AUTHOR GUIDELINES FOR JCC SUBMISSION
SUBTITLE

AUTHOR NAME(S)

[1]*Institute of Information Technology, Vietnam Academy of Science and Technology;*
anonymous@ioit.ac.vn
NEXT AUTHOR AFFILIATION AND EMAIL ADDRESS

Abstract. This paper presents a practical, end-to-end system designed for the banking sector, enabling users to efficiently gather sentiment insights related to various aspects of banks from online sources. The system's real-world applicability lies in its ability to provide actionable sentiment analysis, which can be leveraged by banks for strategic decision-making. On the scientific front, we meticulously constructed a sentiment dataset specific to the banking industry, which was subsequently employed to evaluate the performance of various language models. This comprehensive evaluation not only demonstrates the efficacy of the dataset but also contributes valuable insights into the application of language models for sentiment analysis in the financial domain.

Keywords. Sentiment analysis, Online Sentiment Insights, Text Mining, End-to-End Financial System, Natural Language Processing (NLP), Language Models.

1. INTRODUCTION

In the era of digital transformation, the banking sector is increasingly reliant on advanced technologies to understand customer sentiment and improve service delivery. This paper introduces an end-to-end system specifically designed for the banking industry, which offers a comprehensive solution for sentiment analysis. The system's practical utility lies in its ability to scrape data from various online sources, analyze sentiment, and visually represent the proportion of positive, negative, and neutral sentiments. This visualization empowers banks to gain actionable insights and make informed decisions based on customer feedback.

A core component of the system is a meticulously crafted sentiment dataset tailored for the banking domain. This dataset[1] was developed using the scraped data and serves as the foundation for the sentiment analysis module. To ensure the robustness and reliability of the sentiment analysis, we evaluated the dataset using a range of language models. Specifically, we conducted inference using large language models (LLMs) such as LLama 3.1[4] and Vistral[8], and compared their performance with fine-tuned smaller models like PhoBert[9], ViT5[11] and BartPho[13].

Our findings suggest that the inherent knowledge embedded within LLMs does not significantly enhance their performance in this specific application. In contrast, the fine-tuned smaller models, when trained on the domain-specific dataset, demonstrated superior results. This observation challenges the assumption that larger models are inherently more capable and highlights the importance of domain-specific training for sentiment analysis in the financial sector.

---

[1]Our dataset is available at `https://huggingface.co/datasets/iaiuet/banking_sentiment_vietnamese`

The contributions of this paper are threefold: first, we present a highly practical system that can be readily applied in the banking industry to derive sentiment insights; second, we provide a new, domain-specific sentiment dataset for the banking sector; and third, we offer a thorough evaluation of different language models, contributing valuable knowledge to the field of natural language processing (NLP) and sentiment analysis.

2.   RELATED WORK

All text must be in a one-column format. The main title (on the first page) should begin 1.0 inch (2.54 cm) from the top edge of the page. The second and following pages should begin 1.0 inch (2.54 cm) from the top edge. On all pages, the bottom margin should be 1-1/8 inches (2.86 cm) from the bottom edge of the page for 8.5×11-inch paper; for A4 paper, approximately 1-5/8 inches (4.13 cm) from the bottom edge of the page.

2.1.   Sentiment Analysis in Banking

All printed material, including text, illustrations, and charts, must be kept within a print area 6-7/8 inches (17.5 cm) wide by 8-7/8 inches (22.54 cm) high.

2.2.   Language Models for Sentiment Analysis

Wherever Times is specified, Times Roman may also be used. If neither is available on your word processor, please use the font closest in appearance to Times to which you have access.

2.3.   Dataset Development for Sentiment Analysis

3.   System Design and Implementation

Our proposed system for analyzing public sentiment towards banks operates through a structured pipeline that integrates data scraping, processing, clustering, sentiment analysis, and visualization. This end-to-end approach ensures that data is systematically collected from relevant online sources, processed and stored efficiently, and then analyzed to identify sentiment trends. The final results are presented in an intuitive visual format, allowing users to easily interpret the proportions of positive, negative, and neutral sentiments. This comprehensive pipeline enhances the accuracy and practicality of sentiment analysis, making it a valuable tool for stakeholders in the banking sector.

