Development of a Movie Recommendation System Using SVD and KNN Algorithms

1. Executive Summary:

This project focuses on developing a movie recommendation system utilizing Singular Value Decomposition (SVD) and K-Nearest Neighbors (KNN) algorithms. By leveraging these algorithms, the system aims to provide personalized movie recommendations to users based on their past viewing preferences and ratings. The integration of SVD and KNN ensures that the system captures both the latent factors in user-item interactions and the similarity between users or items, leading to accurate and reliable recommendations.

2. Problem Statement:

Background

In the era of digital streaming, users face an overwhelming choice of movies, making it difficult to find films that match their preferences. Traditional search methods fall short in providing personalized recommendations, leading to user dissatisfaction and decreased engagement.

Objective

The goal is to develop a movie recommendation system using Singular Value Decomposition (SVD) and K-Nearest Neighbors (KNN) algorithms. This system will predict user preferences and offer personalized movie suggestions to enhance user experience and increase engagement.

Scope

The project involves collecting and preprocessing movie rating data, implementing SVD and KNN algorithms, and developing a hybrid model. The system will be evaluated using metrics like RMSE and precision, and it will address challenges such as data sparsity and the cold start problem.

.

3. Data Sources:

Kaggle

Primary Data: Movie names

Secondary Data: Movie Rating

4. Methodology:

- **Data Collection and Preprocessing**:
    - Gather movie rating data from a reliable
    - Clean the dataset by handling missing values and normalizing ratings.
- **Model Implementation**:
    - **SVD (Singular Value Decomposition)**:
        - Decompose the user-item rating matrix into three matrices: U (user features), Σ (singular values), and V^T (item features).
        - Use these decomposed matrices to predict missing ratings and recommend movies.
    - **KNN (K-Nearest Neighbors)**:
        - Calculate the similarity between users or items using metrics like cosine similarity or Pearson correlation.
        - Identify the K-nearest neighbors for a given user or item and recommend movies based on their ratings.
- **Hybrid Approach**:
    - Combine SVD and KNN recommendations to leverage the strengths of both algorithms.
    - Implement a weighted averaging or a hybrid ensemble method to generate the final recommendation list.
- **Evaluation**:
    - Split the dataset into training and testing sets.
    - Evaluate the recommendation system using metrics like Root Mean Squared Error (RMSE) for predicted ratings and precision, recall, and F1-score for recommendation accuracy.

5. Expected Outcomes:

Improved Recommendation Accuracy: By using SVD to capture latent factors and KNN to identify similar users/items, the system should provide highly accurate movie recommendations.
Enhanced User Experience: Users will receive personalized movie suggestions that align with their preferences, leading to increased satisfaction and engagement with the platform.

Scalability: The hybrid model will be designed to efficiently handle large datasets, ensuring scalability for platforms with extensive movie libraries and user bases.

6. Risks and Challenges:

Data Sparsity: Sparse user-item matrices can affect the performance of SVD and KNN, potentially leading to less accurate recommendations.
Scalability Issues: KNN's computational complexity might become a bottleneck with a growing number of users and movies.
Cold Start Problem: Recommending movies for new users with no or minimal rating history remains a challenge.

7. Conclusion:

this project demonstrates the potential of combining SVD and KNN algorithms to create a robust movie recommendation system. By addressing both the latent factors in user preferences and the similarity between users or items, the system aims to deliver highly personalized and accurate movie recommendations. Despite potential challenges such as data sparsity and the cold start problem, the hybrid approach provides a promising solution to enhance user experience in the realm of digital content consumption. Future work may involve integrating additional data sources, such as user demographics and movie metadata, to further improve recommendation accuracy and address existing limitations.