**Character Similarity Model Evaluation**

This project implements a character similarity recommendation system that identifies similar characters across different movies based on their dialogue patterns. The system uses natural language processing techniques to analyze character dialogue and generate personalized recommendations.

**Core Model Components**

* **Sentence Transformer Model**
  + Model Name: all-MiniLM-L6-v2
  + Architecture: A distilled version of BERT optimized for generating sentence embeddings
  + Dimensions: 384-dimensional dense vector embeddings
  + Training: Pre-trained on a combination of over 1 billion sentence pairs
  + Purpose: Converts character dialogue into meaningful embeddings
* **Text Preprocessing Pipeline**
  + Lowercase Conversion: Normalizes all text to lowercase
  + Whitespace Normalization: Replaces multiple spaces with single spaces
  + Text Cleaning: Removes extraneous characters and standardizes format
* **Similarity Computation**
  + Method: Cosine similarity between character dialogue embeddings
  + Threshold: 0.5 (determined through distribution analysis)
  + Cross-Movie Filter: Excludes recommendations from the same movie

**Performance Metrics**

* Accuracy: 0.62 (62% of predictions are correct)
* Precision: 0.60 (60% of recommended character pairs are genuinely similar)
* Recall: 1.0 (100% of truly similar character pairs are captured)
* F1 Score: 0.75 (balanced measure of precision and recall)

**Analysis and Insights**

During initial model testing, a threshold of 0.7 was used for the Cosine similarity. The resulting performance metrics were:

* Accuracy: 0.67 (67% of predictions are correct)
* Precision: 0.66 (66% of recommended character pairs are genuinely similar)
* Recall: 1.0 (100% of truly similar character pairs are captured)
* F1 Score: 0.79 (balanced measure of precision and recall)

A Distribution of Cosine Similarity Scores was generated to determine the appropriate threshold level and potential impact on performance.

A graph of a distribution of cosine

AI-generated content may be incorrect.

The slight decrease in precision and F1 score is a reasonable trade-off for achieving the project goal of cross-movie recommendations based on character similarity. The system maintains perfect recall while still offering good precision, making it effective for helping users discover new movies based on character preferences.

This approach creates a unique recommendation experience that differentiates the recommendation platform from others that might only recommend based on genre, director, or overall movie similarity.

**Error Analysis**

The error analysis identified cases where there is a mismatch between y\_true (actual) and y\_pred (predicted). The total error cases were 8,152 with a total sample size of 21,679. The mismatches occurred where the similarity score was close to the 0.5 threshold level.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| source\_movieID | source\_movieTitle | source\_characterName | target\_movieTitle | similarity | y\_true | y\_pred |
| m517 | Sleepy Hollow | Old Crone | The Princess Bride | 0.495839 | 0 | 1 |
| m517 | Sleepy Hollow | Old Crone | The Third Man | 0.492547 | 0 | 1 |
| m197 | Star Trek VI | Martia’s Voice | His Gal Friday | 0.478570 | 0 | 1 |
| m189 | Spider-Man | Roz | Star Trek Generations | 0.455963 | 0 | 1 |
| m189 | Spider-Man | Roz | Power Rangers | 0.453759 | 0 | 1 |

**Strategies for Error Minimization**

* Model Refinement
  + Threshold Tuning: Establish a tiered system within the model instead of a fixed 0.5 threshold for Cosine similarity
  + Model Fine-Tuning: Fine-tune the Sentence Transformer on a domain specific dataset where embeddings more accurately represent the nuances in character dialogue
* Data Quality Improvement
  + Data Augmentation: Increase the dataset with additional dialogue samples from the movie scripts. The current dataset is using selected lines and conversations which are a subset of the entire script.
  + Enhanced Preprocessing: Refine text cleaning by handling slang and context specific phrases to ensure embeddings effectively capture dialogue nuances.

**Recommendations for Implementation**

* Consider a Tiered Recommendation
  + “Strong Character Matches” with similarity > 0.65
  + “Similar Characters You Might Like” with similarity >= 0.5 and <= 0.65
* Add Contextual Information
  + Include brief character descriptions
  + Include character traits
  + Include memorable character quotes
* Feedback Mechanism
  + Thumbs Up/Down for recommendations