**Collaborative-Based Recommendation System**

**Introduction**

The goal of this project is to design a collaborative-based recommendation system for suggesting personalized posts to users. Using **user-item interaction data** such as views, likes, ratings, and inspirations, we build a recommendation system based on collaborative filtering techniques.

The key algorithm used is **Singular Value Decomposition (SVD)**, which excels in accuracy as measured by **Root Mean Square Error (RMSE)** and **Mean Absolute Error (MAE)**. Additionally, the system is evaluated against other algorithms like **KNNBasic** and **NMF** for comparison.

**Step-by-Step Approach**

**1. Data Preparation**

* **Dataset**:
  + viewed\_posts.csv: Records of posts viewed by users.
  + liked\_posts.csv: Posts liked by users.
  + inspired\_posts.csv: Posts that inspired users.
  + rated\_posts.csv: User ratings for posts.
  + all\_posts.csv: Metadata about posts (e.g., title, category).
  + all\_users.csv: Metadata about users.
* **Data Cleaning**:
  + Convert timestamps to datetime format (viewed\_at, liked\_at, rated\_at, inspired\_at).
  + Merge interaction data with all\_posts for enriching with metadata.
  + Handle missing values:
    - Fill missing ratings with 0.
    - Replace missing average ratings with the column mean.
* **Feature Engineering**:
  + Normalize ratings to a range of [0, 1] for consistency in collaborative filtering.
  + Combine text features for posts (e.g., category, title) for potential hybrid extensions.

**2. Model Development**

* **Collaborative Filtering Approach**:
  + The user-item interaction matrix is created using user\_id, post\_id, and normalized\_rating.
* **Algorithms Implemented**:
  + **SVD**:
    - A matrix factorization technique that reduces dimensionality and predicts user ratings for items.
  + **KNNBasic**:
    - Finds the k-nearest neighbors for items or users based on similarity.
  + **NMF**:
    - A matrix factorization method constrained to non-negative values, making it interpretable.
* **Hyperparameter Tuning**:
  + **Grid Search** was performed on n\_factors, n\_epochs, lr\_all, and reg\_all to optimize SVD performance.
  + Best parameters:
    - RMSE: {'n\_factors': 50, 'n\_epochs': 20, 'lr\_all': 0.005, 'reg\_all': 0.4}
    - MAE: {'n\_factors': 50, 'n\_epochs': 20, 'lr\_all': 0.01, 'reg\_all': 0.2}

**3. Evaluation Metrics**

* **Root Mean Square Error (RMSE)**:
  + Measures accuracy with sensitivity to large errors.
  + Best RMSE: **0.1637** (SVD).
* **Mean Absolute Error (MAE)**:
  + Measures the average magnitude of prediction errors.
  + Best MAE: **0.0884** (NMF).
* **Insights**:
  + SVD strikes a balance between RMSE and MAE, making it suitable for systems requiring precise predictions.
  + NMF minimizes average errors (MAE) but has slightly higher RMSE.
  + KNNBasic performed less effectively due to high data sparsity.

**5. Results and Insights**

* SVD consistently outperformed KNNBasic and NMF in RMSE while remaining competitive in MAE.
* The system was successfully tested and validated using cross-validation on five splits.
* Recommendations are tailored to user interactions, showing meaningful results.

**6. Conclusion**

* The collaborative-based recommendation system, powered by SVD, provides accurate and efficient personalized recommendations.
* Future improvements include exploring hybrid models by integrating content-based filtering with collaborative filtering.