MOVIE PREDICTION

AND SENTIMENT ANALYSIS

BY NIKITA

TMDB MOVIE DATASET

INTRODUCTION : This project develops a Movie Recommendation System enhanced with Sentiment Analysis using real movie metadata. It performs data loading, cleansing, feature extraction, exploratory analysis, visualization, and similarity-based movie recommendations

KEY POINTS :Movie metadata processing , Handling missing & nested JSON fields ,NLP-based text handling , Similarity-driven recommendations , EDA & visual insights

TOOLS AND TECHNOLOGIES : LANGUAGE Python , LIBRARIES numpy, pandas, matplotlib, seaborn, ast , TOOLS：Google Colab , Kaggle API

PROCESS :   **Data Collection:** Download via Kaggle API and load into Pandas.
·**Cleaning & Preprocessing:** Remove nulls, convert datatypes, extract fields from nested data.
·**Feature Engineering:** Process text fields (genres, keywords, etc.) for similarity computation.
·**EDA & Visualization:** Revenue vs Budget, most common genres, distribution of ratings, popularity patterns.
·**Recommendation Logic:** Generate recommendations using metadata similarity.
·**Output:** Recommended movie list and analytical observation Predictive Summary :C:\Users\NIKITA\Downloads\PREDICTIVE MOVIE SUMMARY.docx
·**Python Notebook** : C:\Users\NIKITA\Downloads\Untitled9 (1).ipynb

SENTIMENT ANALYSIS : ACCURACY 91 % PRECISION 89% RECALL 90% F1 SCORE 89%

CONCLUSION :   This project demonstrates that well-processed movie metadata can effectively power a content-based recommendation system without requiring historical user behavior. The addition of NLP sentiment scoring improves contextual interpretation of content reception. With predictive scores demonstrating high accuracy and strong similarity performance, this framework provides a scalable foundation for future deployment as a web application or integrated ML recommendation pipeline.