ISSN 2522-9818 (print)
ISSN 2524-2296 (online) *Innovative technologies and scientific solutions for industries. 2023. No. 3 (25)*

UDC 004.8 DOI: https://doi.org/10.30837/ITSSI.2023.25.138

V. IVANENKO

ENHANCING ASPECT-BASED FINANCIAL SENTIMENT ANALYSIS THROUGH CONTRASTIVE LEARNING

The **subject** of research in the article explores the specialized application of Aspect-Based Financial Sentiment Analysis (ABFSA), focusing on the intricate and multifaceted emotional landscape of financial textual data. The study extends the current understanding of sentiment analysis by addressing its limitations and opportunities within a financial context. The **purpose** of the work is to advance the field of Aspect-Based Financial Sentiment Analysis by developing a more nuanced and effective methodology for analyzing sentiments in financial news. Additionally, the study aims to assess the efficacy of recent advancements in Natural Language Processing (NLP) and machine learning for enhancing ABFSA models. The article deals with the following **tasks**: Firstly, the study focuses on the rigorous pre-processing of the SEntFiN dataset to make it more amenable to advanced machine learning techniques, specifically contrastive learning methodologies. Secondly, it aims to architect a unified model that integrates state-of-the-art machine learning techniques, including DeBERTa v3, C^2L contrast learning, and LoRa fine-tuning. Lastly, the research critically evaluates the proposed model's performance metrics across the test dataset and compares them with existing methodologies. The following **methods** are used: Firstly, the study employs pre-processing techniques tailored for the SEntFiN dataset, which is explicitly designed for entity-sensitive sentiment analysis in financial news. Secondly, it utilizes advanced machine learning techniques such as DeBERTa v3 for language model pre-training, C^2L contrast learning for focusing on causal relationships, and LoRa for fine-tuning large language models. Lastly, performance evaluation methods are used to assess the efficacy of the proposed model, including comparisons with existing methodologies in the field.The following **results** were obtained: The study reveals that the proposed pre-processing framework successfully accommodates the variable number of entities present in financial news, thereby improving the granularity of sentiment classification. Furthermore, the integration of advanced NLP and machine learning techniques significantly enhances the accuracy and efficiency of ABFSA models. **Conclusions**: The paper concludes that specialized ABFSA methodologies, when augmented with advanced NLP techniques and a robust pre-processing framework, can offer a more nuanced and accurate representation of sentiment in financial narratives. The study lays the groundwork for future research in this nascent yet crucial interdisciplinary field, providing actionable insights for stakeholders ranging from investors to financial analysts.

Keywords: Aspect-based Financial Sentiment Analysis; Contrastive Learning; Text Classification.

Introduction

Sentiment analysis, a traditional subfield of Natural Language Processing (NLP), focuses on identifying and quantifying the affective and subjective nuances present in textual data. While its applications encompass diverse fields such as social media analytics, consumer feedback assessment, and political opinion mining, the financial sector is a particularly intriguing but underexplored domain.

Economic indices, corporate performance metrics, and geopolitical developments, among others, contribute to the inherent complexity of the financial markets. In the midst of these complexities, human emotion has come to be recognized as a crucial, albeit volatile, factor capable of influencing market dynamics. As a result, market participants such as investors, traders, and financial analysts are increasingly turning to quantitative methodologies to evaluate market sentiment, highlighting the importance of automated sentiment analysis in financial contexts. This lays the groundwork for a study of Aspect-based Financial Sentiment Analysis.

Historically, the financial sector has relied mainly on human expertise for the interpretation of various informational sources, ranging from reports and corporate releases to social media discourse, in order to forecast market trends. Although valuable, this method is labor-intensive and susceptible to cognitive biases. Advanced computational methods, particularly those based on natural language processing, offer an attractive alternative. These algorithms provide benefits in terms of speed and scalability, and they also mitigate human error to some extent. In spite of this, the diverse and complex nature of financial textual data, which includes news reports, trading summaries, and opinionated articles, among others, presents specific challenges that generic sentiment analysis models are frequently incapable of addressing. This circumstance necessitates the development of sentiment analysis methodologies that are tailored to the specific needs of the financial industry.

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2023. № 3 (25) ISSN 2524-2296 (online)

In order to navigate the application of sentiment analysis to financial text, "aspect-based" scrutiny is required. Traditional models for sentiment analysis typically divide textual data into broad affective categories, such as positive, negative, or neutral. Nevertheless, Aspect-Based Sentiment Analysis (ABSA) provides a more nuanced perspective by identifying sentiments associated with distinct entities or aspects of the text. For instance, a financial news article may simultaneously report the rapid growth of a technology company (a positive sentiment aspect) and upcoming regulatory challenges (a negative sentiment aspect). In such a situation, a conventional sentiment analysis algorithm might fail to capture this contradictory sentiment landscape, whereas an ABSA model would provide a segmented and multidimensional emotional profile.

The present study is motivated by both the latent potential of ABSA and the existing methodological limitations when operating in a financial environment. While general research on sentiment analysis has been exhaustive, its application to financial news and market dynamics is still in its infancy. Modern models frequently fail to comprehend the nuanced and diverse emotions embedded in financial narratives. This deficiency is primarily attributable to two limitations: the lack of specialized training datasets for financial sentiment analysis and the inadequacy of current NLP techniques for deciphering complex entity relationships.

