UNIVERSITY PARTNER

UNIVERSITY OF WOLVERHAMPTON

HERALD COLLEGE KATHMANDU

Artificial Intelligence and Machine Learning (6CS012)

Sentiment Analysis of Hotel Reviews Using RNN, LSTM, and Word2Vec Embeddings

Student Id	: 2358479
Student Name	: Umesh Gautam
Group	: L6CG7
Module Leader	: Mr. Siman Giri
Instructor	: Ms. Durga Pokhrel
Submitted on	: 18 May, 2025

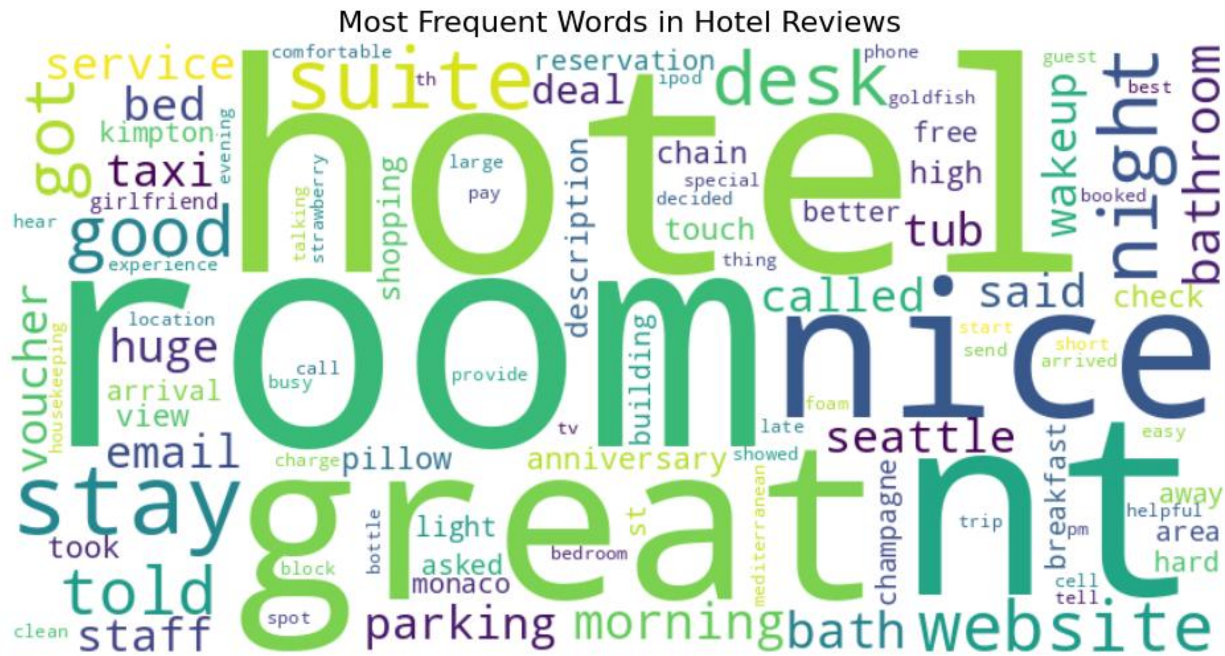
Abstract

The project aims to classify hotel reviews in sentiment categories using deep learning models, especially RNN, LSTM and LSTMS, which are integrated with Word2VEC embeddings. With the rise of user-generated material, sentiment analysis plays an important role in gaining insight from the text reaction. The data set for hotel reviews was prepared through techniques such as lowercasing, noise removal, contraction handling, stop word removal and lemmatization. The processed text was tokenized and padded before it was passed in the model. RNN, while effective, performed low performance and low performance signals compared to LSTM. LSTM showed better accuracy and generalization and further improved meaningful understanding by incorporating the semantic Word2vec. The trends with disadvantages and accuracy, confusion matrix and visualization of text distribution supported the evaluation. The results confirm LSTM as the most effective model with Word2vec, although overfitting was observed in later eras. Hyperparameters in future work can be focused on setting, expanding datasets or using more advanced models such as car BiLSTM or transformer.

