Aspect-Based Sentiment Analysis of Financial News Using Transformer Models

Student Name: Srikar Dayala

Student ID: 23038647

1. Introduction

Fine-grained sentiment analysis of financial news is challenging because sentiment is often expressed toward specific financial entities rather than entire headlines, and multiple entities may appear with differing sentiments. To address this challenge, the **SEntFiN** dataset provides human-annotated, entity-level sentiment labels for financial news headlines (Sinha et al., 2022).

This project focuses on **aspect-based sentiment analysis (ABSA)**, where sentiment (positive, neutral, or negative) is predicted for a given financial entity within a headline. A traditional **Logistic Regression** model is used as a baseline and compared with two transformer-based models: **FinBERT**, which is pre-trained on financial text, and **RoBERTa**, a general-purpose transformer model. The goal is to evaluate whether contextual transformer models outperform feature-based approaches on this task.

2. Dataset and Task Description

The experiments use the **SEntFiN 1.1** dataset (Sinha et al., 2022), which contains **10,700+ financial news headlines** with entity-level sentiment annotations. Approximately **2,800 headlines include multiple entities**, often with conflicting sentiments. The dataset is relatively balanced, with around **4,100 positive**, **3,200 negative**, and **4,500 neutral** entity-level labels.

Each headline was converted into a **sentence–aspect pair** by associating the headline with a single target entity and its sentiment. The task is formulated as a **three-class classification problem**. The data was split into **80% training and 20% testing sets** using stratified sampling to preserve class distribution.

3. Methodology

3.1 Logistic Regression Baseline: A Logistic Regression classifier was implemented using TF-IDF unigram and bigram features, limited to 5,000 features. This bag-of-words approach does not capture contextual or semantic relationships between words and therefore serves as a strong but interpretable baseline against which the benefits of contextual transformer-based models can be assessed.

3.2 FinBERT: FinBERT is a transformer model pre-trained on large-scale financial text and made available through Hugging Face (Araci, 2019). It was fine-tuned for aspect-based sentiment classification using sentence–aspect pairs as input, allowing sentiment to be

conditioned explicitly on the target entity. Training was performed for three epochs using the AdamW optimiser with a learning rate of 2e-5, batch size of 16, and a maximum sequence length of 128 tokens, with a linear learning-rate scheduler and warm-up. This configuration follows standard best practices for transformer fine-tuning and balances convergence speed with overfitting risk.

3.3 RoBERTa: RoBERTa is a robustly optimised, general-purpose transformer model (Hugging Face, n.d.). It was fine-tuned using the same training configuration as FinBERT to ensure a fair and controlled comparison between a domain-specific and a general-purpose language model. Sentence–aspect pairs were formatted using RoBERTa-specific separator tokens, ensuring compatibility with the model's pre-training scheme.

4. Results and Discussion

4.1 Quantitative Performance Comparison

Model performance was evaluated using accuracy, weighted precision, weighted recall, and weighted F1-score. Weighted metrics were used to account for residual class imbalance and ensure that performance across all sentiment classes was reflected fairly. Table 1 summarises the results.

Table 1: Quantitative Performance Comparison

Model	Accuracy	Precision (Weighted)	Recall (Weighted)	F1-score (Weighted)
Logistic Regression	0.779	0.782	0.779	0.778
FinBERT	0.880	0.880	0.880	0.880
RoBERTa	0.899	0.899	0.899	0.898

Both transformer-based models substantially outperform the Logistic Regression baseline, demonstrating the importance of contextual representations for entity-level sentiment analysis. FinBERT shows strong gains due to financial-domain pretraining; however, RoBERTa achieves the highest overall performance, suggesting that robust general-purpose language representations can effectively capture nuanced sentiment cues in financial news. The relatively small performance gap between FinBERT and RoBERTa further indicates that domain-specific pretraining is beneficial but not strictly necessary for this task.

4.2 Confusion Matrix Analysis (RoBERTa)

Figure 1 presents the confusion matrix for the RoBERTa model. The majority of predictions lie along the diagonal, indicating strong classification performance across all sentiment classes. Negative and positive sentiments are identified with high accuracy, reflecting effective recognition of explicit sentiment cues. Most misclassifications occur between the neutral and positive classes, highlighting the inherent ambiguity of weakly positive financial language.
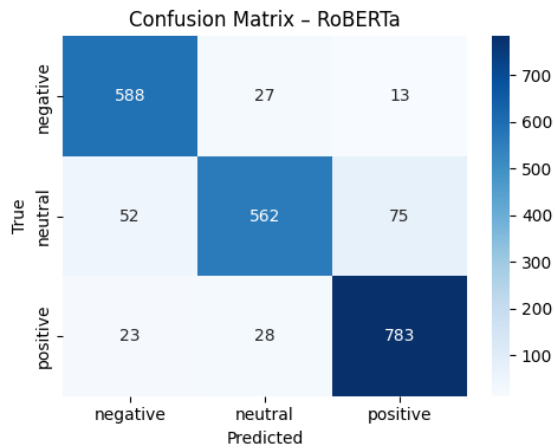
Figure 1: Confusion Matrix – RoBERTa

4.3 Qualitative Analysis: Actual vs Predicted

Qualitative analysis was conducted by comparing actual and predicted sentiment labels for selected test examples (Figure 2). RoBERTa correctly classifies headlines with explicit positive cues, such as earnings growth announcements, but struggles with weakly positive or cautiously phrased statements. For example, *"Outlook not bleak for Infosys going ahead"* is predicted as neutral despite a positive ground-truth label, illustrating the difficulty of detecting implicit sentiment in financial news.

```
=== RoBERTa: Actual vs Predicted ===

Sentence   : Outlook not bleak for Infosys going ahead: Prakash Diwan, Altamount Capital
Aspect     : Infosys
Actual     : positive
Predicted  : neutral
-------------------------------------------------------------------
Sentence   : Colgate-Palmolive Q2 net profit up 11.8% at Rs 100 cr
Aspect     : Colgate-Palmolive
Actual     : positive
Predicted  : positive
-------------------------------------------------------------------
Sentence   : HPCL Q4 net up 48 pc at 1,122.66 crore
Aspect     : HPCL
Actual     : positive
Predicted  : positive
```

Figure 2: Actual vs Predicted Sentiment Examples (RoBERTa)

5. Conclusion

This study evaluates aspect-based sentiment analysis on the SEntFiN dataset using a traditional Logistic Regression baseline and two transformer-based models. The results demonstrate that transformer models consistently outperform feature-based approaches by capturing contextual and entity-specific sentiment cues. FinBERT achieves strong performance due to financial-domain pretraining, while RoBERTa attains the highest overall accuracy, indicating robust generalisation. Remaining errors largely stem from ambiguity between neutral and weakly positive sentiment, reflecting limitations of discrete sentiment labels in financial text. Overall, transformer-based models offer an effective and reliable approach for entity-level sentiment analysis of financial news.

References

Araci, D.T. (2019) *FinBERT: Financial sentiment analysis with pre-trained language models*, *arXiv preprint*, arXiv:1908.10063. Available at: https://arxiv.org/abs/1908.10063

Hugging Face (n.d.) *RoBERTa* [Online]. Available at: https://huggingface.co/docs/transformers/en/model_doc/roberta

Sinha, A. (2022) *Aspect-Based Sentiment Analysis for Financial News* [Online]. Available at: https://www.kaggle.com/datasets/ankurzing/aspect-based-sentiment-analysis-for-financial-news