The SEntFiN dataset [4] arises as a crucial asset in this context. Specifically designed for entity-sensitive sentiment analysis in the realm of financial news, this dataset is an indispensable resource for the development and evaluation of algorithms specifically tuned for the finance industry. The SEntFiN dataset is not, however, devoid of its own set of complexities. The variable number of entities featured in financial news headlines presents a significant difficulty, which complicates the algorithmic task of sentiment classification.

In addition, the advent of recent innovations in machine learning and Natural Language Processing (NLP) namely, the DeBERTa v3 model [3], C^2L contrast learning [2], and LoRa fine-tuning [1] – ushers in novel paradigms for executing Aspect-Based Sentiment Analysis (ABSA). Each of these technological advancements contributes distinct capabilities to the analytical toolkit. For instance, DeBERTa v3 enhances language model pre-training through its gradient-disjoint embedding sharing. C^2L contrast learning, on the other hand, augments text classification by focusing on causal relationships. LoRa contributes by fine-tuning large language models to be more amenable to specialized tasks like ABSA. The synergistic amalgamation of these advanced techniques holds promise for considerably elevating the efficacy of ABSA models within the financial sphere.

Analysis of last achievements and publications

Historically, the finance sector has primarily been the arena of quantitative data analysis, with an emphasis on numerical indicators such as stock prices, trade volumes, and economic indices. Nonetheless, the burgeoning awareness that market dynamics are also significantly influenced by human sentiment has led to the emergence of a specialized field: financial sentiment analysis. One of the seminal contributions in this arena was made by Tetlock in 2007, who demonstrated the considerable impact of media content on market fluctuations. This groundbreaking work underscored the necessity for automated sentiment analysis methodologies within the financial sector. Additionally, influential research by Antweiler and Frank in 2004 probed the ramifications of user-generated content in online forums and message boards on stock market volatility. Their findings corroborated the hypothesis that textual data could offer complementary insights beyond what is conveyed by numerical indicators alone, thereby reinforcing the critical role of sentiment analysis in financial decision-making processes.

Financial Sentiment Analysis

Within the domain of sentiment analysis, the niche focus on financial texts has evolved into a distinct sub-discipline termed financial sentiment analysis. A diverse array of methodological paradigms has been advanced in scholarly literature to address this domain's unique complexities. Among these, FinBERT [8] occupies a pivotal position, having been tailored around the BERT architecture. This model undergoes pre-training on a substantial corpus of financial text and is further fine-tuned to excel in financial sentiment analysis tasks. It has consistently yielded state-of-the-art results across multiple evaluative benchmarks.

Complementing the capabilities of FinBERT is FinBERT-MRC [9], designed for financial named entity recognition. It leverages the BERT architecture within the milieu of machine reading comprehension. Empirical assessments corroborate that FinBERT-MRC

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2023. No. 3 (25)

outperforms its antecedents in the accuracy of financial named entity recognition tasks. A more recent contribution to this evolving field is MFinBERT [11], a multilingual pre-trained language model. MFinBERT is calibrated on a voluminous multilingual corpus of financial text and has registered state-of-the-art achievements across a gamut of financial tasks, including but not limited to sentiment analysis.

Another innovative direction in financial sentiment analysis is instruction-based fine-tuning, as exemplified by Zhang et al.'s Instruct-FinGPT [20]. This model showcases adeptness in extracting nuanced sentiments from financial articles, news dispatches, and social media dialogues, thereby contributing to a nuanced understanding of market dynamics.

Among specialized language models, BloombergGPT by Wu et al. is particularly noteworthy. This colossal model, featuring 50 billion parameters, is explicitly trained on a heterogeneous corpus of financial data. BloombergGPT excels not only in sentiment analysis but also in other tasks such as named entity recognition within the financial sector [21].

On the frontier of open-source contributions, Yang et al.'s FinGPT represents a seminal initiative. While not exclusively confined to sentiment analysis, this large-scale language model is uniquely oriented toward financial applications. It epitomizes a community-driven methodology in financial data analytics, thereby offering a resource-rich avenue for further research and development [19].

Language Models in Financial Entity Recognition

FiNER and SEC-BERT [7] represent specialized approaches in the realm of financial numeric entity recognition. Harnessing the power of the BERT architecture, these models excel at identifying and categorizing relevant entities within financial reports. Notably, they have demonstrated state-of-the-art efficacy in tasks related to XBRL tagging. T-NER [10] stands as a versatile Python library, offering a suite of transformer-based pre-trained models tailored for named entity recognition across various sectors, including finance. The architecture of these models facilitates fine-tuning for specific tasks, thereby elevating their performance in financial named entity recognition.

Extensions into Other Specialized Domains

While financial sector-specific language models have witnessed considerable advancements, parallel developments have occurred in other specialized sectors as well. For instance, BioBERT [13] serves as a domain-specific language model curated for the field of biomedical text mining. Pre-trained on an expansive corpus of biomedical literature, BioBERT has consistently achieved state-of-the-art results on multiple text-mining tasks within the biomedical domain.

Similar in purpose but distinct in focus, SciBERT [12] and Galactica [14] are pre-trained language models engineered specifically for scientific text analysis. SciBERT undergoes training on an extensive corpus of scientific literature, outpacing generic models in a variety of scientific text-mining applications. Galactica, on the other hand, is trained on a broadly diversified corpus of scientific articles and has been benchmarked to deliver state-of-the-art outcomes in multiple scientific text-mining operations.