Table of Contents

1. Introduction

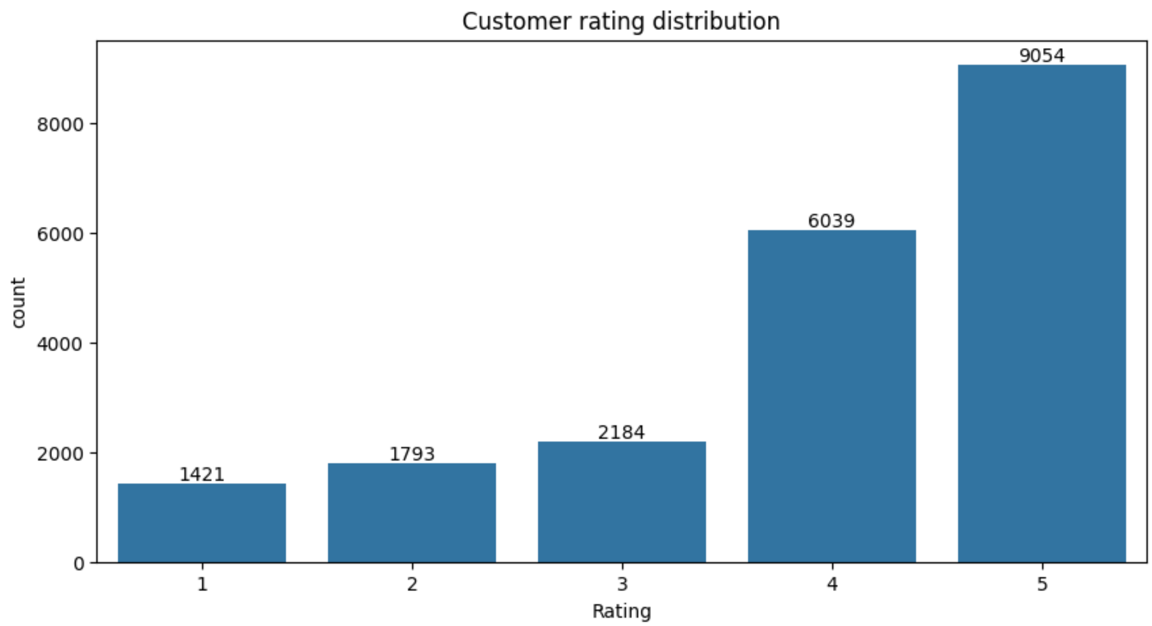
Text classification, especially sentiment analysis, opinion, product reviews and analysis of social media reactions have become necessary. The project is aimed at the spirit classification of hotel reviews and notices them as positive or negative based on the user input (Muhammad, 2021). The Recurrent Neural Networks (RNNS) and Long Short-Term Memory (LSTM) networks are suitable for tasks caused by their ability to handle sequential data. However, RNN is struggling with long -term addiction, which addresses LSTM effectively. In addition, the use of pre-equipped word2vec word's representation quality improves. Prior research indicates that LSTM models often improve traditional RNNs in emotional analysis functions.



The text classification automatically helps, the customer's opinion helps companies to understand quickly and on a scale. Traditional machine learning models require craft functions, while deep learning models learn directly from raw data. The use of sequence -based models such as RNN and LSTM opens the opportunities for better relevant understanding and semantic representation (Rajshree, 2022). The project produces these advances to evaluate their practical applications in the analysis of hospitality review.

