NAME: COVENANT OLUDAPOMOLA OJO

STUDENT ID: 23013682

COURSE: APPLIED DATA SCIENCE 2

EVALUATION OF MULTIMODAL IMDB GENRE CLASSIFICATION

1. INTRODUCTION

This project utilized CNN and LSTM models to classify movie genres based on multimodal data from the IMDB dataset. The CNN analyzed film posters, while the LSTM processed overviews. This report evaluates both models' performances, explores key observations, and compares their capabilities using specific examples.

MODEL PERFORMANCE AND RESULTS

CNN MODEL

METRICS AND TRENDS:

The CNN achieved consistent improvements in training precision and recall. However, the validation metrics slightly lagged, indicating potential overfitting.

Loss and accuracy plots revealed steady progress, though diminishing returns were observed after the 30th epoch.

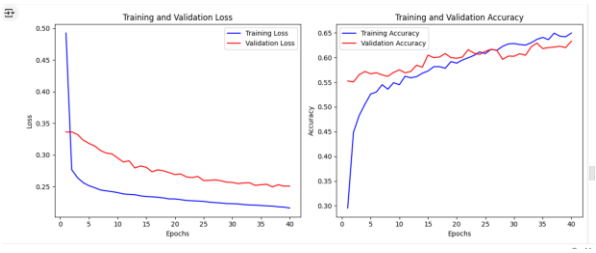


Figure 1: Training and validation (loss and accuracy)

CLASSIFICATION ANALYSIS:

Correct Samples: Posters with clear visual genre indicators (e.g., distinct action scenes or romantic imagery) were accurately classified.

Incorrect Samples: Ambiguous visuals or overlapping genres (e.g., drama and thriller) led to misclassifications.

Example Evaluation:

Film X: Accurately classified as Animation, Family, and Adventure. The colorful and animated visual cues were clear indicators for these genres.

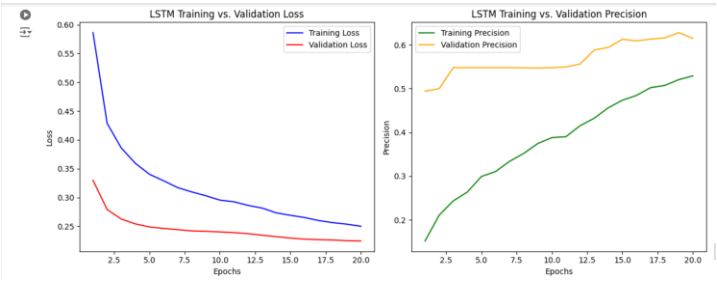
Film Y: Misclassified as Comedy and Action, while the actual genres were Romance and Drama. The poster's neutral tone lacked definitive clues, causing the misclassification.

LSTM MODEL

METRICS AND TRENDS:

The LSTM demonstrated higher validation precision and recall compared to the CNN. Training and validation metrics aligned closely, reflecting robust generalization.

Loss plots indicated steady progress without significant overfitting.



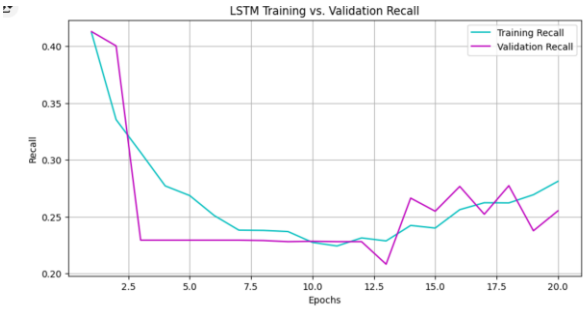


Figure 2: LSTM evaluation

Strengths and Limitations:

Captured nuanced text patterns effectively, leading to accurate genre classification.

Struggled with vague or inconsistent overviews, resulting in occasional misclassifications.

Example Evaluation:

Film C: Correctly classified as Fantasy, Adventure, and Family.

Film D: Misclassified as Historical Drama; actual genres were Biography and War.

COMPARISON BETWEEN CNN AND LSTM

PERFORMANCE:

LSTM consistently outperformed CNN due to its ability to leverage detailed textual data.

Strengths:

CNN excelled with visually distinct genres, making it ideal for visual-based classification tasks.

LSTM provided richer insights from textual descriptions, demonstrating higher flexibility and accuracy.

Weaknesses:

CNN struggled with subtle or ambiguous poster designs.

LSTM's performance depended on the quality of the provided text.

CRITICAL OBSERVATIONS

For three selected films, the following comparisons were made:

Film ID: tt0117924

Actual Genres: Action, Sci-Fi, Drama

CNN Prediction: Action, Drama, Romance

LSTM Prediction: Action, Sci-Fi, Drama

Film ID: tt0829193

Actual Genres: Comedy, Romance

CNN Prediction: Action, Comedy, Drama

LSTM Prediction: Comedy, Romance, Drama

Film ID: tt0878835

Actual Genres: Biography, War

CNN Prediction: Drama, Action, History

LSTM Prediction: Biography, War, Drama

Analysis:

CNN’s predictions were accurate for visually strong genres but diverged for subtle overlaps.

LSTM consistently aligned closer to ground truth due to its semantic understanding.

CONCLUSION

The analysis highlights the complementary strengths of CNN and LSTM models. While CNN thrives on visual clarity, LSTM excels in capturing complex textual nuances. Future research could explore ensemble approaches, combining both modalities for enhanced genre classification accuracy.