UNDERSTANDING READEREMOTIONS THROUGHASPECT-BASED SENTIMENT ANALYSIS

Surabodee Phasuk , Rapin Narach , Suttipot Prathoumthong
Department of Computer ScienceFaculty of Science and Technology.
Thammasat University ,Pathum Thani, Thailand

Abstract— Based Sentiment Analysis (ABSA) identifies sentiment polarity toward specific aspects rather than assigning a single overall sentiment to a text. This project compares two modeling approaches for ABSA: (1) a traditional Logistic Regression model using TF-IDF and n-gram features, and (2) a Bidirectional Long Short-Term Memory (Bi-LSTM) neural network. A dataset containing 21,605 review–aspect entries was cleaned and reduced to 20,684 valid samples. Both models were evaluated on accuracy and their ability to handle complex sentiment structures. The Logistic Regression baseline achieved marginally higher accuracy (68%) than the Bi-LSTM (67%). However, detailed analysis—especially with conflict cases—shows that the Bi-LSTM model better understands contextual sentiment relationships. These findings highlight the limitations of accuracy-only evaluation and emphasize the importance of contextual modeling in ABSA.

Keywords— Aspect-Based Sentiment Analysis, Bi-LSTM, Logistic Regression, NLP, Deep Learning

I. INTRODUCTION

Sentiment analysis traditionally predicts a single sentiment for an entire sentence, but real-world reviews often express different opinions regarding multiple aspects of a product or service. Aspect-Based Sentiment Analysis (ABSA) addresses this limitation by determining the sentiment associated with each specific aspect mentioned.

Machine Learning (ML) and Deep Learning (DL) approaches are widely used in sentiment classification. ML models such as Logistic Regression are simple and efficient but rely heavily on handcrafted features. DL models such as LSTM and Bi-LSTM capture long-term dependencies and understand sentence structure more effectively. This work compares Logistic Regression and Bi-LSTM to determine which performs better for ABSA under controlled preprocessing and evaluation conditions.

II. RELATED WORK

Traditional sentiment classification approaches use bag-of-word features, TF-IDF, and algorithms such as SVM and Logistic Regression. These methods work well for general sentiment tasks but lack the ability to handle complex sentencestructures.

Deep learning introduced LSTM and Bi-LSTM architectures capable of learning sequential dependencies and contextual relationships between words, which significantly improved ABSA performance. More recent approaches incorporate attention mechanisms and Transformer-based models such as BERT, which push state-of-the-art performance further. However, this study limits its comparison to classical ML and LSTM-based DL to understand their fundamental differences.

III. DATA AND PREPROCESSING

A. Dataset Overview

The original dataset contained 21,605 review rows. After removing missing entries, short sentences, and incorrect sentiment labels, 20,684 samples remained.

Each entry contained:

- Aspect term
- Review sentence
- Sentiment label (Positive = 1, Neutral = 0, Negative = -1)

Data was split into:

- Training set: 16,547 samples (80%)
- Test set: 4,137 samples (20%)

B. Preprocessing Pipeline

The preprocessing steps included:

1. Lowercasing and text normalization
2. Tokenization
3. Optional stopword removal (kept for LSTM, removed for TF-IDF)
4. TF-IDF (for baseline) using n-grams (1,2)

5. Integer encoding and padding to max length of 100 (for LSTM)
6. Constructing ABSA-specific input format: “aspect <sep> sentence” to explicitly signal the aspect
7. Lemmatization for baseline; raw tokens preserved for LSTM
8. Train-test split with a fixed random seed

C. *Class Imbalance Handling*

The dataset exhibited imbalance, with positive reviews being themajority class.
For the Bi-LSTM training:

- Neutral weighted ×1.33
- Negative weighted ×1.20
- Positive weighted ×1.00
  Baseline model used class_weight='balanced' as needed.

IV. Models

A. *Logistic Regression (Baseline)*

bigrams) as input features.
Key parameters:

- N-grams: (1,2)
- Max features: 50,000
- Regularization: L2
- Solver: lbfgs
- Multi-class: softmax-like probabilistic output

Advantages: low complexity, interpretable, efficient.
Weaknesses: cannot model word order or long-range dependencies.

B. *Bi-LSTM Model*

Architecture:

- Embedding layer (dim = 300)
- Bidirectional LSTM (128 units, dropout = 0.3)
- Dense layer (64 units, ReLU)
- Output softmax layer (3 classes)
- Loss: Categorical Cross-Entropy
- Optimizer: Adam (lr = 0.001)
- Epochs: 10–20 with early stopping
- Batch size: 32

Bi-LSTM processes sentences forward and backward, giving it better contextual awareness.

```
Model: "sequential"
```

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 100, 100)	2,000,000
bidirectional (Bidirectional)	(None, 128)	84,480
dense (Dense)	(None, 32)	4,128
dropout (Dropout)	(None, 32)	0
dense_1 (Dense)	(None, 3)	99

```
Total params: 2,088,709 (7.97 MB)

Trainable params: 2,088,707 (7.97 MB)

Non-trainable params: 0 (0.00 B)

Optimizer params: 2 (12.00 B)
```

Figure 1. Model Architecture of the Bidirectional LSTM for ABSA

V. Evaluation Methodology

A. Metrics

The following were used:

- Accuracy
- Precision, Recall, F1-score (per class)
- Macro-averaged F1
- Confusion matrices
- Manual testing with challenge set
- LIME explainability analysis

B. Challenge Set for Deeper Understanding

Four scenarios designed:

1. Positive/Positive
2. Negative/Negative
3. Neutral/Neutral
4. Positive/Negative (conflicting sentiment → most difficult)

These cases highlight model behavior beyond accuracy.