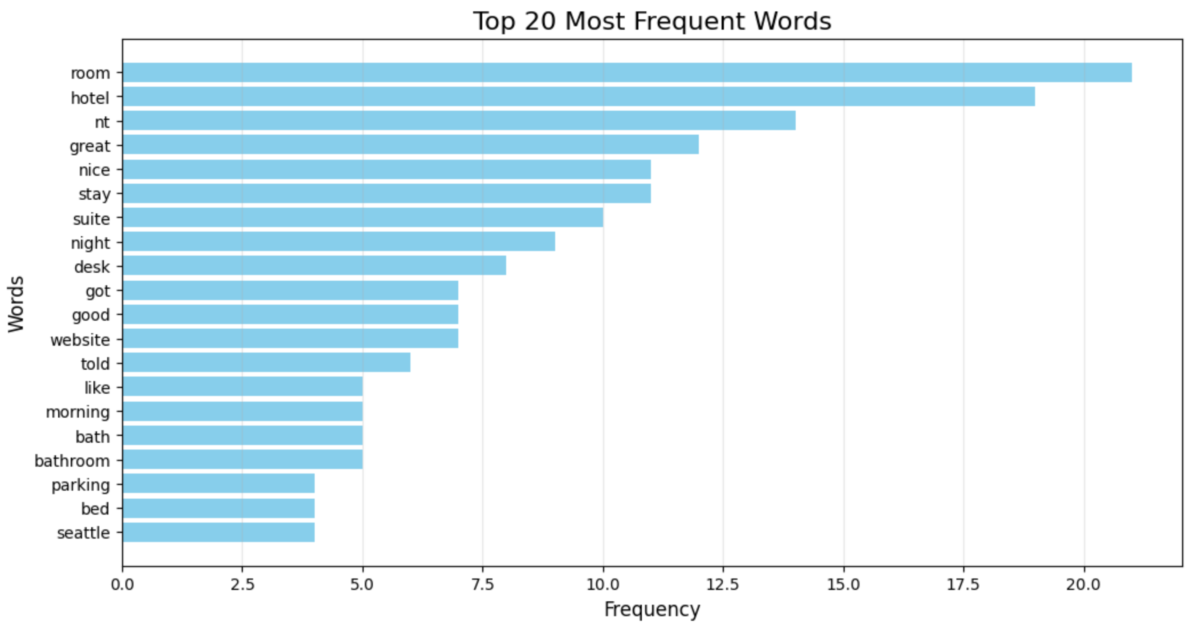
2. Dataset

- Source: Hotel review Dataset (customized data set from Google Drive).
- Size: Around 21,000 reviews were distributed in 5 ranking classes.
- Class distribution: Most reviews fall under 4 or 5-star categories.