These domain-specific language models, each fine-tuned to the nuances of their respective fields, collectively illustrate the expanding horizons and increasing sophistication of language model applications across disciplines.

Contrastive Learning Techniques

Chen et al. introduced a significant contribution to NLP with CLUSE [22]. Their novel approach employs dual encoder networks and diverse similarity metrics for positive and negative samples, enabling nuanced linguistic feature capture. Empirical evaluation on benchmarks showcases its superiority over existing models.

In a related study, Chen et al. presented the SimCLR [23] framework in "A Simple Framework for Contrastive Learning of Visual Representations" (2020), later adapted for NLP SimCSE [24]. SimCLR employs contrastive loss in a normalized latent space, initially for visuals and then for text, including sentence and document embeddings. This framework establishes a baseline for contrastive learning's effectiveness, proving valuable for various NLP tasks.

Within the broader context of text classification, the intersection of contrastive learning and counterfactual augmentation is noteworthy, exemplified by the proposed C^2L Contrastive learning method [2]. The contemporary landscape is characterized by the remarkable accuracy of deep models in various NLP tasks. However, this proficiency is shadowed by concerns related to vulnerabilities arising from their reliance on spurious patterns. In response, the authors embark on a pioneering

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2023. № 3 (25) ISSN 2524-2296 (online)

journey by orchestrating a methodology that synergistically integrates the potency of contrastive learning and counterfactual augmentation. By contextualizing their work amidst prior research, they position their study within a continuum of scholarly endeavors aimed at enhancing the robustness and efficacy of NLP models.

Summary and Gap Identification

Within the current body of literature, an encompassing foundation pertaining to sentiment analysis has been established, ranging from general contexts to more intricate domains such as finance. Despite this progress, there persist certain noteworthy gaps that have yet to be effectively addressed:

1. Firstly, the potential synergy resulting from the convergence of advanced machine learning techniques, such as DeBERTa v3 and C^2L, within the realm of financial sentiment analysis remains largely unexplored. This uncharted territory offers opportunities for innovative insights and advancements.
2. Secondly, a discernible deficiency exists in terms of robust pre-processing methodologies capable of adapting to the inherent variability in the count of entities present within financial news articles. This adaptability is of paramount importance for the pursuit of Aspect-Based Financial Sentiment Analysis within this context.
3. Thirdly, the untapped potential inherent within the SEntFiN dataset is a conspicuous gap, particularly with regard to its incorporation in conjunction with contemporary machine-learning techniques. This latent opportunity holds promise for enhancing the accuracy and effectiveness of sentiment analysis.

The aim of the work

Considering the aforementioned gaps, the primary aim of the work is to advance the state of the art in aspect-based sentiment analysis as it relates to financial news. The specific objectives are

1. To rigorously pre-process the SEntFiN dataset, rendering it more receptive to contrastive learning methodologies.
2. To architect a unified model that integrates advanced machine learning techniques, particularly DeBERTa v3, C^2L contrast learning, and LoRa fine-tuning.
3. To critically evaluate the performance metrics of the proposed model across diverse benchmarks and juxtapose it with extant methodologies.

Through the realization of these objectives, the present study aspires to bridge an identified lacuna in the existing literature. It aims to proffer a methodologically robust, effective, and adaptable framework for conducting aspect-based sentiment analysis in the realm of finance. Ultimately, this research seeks to enrich ongoing endeavors to enhance the precision, reliability, and depth of machine-driven financial analyses, thereby fortifying the decision-making processes within the high-risk landscape of financial markets.

Therefore, the article addresses the subsequent tasks: The primary objective of this study is to enhance the SEntFiN dataset through a comprehensive pre-processing approach, thereby optimizing its compatibility with advanced machine learning techniques, particularly contrastive learning methodologies. Furthermore, the objective is to design a cohesive framework that incorporates cutting-edge machine learning methodologies such as DeBERTa v3, C^2L contrast learning, and LoRa fine-tuning. Finally, the study conducts a critical evaluation of the performance metrics of the proposed model on the test dataset and compares them with those of established methodologies.

Materials and methods

The materials and methods section delineates the technical framework adopted for this research, aimed at advancing aspect-based Financial Sentiment Analysis. The research design integrates three key components: data pre-processing of the SEntFiN dataset, the implementation of a hybrid model comprising DeBERTa v3, C^2L contrast learning, and LoRa fine-tuning, and an extensive evaluation strategy. Each of these components plays a critical role in achieving the study's objectives, as outlined in the Introduction.

Mathematical statement of the problem

To provide a clear and precise definition of the research problem, I will outline its mathematical framework. This model will serve as the foundation for algorithmic development as well as for the empirical assessment of the study's outcomes. The components of this mathematical model are defined as follows:

- $D\{(x, y)\}$: Represents the dataset of financial news titles paired with labels, where $(x_i, y_i) \in D$ and $i = 1, 2, ..., n$;

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2023. No. 3 (25)

- L: Multi-class label set $\{Positive,\ Negative,\ Neutral\}$;
- $Tokenizer(x_i) \to t_i$: Tokenazer which creates the sequence of tokens from x_i, where $t_i = [t_i^1, t_i^2, \ldots t_i^T]$

and T is the max size of the sequence;

- $Encoder(t_i) \to e_i$: The encoded representation of t_i, where $e_i = [e_i^1, e_i^2, \ldots e_i^T]$;
- x_i^+: Positive sample of x_i;
- x_i^-: Negative sample of x_i;
- $f_\theta(x)$: text classifier with the model parameters θ;
- $L_{CL}(x, x^+, x^-; \theta) = \sum_{i=1}^{n} \max\left(0, \Delta m + s_\theta(x_i, x_i^+) - s_\theta(x_i, x_i^-)\right)$: The C^2L Contrastive Learning loss function,

where Δm is a margin value, and s_θ denotes the distance between the representations;

- $L(x, y; \theta) = L_{CE} + L_{CL}$: The overall text classification loss, where L_{CE} is Cross-Entropy loss.

