IMDB Movie Review Sentiment Analysis Using RNN

Objective

This project focuses on sentiment classification of movie reviews from the IMDB dataset using a Recurrent Neural Network (RNN). The model predicts whether a review expresses a positive or negative sentiment.

Dataset Overview

- **The IMDB dataset consists of 50,000 movie reviews, labeled as positive (1) or negative (0).**
- **The top 10,000 most frequent words are retained.**
- **Each review is padded to a fixed length of 500 words.**

Model Architecture

- **Embedding Layer: Converts words into 32-dimensional dense vectors.**
- **SimpleRNN Layer: Processes sequential data to learn relationships between words.**
- **Dense Output Layer: Uses sigmoid activation for binary classification (Positive/Negative).**

Training & Evaluation

- **Optimizer: Adam**
- **Loss Function: Binary Crossentropy**
- **Batch Size: 128**
- **Epochs: 10**
- **Validation Split: 20% of training data**
- **Test Accuracy: Achieved an accuracy of ~85% on unseen test data.**

Visualization & Predictions

- **Accuracy Trends: A graph is plotted to show training vs validation accuracy over epochs.**
- **Custom Review Predictions: Users can input their own movie reviews to get sentiment predictions with confidence scores.**

Future Enhancements

- **Use LSTMs or GRUs for improved handling of long-term dependencies.**
- **Implement pre-trained embeddings (e.g., GloVe, Word2Vec).**
- **Apply Transformer models like BERT for better sentiment analysis.**