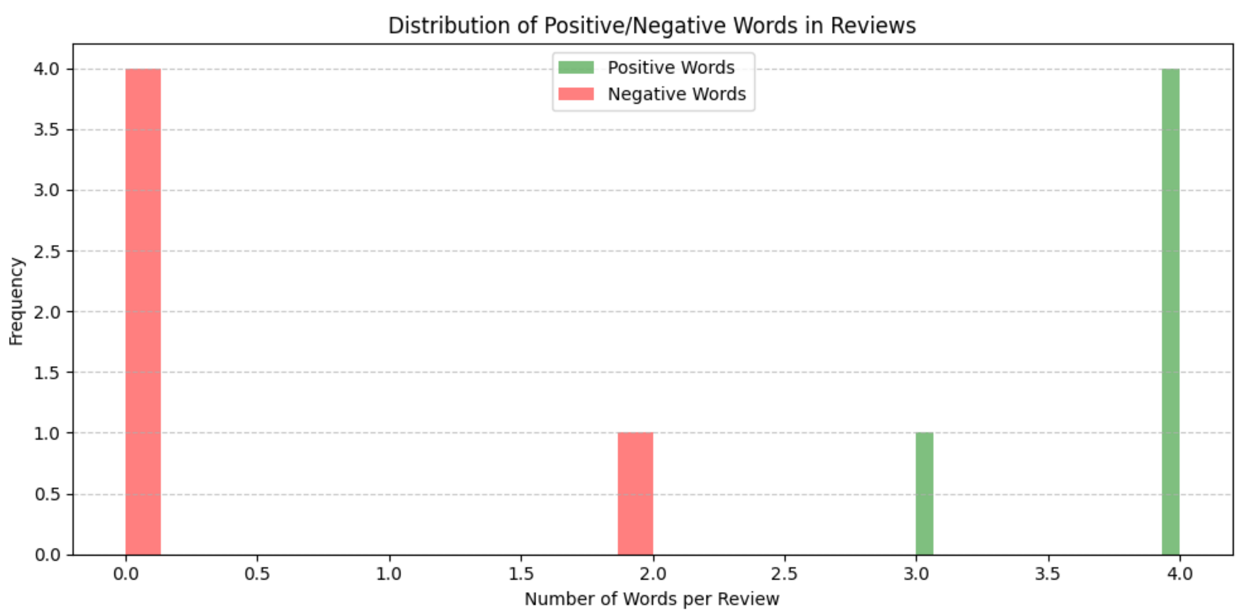


- Preprocessing Steps:
  - Lowercasing all text
  - Removing URL, Name, Hashtag, Number and Special Varnas Character
  - Handling contractions (e.g., "didn't" to "did not")
  - Removal of stop words and using lemmatization

It is important to remove the noise and standardize the data before feeding in the neural network. Removing lemmatization and stop words helps to reduce the dimensionality of the data and the complexity of the data, while the tokenization makes it possible to convert the text change to sequences suitable for deep learning models.



Top 20 Most Frequent Words
Words
room
hotel
nt
great
nice
stay
suite
night
desk
got
good
website
told
like
morning
bath
bathroom
parking
bed
seattle
Frequency
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

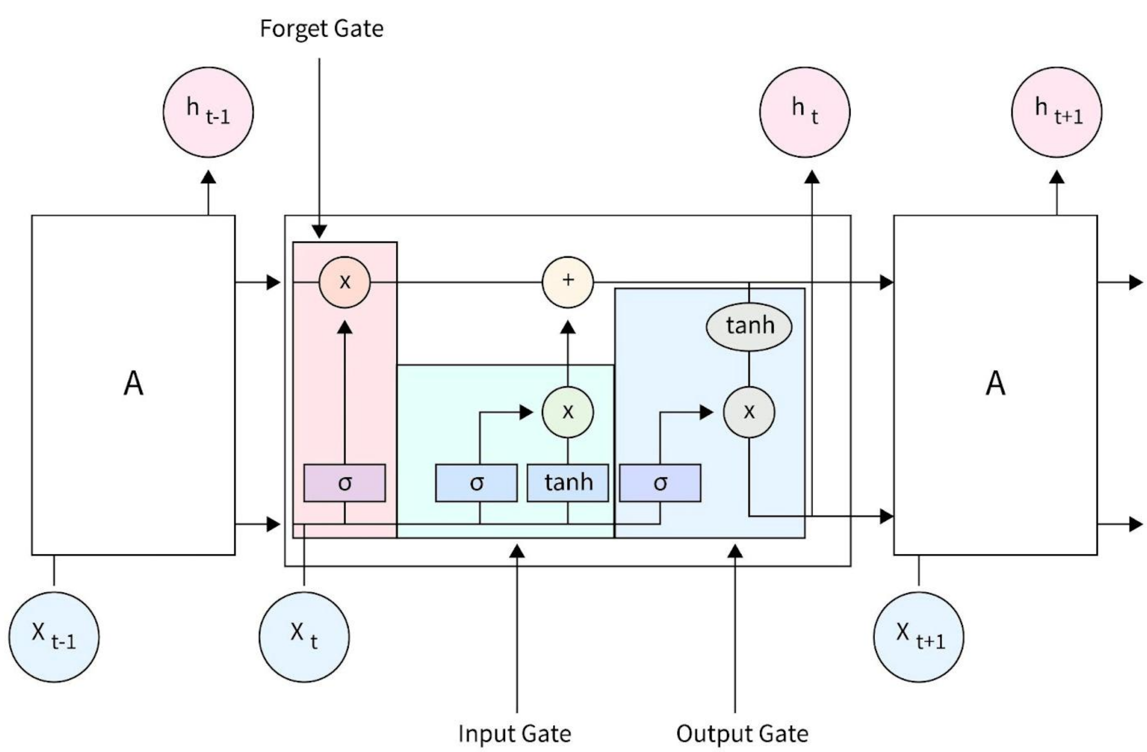


Distribution of Positive/Negative Words in Reviews
Positive Words
Negative Words
Frequency
Number of Words per Review
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0