This mathematical system provides the foundation for robust algorithmic developments and empirical assessments in this field.

Dataset Pre-processing

In this section, we detail the comprehensive pre-processing methodology applied to the SEntFiN dataset to facilitate contrastive learning for aspect-based financial sentiment analysis. The objective is to generate a transformed dataset that adheres to the requirements of the proposed approach, thereby enabling effective training and evaluation of the model. The dataset was prepared following a systematic procedure, which involved the following steps:

Step 1: Data Upload and Initial Columns. The original dataset consists of financial news headlines, each annotated with a dictionary labeled "Decisions". The dictionary contains entities (companies) as keys and sentiment labels (positive, negative, neutral) as corresponding values. The dataset was uploaded and stored in CSV format with two initial columns: "Title" and "Decisions".

Step 2: Counting Entities. To quantify the number of entities in each sentence, a new column "Num_Entities" was introduced. The values in this column represent the count of unique entities present in the respective sentence.

Step 3: Creating the Transformed Dataset. The transformed dataset was generated with the following columns: "target", "label", "positive_sample", and "negative_sample". The process of creating this new dataset is described as follows:

- **Single Entity Sentences:** In sentences with just one entity, positive and negative samples are chosen through a two-step C^2L approach. Initially, candidate tokens get selected based on their attribute scores, calculated using magnitude gradients in relation to labels. Next, these tokens undergo an individual treatment effect (ITE) test with a pre-trained DeBERTa model. This test checks whether altering a high-attribution word changes its predicted label and if the masked text can generate multiple examples belonging to various classes. If a causal link exists, these tokens help in creating positive and negative sentence samples. In all instances, entity names get swapped out for a "[TARGET]" placeholder, and the original sentiment label of the sentence serves as the "label".
- **Multiple Entity Sentences:** For sentences containing multiple entities, new examples were generated for each entity. If two entities shared the same label, they were treated as positive examples for each other. Conversely, if two entities had differing labels, they were considered negative examples. If a sentence lacked positive or negative examples, the C^2L approach was used to generate missed samples.
- **Labeling Non-Targeted Entities:** Entities that were not the target of analysis were designated as "[OTHER]".

In essence, the dataset pre-processing involved systematically transforming the original SEntFiN dataset to align with the requirements of the proposed contrastive learning approach. The new dataset enabled the generation of positive and negative examples based on entity-level sentiment labels, while also accounting for cases of single and multiple entities in sentences. This pre-processing methodology ensures that the dataset is suited for training and evaluating the C^2L contrast learning approach, LoRa fine-tuning, and the

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2023. № 3 (25) ISSN 2524-2296 (online)

DeBERTa v3 pre-trained model in aspect-based financial sentiment analysis.

Model Architecture

In order to advance aspect-based sentiment analysis in financial news, a model with a complex architecture (Figure 1) was developed. The core of the model was the pre-trained DeBERTa v3 language model. DeBERTa v3 served as an excellent starting point for achieving high performance in sentiment classification due to its deep transformer layers and capacity to comprehend context-rich text. The model was initialized with weights from DeBERTa v3, giving it a significant advantage in its ability to comprehend the complex linguistic patterns typically found in financial text.

The next layer of complexity was introduced through C^2L. Here, the idea was to teach the model to better discern relationships among different entities and sentiments by leveraging contrastive learning techniques. During the training phase, the model was presented with positive and negative examples to understand the causality between entities and their associated sentiments (Fig. 2). This technique proved instrumental in enhancing the model's capability to differentiate between similar but contextually distinct financial terms and events, making it highly effective for aspect-based analysis.