VI. Results

A. Overall Performance

Model	**Accuracy**
Logistic Regression	68%
Bi-LSTM	67%

```
==================================================
📊 รายงานผลโมเดล : Baseline (Logistic Regression)
==================================================
              precision    recall  f1-score   support

    Negative       0.67      0.62      0.65      1195
     Neutral       0.58      0.40      0.47      1025
    Positive       0.72      0.87      0.79      1917

    accuracy                           0.68      4137
   macro avg       0.66      0.63      0.63      4137
weighted avg       0.67      0.68      0.67      4137


==================================================
📊 รายงานผลโมเดล: Bidirectional LSTM (ของคุณ)
==================================================
              precision    recall  f1-score   support

    Negative       0.62      0.65      0.63      1195
     Neutral       0.52      0.48      0.50      1025
    Positive       0.77      0.77      0.77      1917

    accuracy                           0.67      4137
   macro avg       0.64      0.64      0.64      4137
weighted avg       0.66      0.67      0.66      4137
```

Figure 2. Performance Comparison Between LR and Bi-LSTM

B. Challenge Set Results

1. Positive–Positive Sentence

Both models predicted correctly; Bi-LSTM showed higher confidence.

2. Negative–Negative Sentence

Both models predicted correctly; Bi-LSTM confidence was higher.

3. Neutral–Neutral Sentence

Both models predicted correctly; Bi-LSTM demonstrated stronger stability.

4. Positive–Negative

(Conflict Case)
Sentence: "The plot was amazing, but the characters were boring."

- Logistic Regression → predicted *Neutral* for both aspects (0% correct).
- Bi-LSTM → predicted *Positive* for plot (correct) but *Positive* for characters (incorrect).

Interpretation:

- Logistic Regression averages positive & negative words → fails in mixed-sentiment cases.
- Bi-LSTM recognizes the sentiment of "amazing" strongly but is overly influenced by the earlier positive term.

C. LIME Analysis

LIME reveals:

- Strong positive terms ("amazing") dominate model attention.
- Negative terms ("boring", "terrible") push prediction negative.
- Bi-LSTM recognizes discourse markers ("but") but not strongly enough.
- Logistic Regression treats the entire sentence globally, losing aspect focus.

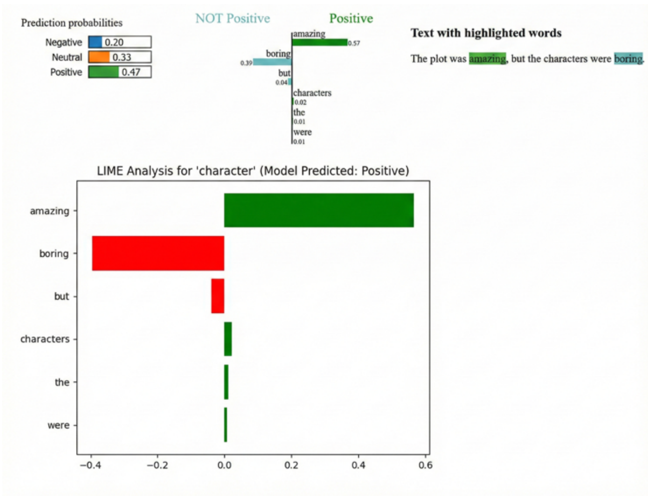

Figure 3. LIME Explanation for the Conflict Case

VII. Conclusion

This study compared Logistic Regression and Bi-LSTM for Aspect-Based Sentiment Analysis. Logistic Regression achieved slightly higher accuracy but performed poorly in conflict scenarios and complex contextual situations. Bi-LSTM, despite lower accuracy, demonstrated stronger understanding of aspect-level relationships and better handled mixed sentiments.

Therefore, for real-world ABSA tasks, models must be evaluated on contextual reasoning—not just accuracy. Bi-LSTM and Transformer-based models are recommended for deployments requiring deeper semantic understanding.

REFERENCES

[1] B. S. Harish, A. G. S. Badrinath, M. J. M., "Classification of Amazon Book Reviews Based on Sentiment Analysis," in International Conference on Intelligent Systems (ICIS), 2018.[Online]Available:https://www.researchgate.net/profile/BHarish/publication/323503127_Classification_of_Amazon_Book_Reviews_Based_on_Sentiment_Analysis/links/5aa8d70caca2726f41b194f9/Classification-of-Amazon-Book-Reviews-Based-on-Sentiment-Analysis.pdf

[2] J. Weng, "Comparison of machine learning-based book comments sentiment analysis for constructing recommendation system," in Proceedings of the 4th International Conference on Signal Processing and Machine Learning,2024,Art.No.10907.

[3] L. Lyuhao, H. Ashraf, H. Jazri, and T. Parashar, "Emotion Prediction in Amazon Kindle Book Reviews: Leveraging RNN for Sentiment Analysis," in 2024 IEEE 1st Karachi Section Humanitarian Technology Conference (KHI-HTC), 2024,pp.1–6.

[4] N. I. Lestari, I. A. Aziz, S. M. Taib, and M. R. Habibi, "Aspect-Based Sentiment Analysis for Mobile App Review Using Convolutional Neural Network (CNN) and Word2Vec," in 2024 IEEE 7th International Conference on Electrical, Electronics and System Engineering (ICEESE), 2024,pp.1–6.

[5] V. R. Azhaguramyaa, J. Janet, G. R. Madhavan, S. Balakrishnan, and K. Arunkumar, "Sentiment Analysis on Book Reviews Using Machine Learning Techniques," in 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), 2022, pp. 1–6.

[6] M. Nagadia, "Amazon Kindle Book Review for Sentiment Analysis," Kaggle. [Online]. Available: https://www.kaggle.com/datasets/meetnagadia/amazon-kindle-book-review-for-sentiment-analysis/data