.kaggle.com/datasets/tmdb/tmdb-movie-metadata?select=tmdb_5000_movies.csv

Proposed Solution

To recommend other movies related to the given movie we'll have to find similarity between these movies. As we don't have numerical data so I vectorized the data based on some frequently used words in the data. Then I used cosine similarity to find 10 nearest or most similar movies to our movie using cosine distance.

Methodology & Progress (Mention the work done week-wise)

Week-1	Basics, Overview, Python Basics
Week-2	Previewing Data, Pre-Processing, EDA
Week-3	NLP Basics, NLTK
Week-4	Basic ML models, feature engineering
Week-5	Report and Code Submission

Results

https://github.com/Dheeraj-2003/Movie-Recommendation-System-Wids

Learning Value

Learnt about several libraries of python including numpy, pandas, nltk,ast
And concepts of natural language processing, cosine similarities.

Tech-stack Used

Python libraries
-nltk
-scikit-learn
-numpy
-pandas

References and Citations

https://www.kaggle.com/datasets/tmdb/tmdb-movie-metadata?select=tmdb_5000_movies.csv