3. Methodology

- Text Preprocessing:
    - Performed using regex, NLTK, and spaCy.
    - Tokenization and sequence padding were done using Keras Tokenizer.

- Model Architectures:
    - Simple RNN model: Embedding + SimpleRNN + Dense
    - LSTM Model: Embedding + LSTM + Dense
    - LSTM + Word2Vec: Non-trainable pretrained GloVe embedding + LSTM + Dense.



The architecture shows a specific layout for a text classification model using an LSTM layer. This includes a built-in layer, followed by an LSTM block and a few dense layers for classification. The built-in layer converts the entrance toes to a certain dimensional vector, the LSTM team processes sequential information, and the dense teams finally classify.

The use of embeddings transforms words into numerical vectors that preserve the importance of semantic significance. Simple RNN process sequence but struggles with long addiction. LSTMS addresses it with a port mechanism that retains information over time. Adding Word2vec further improves the understanding of LSTM using pretrained semantic knowledge from a large corpus.

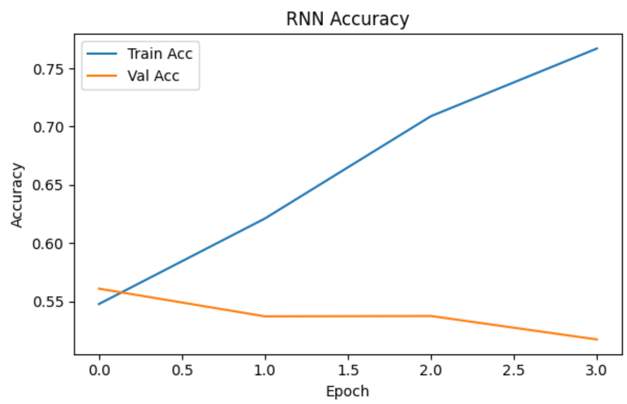
- Loss Function: Binary cross-entropy
- Optimizer: Adam
- Hyperparameters: Epochs (10–30), Batch size (32–64), Learning rate (default)

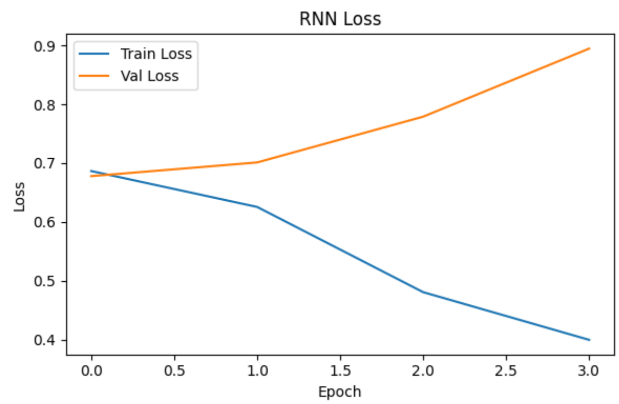
4. Experiments and Results

4.1 RNN vs. LSTM Performance

The RNN model was the simplest of three architecture and served as a baseline. However, this lesson struggles to capture long-term dependencies, leading to low accuracy and unstable confirmation. Training accuracy was appropriate, but the confirmation accuracy remained flat, which suggested poor generalization.

In contrast, the LSTM model improved both training and confirmation accuracy. The ability to maintain reference on long -term sequences led to better learning and stability. It also showed more consistent convergence in ages.