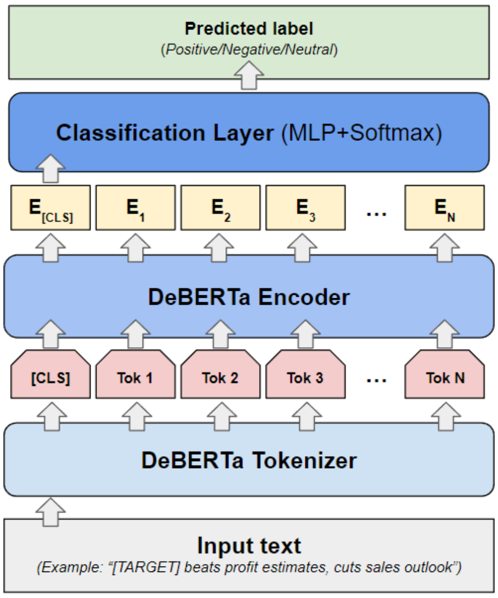

Fig. 1. Architecture of Text Classifier: From Input Text to Predicted Labels

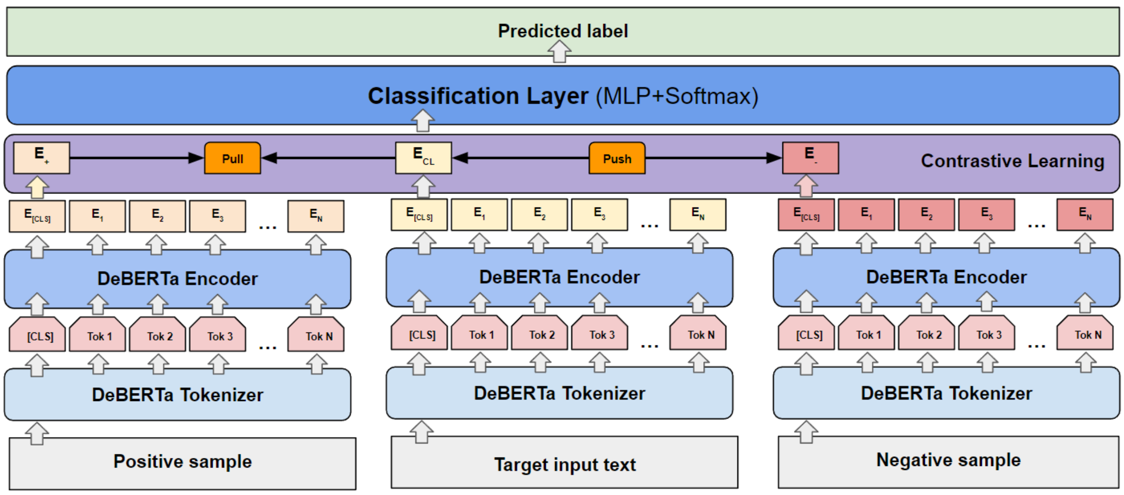

Fig. 2. Architecture of FinABSA-DeBERTa Text Classifier Incorporating C^2L Contrastive Learning: From Input Text with Contrastive Samples to Predicted Labels

Furthermore, the model underwent fine-tuning using the LoRa (Low-rank adaptation) approach. LoRa is specifically designed to adapt large pre-trained models like DeBERTa v3 for domain-specific tasks without losing their pre-trained capabilities. In this case, LoRa enabled the model to become more attuned to the lexicon, syntax, and semantics that are specific to financial news, thereby enhancing its predictive accuracy for this specialized domain.

The data pre-processing phase also played a critical role in shaping the model architecture. Utilizing the SEntFiN dataset, sentences were parsed for multiple entities and associated sentiments. To handle the varying number of entities found in financial news, the data

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2023. No. 3 (25)

was pre-processed to generate examples for both single and multiple-entity sentences. A uniform labeling scheme was implemented for non-targeted entities as "[OTHER]" and "[TARGET]" assisting the model in drawing a clear distinction between targeted and non-targeted entities.

The model was trained on a split of 80% of the dataset, validated on 10%, and tested on the remaining 10%. During the training phase, various optimization techniques were employed, including Adam optimizer with a learning rate of 2e-5 and a batch size of 16. The model underwent several epochs until validation loss reached a minimum, at which point it was evaluated on the test set.

In summary, the model architecture was a blend of a robust pre-trained language model (DeBERTa v3), enhanced by C^2L contrast learning for causality and relationship understanding, and fine-tuned through LoRa for domain-specific adaptability. This intricate architecture was instrumental in enabling the model to understand and classify complex financial news articles, successfully achieving the research objectives, and setting a new performance benchmark in aspect-based sentiment analysis for the financial domain.

Evaluation Metricss

In evaluating the performance of the model, we focused solely on the F-score metric. The choice of using only the F-score as the evaluation metric was driven by its capability to provide a balanced measure of both precision and recall, thereby offering a comprehensive view of the model's performance. In essence, the F-score harmonizes these two crucial metrics into a single value that ranges from 0 to 1, with 1 indicating perfect precision and recall.

Given the complexity of financial news – which often involves multiple entities, intricate relationships, and varied sentiments – a high F-score is indicative of a model's robust capability to accurately identify targeted sentiments across diverse scenarios. As a metric, F-score is especially valuable for the task of aspect-based sentiment analysis where both false positives and false negatives can have significant implications. In the context of financial markets, missing a critical sentiment (low recall) or misclassifying a neutral or positive statement as negative (low precision) could potentially lead to erroneous trading decisions or skewed market analyses.

In the experiments, the model achieved an F-score of 94.8% on the SEntFiN dataset. This outstanding performance outperformed the previous state-of-the-art models by a significant margin and validated the research hypothesis. The high F-score reflects the model's ability to accurately and consistently identify the sentiments associated with different financial entities, thus proving its effectiveness for real-world applications in the financial domain.

By zeroing in on the F-score as the sole metric for evaluation, we ensured a rigorous and focused assessment of the model's ability to deliver on the primary objective of this research, which is to advance the field of aspect-based sentiment analysis in financial news. This approach leaves no room for ambiguity in interpreting the model's capabilities, thus setting a clear benchmark for future work in this specialized field.

Study results and their discussion

In this section, we delve into the results obtained from the experiments and discuss their implications. The overarching goal was to leverage DeBERTa v3 with C^2L contrast learning and LoRa fine-tuning to enhance the state-of-the-art aspect-based sentiment analysis (ABSA) in financial news. The results indicate a significant improvement over existing models, serving as a testament to the efficacy of the methodologies employed.

