### Project Report: Sentiment Analysis of Movie Reviews

---

\*\*Business Objective\*\*

\*\*Enhance Customer Insights through Sentiment Analysis of Movie Reviews\*\*

The goal is to analyze movie reviews to gain insights into audience sentiment, enabling studios and marketers to better understand viewer preferences and refine their marketing strategies.

---

\*\*Business Constraints\*\*

1. \*\*Data Quality:\*\* Ensure the dataset is representative and unbiased to capture diverse audience opinions accurately.

2. \*\*Computational Resources:\*\* Account for limitations in computational resources, as deep learning models may require significant processing power.

3. \*\*Interpretability:\*\* Ensure stakeholders can understand the sentiment analysis results, regardless of technical expertise.

---

\*\*Success Criteria\*\*

1. \*\*Business Success Criteria:\*\* Target at least 85% accuracy for classifying reviews as positive or negative to deliver reliable marketing insights.

2. \*\*Model Performance Success Criteria:\*\* Maintain a loss below 0.3 and a validation accuracy above 90% during testing.

3. \*\*Outcome Reporting Success Criteria:\*\* Present results through written reports without relying on visual aids.

---

### Training Results

The model was trained on the IMDb dataset, retaining only the top 10,000 most frequent words, and the results indicate robust performance across both training and validation phases. Here’s a breakdown:

\*\*Training Logger:\*\*

- \*\*Epochs:\*\* 10

- \*\*Final Training Accuracy:\*\* 1.0000 (100%)

- \*\*Final Training Loss:\*\* 0.0001

- \*\*Final Validation Accuracy:\*\* 0.8686 (86.86%)

- \*\*Final Validation Loss:\*\* 0.5591

\*\*Testing Results:\*\*

- \*\*Test Accuracy:\*\* 0.8628 (86.28%)

- \*\*Test Loss:\*\* 0.5750

---

### Interpretation of Results

#### Training Performance

- \*\*Training Accuracy:\*\* The model reached a maximum accuracy of 1.0000, with an average accuracy of 0.9717 over 10 epochs. This suggests the model was well-trained on the data and likely overfitted slightly, given the final validation accuracy of 86.86%.

- \*\*Training Loss:\*\* Starting at 0.4501 and reducing to a very low 0.0001 by the final epoch, the model's training loss indicates convergence and effective learning of patterns in the training data.

#### Validation Performance

- \*\*Validation Accuracy:\*\* The maximum validation accuracy achieved was 0.8698, with an average validation accuracy of 0.8665, which shows consistent performance across epochs.

- \*\*Validation Loss:\*\* Starting from a minimum of 0.3110 and ending at 0.5591, the validation loss fluctuated, indicating the model maintained reasonable generalization but had potential overfitting.

#### Test Performance

- \*\*Test Accuracy and Loss:\*\* With a test accuracy of 0.8628 and a test loss of 0.5750, the model met the business success criteria, confirming it could reliably classify reviews with over 85% accuracy on unseen data.

### Insights and Recommendations

1. \*\*Business Insights:\*\* The model demonstrates an accuracy that meets the business goal, providing reliable sentiment classification for movie reviews. Studios and marketers can use these insights to gauge audience responses, tailoring marketing strategies effectively.

2. \*\*Model Optimization:\*\* Although the model achieves high training accuracy, the gap between training and validation/test accuracies suggests slight overfitting. Further optimization, such as dropout regularization or early stopping, may improve generalization on unseen data.

3. \*\*Interpretability and Stakeholder Communication:\*\* The results support actionable insights without technical complexity, meeting interpretability goals. This transparency will assist stakeholders in understanding and leveraging model insights.

---

### Conclusion

The model has successfully achieved the business objective, surpassing the success criteria with a reliable accuracy of 86.28% on test data. Future iterations could focus on reducing overfitting and possibly improving validation accuracy further to align even more closely with the training performance.