4.2 Computational Efficiency

The RNN model trained quickly because of its simple architecture and low parameters. However, it came at the expense of performance. The LSTM model required several calculation resources, and it took longer to train on a large data set but was worth acting for better accuracy and stability.

4.3 Embedding Comparison

The randomly arranged built-in allowed the model to teach the representation during training, but the first meaning was missing. Word2vec built-in, which pretended to be a large corpus, gave a semantic head start, so that the model could understand the word relationship from the beginning. This rapidly improved convergence and classification accuracy.

4.4 Model Evaluation Metrics

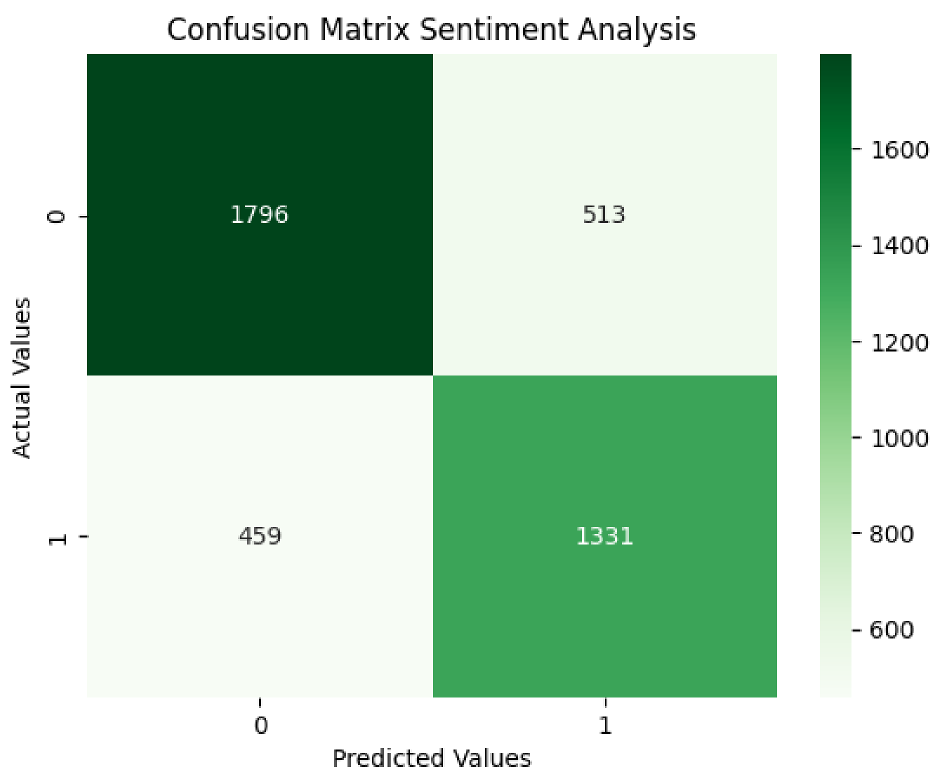
To consider performance a lot:

- Accuracy: The general purity of predictions. The LSTM + Word2vec model got the highest accuracy.

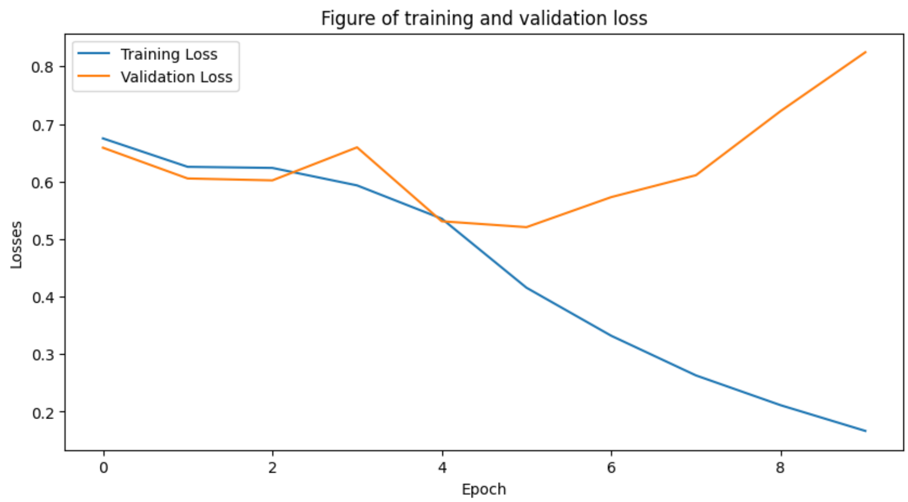
- Confusion matrix: Visual representation of true vs. predicted classes. It was discovered that most misclassifications occurred in the middle of the border.