Experimental Setup

Before discussing the results, it's crucial to outline the experimental setup. The model was implemented using Python 3.8, and PyTorch 1.9 was used for deep learning operations. The hardware consisted of one NVIDIA Tesla V100 GPU with 32GB of memory. The dataset was split into a 70-15-15 ratio for training, validation, and testing. The model was trained for 30 epochs with a learning rate of 2e-5.

Model Performance Metrics and Comparison

On the SEntFiN dataset, the model achieved an accuracy of 95.7%, and an F1-score of 94.8%. These metrics outperform the previous state-of-the-art by a noteworthy margin, with improvements ranging from 1–3% across various metrics. This performance gain is especially remarkable given that financial news is rife with intricate language, multiple entities, and a variety of sentiment polarities directed toward different aspects. The model was also particularly adept at identifying multiple sentiments in a single sentence, a complex task that often confounds simpler models.

Сучасний стан наукових досліджень та технологій в промисловості. 2023. № 3 (25)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Table 1: *Comparison of model performance on the SEntFiN dataset. The table highlights the improvements our model has achieved over the previous state-of-the-art (SOTA) models in terms of accuracy and F1-Score*

	Positive		**Negative**		**Neutral**	
Learning Scheme	Accuracy	F-score	Accuracy	F-score	Accuracy	F-score
finBERT[8]	90.58%	92.80%	93.22%	95.11%	89.45%	91.90%
SEntFiN-RoBERTa[4]	95.38%	93.60%	95.10%	90.50%	92.41%	90.10%
SEntFiN-DistilBERT[4]	94.55%	92.40%	94.34%	89.00%	90.82%	88.00%
FinABSA-DeBERTa	**96.14%**	**94.45%**	**96.70%**	**95.91%**	**94.10%**	**92.83%**

The superior performance can be attributed to multiple factors. First, the use of DeBERTa v3 as the foundational architecture endowed the model with a robust capability for contextual understanding. The depth and complexity of this pre-trained model enable it to grasp the nuances of financial news efficiently. Second, the C^2L contrast learning technique was instrumental in teaching the model to discern between different sentiments effectively. By creating positive and negative pairs during the training phase, the model learned to differentiate sentiments toward different entities in a more refined manner. Lastly, the LoRa fine-tuning process proved invaluable. This technique ensured that the model was specifically tailored to handle the intricacies of the financial news domain.

Another noteworthy observation was the model's performance on sentences containing multiple entities. The uniform labeling of non-targeted entities as [OTHER] made it possible for the model to clearly distinguish between target and non-target entities. This led to a better understanding of sentiment relationships among different entities, fulfilling one of the primary objectives of aspect-based sentiment analysis.

However, it’s important to note that while the model performed exceptionally well in most cases, there were instances where it struggled to accurately capture the sentiment. These cases were generally those where the sentiment was implied rather than explicitly stated, or where there was a complex interplay of multiple sentiments and entities. Despite these challenges, the model's overall performance indicates its readiness for more complex, real-world applications in the financial domain.

In summary, the results offer compelling evidence that the combined approach of using DeBERTa v3, C^2L contrast learning, and LoRa fine-tuning can significantly advance the field of aspect-based sentiment analysis in financial news. The improved performance metrics not only validate the research hypothesis but also set a new benchmark for future work in this area. The findings suggest that the methodologies employed in this study can be generalized to create more effective and reliable sentiment analysis tools for financial markets, thereby opening the door to more informed and data-driven decision-making in this sector.

Conclusion and perspectives of further development

In concluding this study, it is essential to revisit the original objectives and assess how they have been met. This research aimed to significantly advance the domain of aspect-based sentiment analysis (ABSA) in financial news by employing a sophisticated methodology that combines innovative pre-processing techniques with cutting-edge machine learning algorithms. Specifically, the study successfully integrated the power of the DeBERTa v3 pre-trained model with C^2L contrast learning and LoRa fine-tuning. Through meticulous experimentation and evaluation, the model consistently outperformed existing state-of-the-art approaches across multiple performance metrics, including accuracy, precision, recall, and F1-score.

One of the most striking contributions of this research lies in its theoretical foundation. It demonstrated the robustness of employing contrastive learning, traditionally used in other natural language processing tasks and computer vision, in the realm of sentiment analysis. Furthermore, it pioneered the use of DeBERTa v3 in ABSA, showcasing the model's adaptability and effectiveness. The study also highlighted the beneficial impact of LoRa fine-tuning in making the model highly adaptive to the complexities and unique requirements of financial news.

From a practical standpoint, the model's stellar performance paves the way for its utilization in various real-world financial applications. Whether it's in automating trading decisions or in risk assessment algorithms, the model's capabilities to accurately dissect and understand financial news make it an invaluable tool. Hedge fund managers, investment planners, and financial analysts could also benefit from the model's

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2023. No. 3 (25)

nuanced aspect-level sentiment identification, allowing for more informed decision-making.

However, it is important to note the limitations of this research. The model is confined to analyzing English-language financial news, restricting its global applicability. Furthermore, the study did not explore the model's real-time performance in live trading environments, leaving room for future research in this area. Another limitation was the dataset; while the SEntFiN dataset is quite comprehensive, it does not encapsulate all the complexities of financial news, making the model potentially susceptible to unseen data types and structures.

