Report: Music Recommendation System

1. **Stakeholder:**
The stakeholder for this project is the product development team of a music streaming service.

2. **Problem Statement**:
To enhance user experience and engagement by implementing a music recommendation system that provides personalized song suggestions to users based on their preferences.

3. **Dataset Source**:
The dataset used for this project is the Spotify Million Song Dataset available on Kaggle. (https://www.kaggle.com/datasets/notshrirang/spotify-million-song-dataset)

4. **Models Explored**:

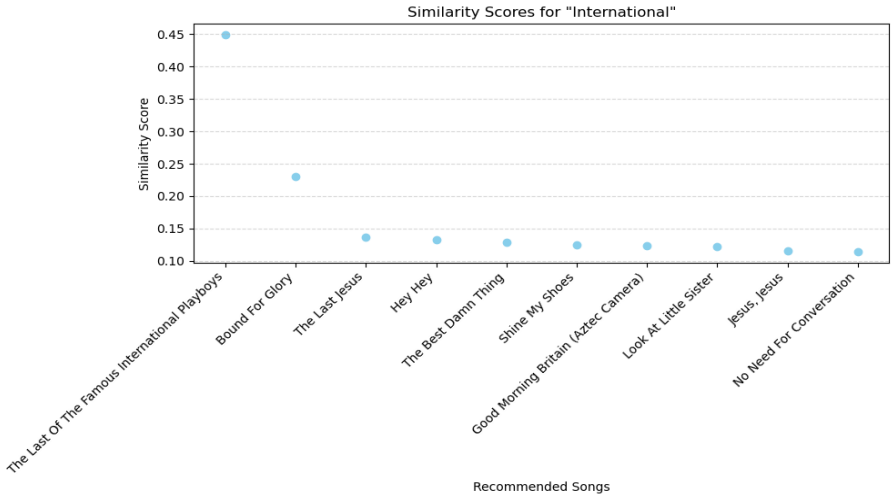
a. **Content-Based Filtering**:
- Model: TF-IDF Vectorization with Cosine Similarity
- Reasoning: Content-based filtering leverages textual features such as song lyrics to recommend songs that are textually similar to the user's preferences. TF-IDF (Term Frequency-Inverse Document Frequency) vectorization is chosen to represent textual features, and cosine similarity is used to measure the similarity between songs.

b. **Collaborative Filtering:**
- Model: Matrix Factorization (SVD - Singular Value Decomposition)
- Reasoning: Collaborative filtering recommends items based on the preferences of users with similar tastes. Matrix factorization techniques such as SVD are effective in capturing latent features and generating recommendations.

5**. Feature Selection/Engineering**:
- The features selected for the content-based filtering model include artist names, song titles, and song lyrics.
- These features are combined into a single feature called "Combined_features" to represent the textual information of each song.

6. **Model Evaluation**:
- For the content-based filtering model, cosine similarity between songs is used as the evaluation metric. Cosine similarity measures the similarity between two non-zero vectors of an inner product space, which in this case, represents the similarity between song vectors.
- For the collaborative filtering model, the Root Mean Square Error (RMSE) or Mean Absolute Error (MAE) could be used as evaluation metrics to measure the accuracy of predicted ratings compared to actual ratings in the test set.

7. A visual representation of the similarity scores between the input song and the recommended songs, helping users to identify similar songs.



8. **Future Work**:
- Experiment with different text preprocessing techniques to further improve the quality of textual features.
- Explore hybrid recommendation approaches that combine content-based and collaborative filtering methods for improved recommendation accuracy.
- Conduct A/B testing to assess the real-world impact of the recommendation system on user engagement and satisfaction.

9. **Recommendation:**
- While the models show promise in providing personalized song recommendations, further evaluation and testing are necessary before deploying the system in production. Consideration should be given to the specific requirements and performance thresholds set by stakeholders.
- The precision metrics should be evaluated against the intended use case to ensure the recommendations meet user expectations.

Conclusion:

This report provides insights into the methodology, models explored, feature engineering, model evaluation, and recommendations for future work of a music recommendation system. The goal is to enhance user experience and engagement on the music streaming platform by providing personalized song recommendations tailored to individual preferences.