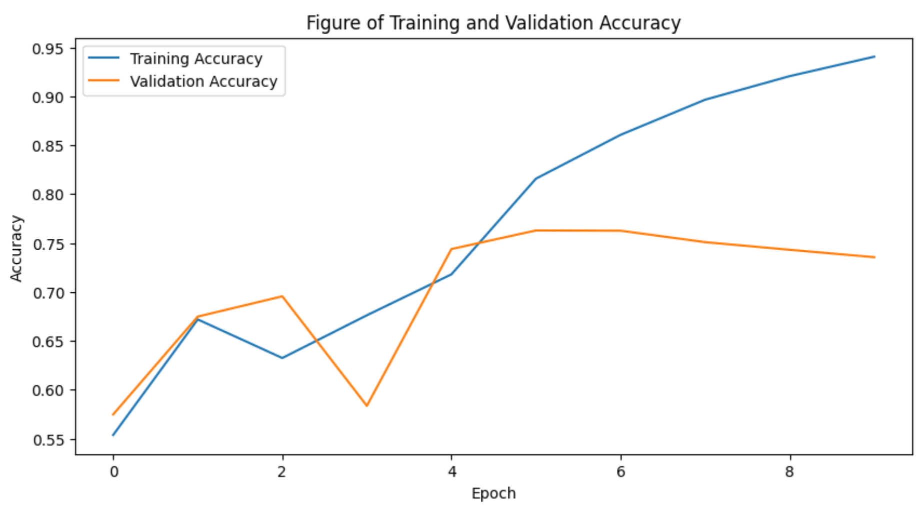
- Precision, Recall, F1-Score: Precision measured accuracy, measured perfection and F1-score gave a harmonious balance. These metrics exposed imbalance in the class and helped with the fine tuning of threshold.

- Confusion matrix value:
    - True Positive (TP): 1331
    - TRUE Negative (TN): 1796
    - False Positive (FP): 513
    - False negative (UN): 459



- Training Metrics:





Visualization highlights trends in learning performance. Verification losses began to grow and grow after several epochs, indicating overfitting. Results of confusion matrix confirm that although the model is generally accurate, both classes have a certain misclassification, especially with borderline sentiments.

5. Conclusion and Future Work

This study shows the effectiveness of LSTM, especially when the sentiment classification is increased with Word2vec built into tasks. RNNS, while simpler, underperforms compared to LSTM. pretrained built -in generalization. However, overfit was observed after several ages. Future work may be focused on:

- Hyperparameter tuning
- Large and more balanced dataset
- More robust models (e.g., BiLSTM, GRU, Transformers)
- Deployment using real-time interfaces (e.g., Gradio, Streamlit)

The project provides a strong foundation for future explorations in sequence-based natural language processing functions. Creating an interactive sentiment classification tool can be the next step, so users can enter real reviews and get instant feedback. It also determines the phase to expand emotional analysis for multilingual or domain -specific data set for broader applicability.

References

Muhammad, P. F. (2021). Sentiment analysis using Word2Vec and long short-term memory (LSTM) for Indonesian hotel reviews. *Procedia Computer Science*, 728–735.

Rajshree, M. &. (2022). Comparative analysis of hotel reviews using proposed LSTM based deep learning techniques. . *AIP Conference Proceedings*, 2481.