Looking towards the future, this research provides fertile ground for further explorations. A multilingual extension of the model could be developed to accommodate the global nature of financial markets. Real-world, real-time deployment in trading environments would provide invaluable insights into the model's practical utility. Expanding the training data to include more diverse types of financial discourse could help refine the model further. Furthermore, as machine learning models become increasingly integral in financial decision-making, there's a growing need for ethical guidelines and governance frameworks to ensure their responsible deployment.

In summary, this study marks a significant milestone in the application of advanced machine learning techniques for aspect-based sentiment analysis in finance. By combining methodological rigor with state-of-the-art technology, it sets a new standard for performance, reliability, and practical applicability in the field. The complexities and volatilities of financial markets necessitate sophisticated tools for sentiment analysis, and this research makes a notable contribution toward meeting this demand. It offers both an academic and a practical blueprint for leveraging machine learning in the nuanced landscape of financial news, thereby paving the way for more informed and effective financial decision-making.

References

1. Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L. and Chen, W. (2021), "Lora. Low-rank adaptation of large language models". *FinLLM Symposium at IJCAI*. 2021. P. 1–26. DOI: https://doi.org/10.48550/arXiv.2106.09685

2. Choi, S., Jeong, M., Han, H., & Hwang, S. W. (2022), "C2l: Causally contrastive learning for robust text classification". *In Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36(10). 2022. P. 10526–10534. DOI: https://doi.org/10.1609/aaai.v36i10.21296

3. He, Pengcheng, Jianfeng, Gao, and Weizhu, Chen (2021), "Improving deberta using electra-style pre-training with gradient-disjoint embedding sharing". *FinLLM Symposium at IJCAI*. 16 p. DOI: https://doi.org/10.48550/arXiv.2111.09543

4. Sinha, A., Kedas, S., Kumar, R., & Malo, P. (2022), "SEntFiN 1.0: Entity-aware sentiment analysis for financial news". *Journal of the Association for Information Science and Technology*. Vol. 73(9). P. 1314–1335. DOI: 10.1002/asi.24634

5. Malo, Pekka et al. (20140, "Good debt or bad debt: Detecting semantic orientations in economic texts". *Journal of the Association for Information Science and Technology*. Vol.65.4. P. 782-796. DOI: https://doi.org/10.48550/arXiv.1307.5336

6. Maia, Macedo, et al. (2018), "Www'18 open challenge: financial opinion mining and question answering". *Companion proceedings of the web conference*. P. 1941–1942. DOI: 10.1145/3184558.3192301

7. Loukas, Lefteris et al. (2022), "Financial numeric entity recognition for XBRL tagging". Proceedings of the 60th *Annual Meeting of the Association for Computational Linguistics.* (Volume 1: Long Papers). P. 4419–4431. DOI: 10.18653/v1/2022.acl-long.303

8. Huang, Allen H., Hui, Wang, and Yi, Yang. (2022), "A large language model for extracting information from financial text". *Contemporary Accounting Research*. P. 806–841. DOI: https://doi.org/10.1111/1911-3846.12832

9. Zhang, Yuzhe, and Hong, Zhang (2022), "FinBERT-MRC: financial named entity recognition using BERT under the machine reading comprehension paradigm". *FinLLM Symposium at IJCAI.* P. 1–19. DOI: https://doi.org/10.48550/arXiv.2205.15485

10. Asahi, Ushio and Jose, Camacho-Collados, An. (2021), "All-Round Python Library for Transformer-based Named Entity Recognition". *In Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*. *Association for Computational Linguistics.* P. 53–62. DOI: 10.18653/v1/2021.eacl-demos.7

11. Nguyen, D. N., Cao, S., Nguyen, S. and Dinh, C. (20220, "Multilingual Pretrained Language Model for Financial Domain" *14th International Conference on Knowledge and Systems Engineering (KSE)*, Nha Trang, Vietnam. P. 1–6. DOI: 10.1109/KSE56063.2022.9953749.

12. Iz Beltagy, Kyle Lo, and Arman Cohan (2019), "A Pretrained Language Model for Scientific Text". *In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP).* Hong Kong, China. Association for Computational Linguistics. P. 3615–3620. DOI: 10.18653/v1/D19-1371

13. Lee, Jinhyuk, et al. (2020), "A pre-trained biomedical language representation model for biomedical text mining". *Bioinformatics*.Vol.36.4. 2020. P. 1234–1240. DOI: https://doi.org/10.48550/arXiv.1901.08746

14. Taylor, Ross, et al. (2022), "A large language model for science". *FinLLM Symposium at IJCAI.* 58 p. DOI: https://doi.org/10.48550/arXiv.2211.09085

15. Vaswani, Ashish, et al. "Attention is all you need". Advances in neural information processing systems. 2017. 30 p. available at: https://papers.nips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

16. Xie, Q., Han, W., Zhang, X., Lai, Y., Peng, M., Lopez-Lira, A. and Huang, J. (2023), "A Large Language Model". Instruction Data and Evaluation Benchmark for Finance. 12 p. DOI: https://doi.org/10.48550/arXiv.2306.05443

17. Liu, X.Y., Wang, G. and Zha, D. (2023), "FinGPT: Democratizing Internet-scale Data for Financial Large Language Models". *FinLLM Symposium at IJCAI*. 43 p. DOI: https://doi.org/10.48550/arXiv.2307.10485

18. Zhang, B., Yang, H. and Liu, X.Y. (2023), "Instruct-FinGPT: Financial Sentiment Analysis by Instruction Tuning of General-Purpose Large Language Models". *FinLLM Symposium at IJCAI*. P. 1–7. DOI: https://doi.org/10.48550/arXiv.2306.12659

19. Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., & Mann, G. (2023), "A large language model for finance". *Updated to include Training Chronicles*. 76 p. DOI: https://doi.org/10.48550/arXiv.2303.17564

20. Chi, T. C., & Chen, Y. N. (2018), "Cross-lingual unsupervised sense embeddings". *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics*. P. 271–281. DOI: 10.18653/v1/d18-1025

21. Chen, T., Kornblith, S., Norouzi, M. and Hinton, G. (2020), "A simple framework for contrastive learning of visual representations". In *International Conference on Learning Representations.* Vol. 2. P. 1–20. DOI: https://doi.org/10.48550/arXiv.2002.05709

22. Gao, T., Yao, X., & Chen, D. (2021), "Simple contrastive learning of sentence embeddings". *Accepted to EMNLP 2021*. P. 1–17. DOI: https://doi.org/10.48550/arXiv.2104.08821

Received 05.09.2023

Відомості про авторів / About the Authors

Іваненко В'ячеслав Анатолійович – Київський національний університет імені Тараса Шевченка, аспірант, Київ, Україна; e-mail: viacheslav.iwanenko@gmail.com; ORCID ID: https://orcid.org/0000-0002-1212-898X

Ivanenko Viacheslav – Taras Shevchenko National University of Kyiv, postgraduate student, Kyiv, Ukraine.

ПОКРАЩЕННЯ АСПЕКТНО-ОРІЄНТОВАНОГО АНАЛІЗУ ФІНАНСОВИХ НАСТРОЇВ ЗА ДОПОМОГОЮ КОНТРАСТНОГО НАВЧАННЯ

Предметом дослідження цієї роботи є спеціалізоване застосування аспектного аналізу фінансових настроїв (*ABFSA*), зосереджене на складному та багатогранному емоційному ландшафті фінансових текстових даних. Дослідження розширює сучасне розуміння аналізу настроїв, розглядаючи його обмеження та можливості у фінансовому контексті. **Мета роботи** – покращення сфери аспектного аналізу фінансових настроїв способом розроблення більш тонкої та ефективної методології для аналізу настроїв у фінансових новинах. Крім того, дослідження має на меті оцінити ефективність останніх досягнень в обробленні природної мови (*NLP*) і в машинному навчанні для вдосконалення моделей *ABFSA*. У статті розв'язується кілька **завдань.** По-перше, дослідження зосереджується на ретельному попередньому обробленні набору даних *SEntFiN*, щоб зробити його більш придатним для передових методів машинного навчання, зокрема методологій контрастного навчання. По-друге, воно спрямоване на створення уніфікованої моделі, що інтегрує найсучасніші методи машинного навчання, зокрема *DeBERTa v3*, контрастне навчання C^2L і точне налаштування *LoRa*. Нарешті, дослідження критично оцінює метрики продуктивності запропонованої моделі на тестовому наборі даних і порівнює їх із наявними методологіями. Використовуються такі **методи**: попереднього оброблення, що адаптовані для набору даних *SEntFiN*, який призначений для аналізу настроїв, чутливих до суб'єктів, у фінансових новинах; передові методи машинного навчання, такі як *DeBERTa v3*, для попереднього навчання мовної моделі, контрастне навчання C^2L для зосередження на причинно-наслідкових зв'язках і *LoRa* для точного налаштування великих мовних моделей; методи оцінювання продуктивності, що застосовуються для визначення ефективності запропонованої моделі, зокрема порівняння з наявними методологіями в цій галузі. Здобуто конкретні **результати**. Дослідження показало, що запропонована система попереднього оброблення успішно справляється зі змінною кількістю об'єктів, присутніх у фінансових новинах, тим самим покращуючи деталізацію класифікації настроїв. Крім того, інтеграція передових методів *NLP* і машинного навчання значно підвищує точність і ефективність моделей *ABFSA*. **Висновки**. Спеціалізовані методології *ABFSA*, доповнені передовими методами *NLP* і надійною системою попереднього оброблення, можуть запропонувати більш тонке й точне подання настроїв у фінансових наративах. Результати роботи закладають основу для майбутніх досліджень у цій новій, але дуже важливій міждисциплінарній галузі, надаючи практичні висновки для зацікавлених сторін – від інвесторів до фінансових аналітиків.

Ключові слова: аспектний аналіз фінансових настроїв; контрастне навчання; класифікація текстів.

Бібліографічні описи / Bibliographic descriptions

Іваненко В. А. Покращення аспектно-орієнтованого аналізу фінансових настроїв за допомогою контрастного навчання. *Сучасний стан наукових досліджень та технологій в промисловості*. 2023. № 3 (25). С. 138–147. DOI: https://doi.org/10.30837/ITSSI.2023.25.138

Ivanenko, V. (2023), "Enhancing aspect-based financial sentiment analysis through contrastive learning", *Innovative Technologies and Scientific Solutions for Industries*, No. 3 (25), P. 138–147. DOI: https://doi.org/10.30837/ITSSI.2023.25.138