3.1.   Data Scraping

In the data scraping phase of our system, we leverage the Scrapy framework[7] to extract information from various online news sources. We aim to gather data from nine distinct websites: *cafef, dantri, kinhtedothi, nhadautu, vietnambiz, vietnamnet, vietstock, vneconomy,* and *vnexpress.*

Each website provides valuable insights into public sentiment towards banking institutions through news articles and reports. However, the available data on these websites may vary in terms of publication dates. Therefore, we establish specific timeframes for data collection from each website to ensure comprehensive coverage.

The designated timeframes for data scraping from each website are as follows:

- **cafef[1]**: Up to January 1, 2000.

- **dantri[2]**: Up to approximately six months prior to the current date.

- **kinhtedothi[6]**: Up to May 16, 2022.

- **nhadautu[10]**: Up to March 25, 2017.

- **vietnambiz[14]**: Up to August 22, 2016.

- **vietnamnet[15]**: Up to September 23, 2010.

- **vietstock[16]**: Up to July 2, 2003.

- **vneconomy[17]**: Up to July 12, 2023.

- **vnexpress[18]**: Up to approximately six months prior to the current date.

By adhering to these timeframes, we aim to collect a diverse range of news articles spanning different periods, enabling comprehensive sentiment analysis and visualization in subsequent stages of our system.

3.2. Processing and Storing

3.2.1. Database Structure

Our database is designed following the Entity-Relationship (ER) model[12], depicted in Figure 1. The ER model establishes relationships between entities, enhancing data integrity and facilitating complex queries. The database comprises four main tables, each serving a specific purpose:

1. **SAC THAI**: This table stores paragraphs of text data along with their corresponding sentiment states and associations with banking topics. It includes the following columns:

    - **id_paragraph**: Primary key identifier for paragraphs

    - **text_paragraph**: Text content of the paragraph

    - **chu_de**: Topic of the paragraph

    - **ngan_hang**: Foreign key referencing the 'NGAN HANG' table

    - **sac_thai**: the sentiment states of the paragraph

    - **id_ban_tin**: Foreign key referencing the 'BAN TIN' table

    - **ngay**: Date of publication of the articles

2. **NGAN HANG**: This table stores information about banks, including their names and aliases. It includes the following columns:

    - **id_ngan_hang**: Primary key identifier for banks

    - **Ten**: Name of the bank

    - **Ten_viettat**: Abbreviated name of the bank

- **Ten_khac**: Other aliases of the bank

3. **BAN TIN**: This table stores news articles related to economic topic. It includes the following columns:

    - **id_ban_tin**: Primary key identifier for news articles

    - **Nguon**: Source of the news article

    - **Ngay**: Date of publication

    - **text**: Text content of the news article

    - **url**: URL of the news article

    - **url_hash**: Hash value of the URL

4. **CHU DE**: This table stores information about topics, including their names and synonyms. It includes the following columns:

    - **id_chu_de**: Primary key identifier for topics

    - **Ten**: Name of the topic

    - **Dong_nghia**: Synonyms of the topic

    - **Tich_cuc**: Keywords for positive sentiments about the topic

    - **Tieu_cuc**: Keywords for negative sentiments about the topic

    - **id_chu_de_con**: Identifier for subtopics

3.2.2.   Data Classification

Upon retrieval, the text data undergoes NLP-based classification to identify its relevance to economic topics. The module use standard binary-classification pipeline: feature extraction with TF-IDF (Term Frequency-Inverse Document Frequency), and classify using SVM (Support Vector Machine). This classification step ensures that only economically significant data is further processed and analyzed. The performance of this classification model was evaluated using a test set, and the results are shown in Table 1. Subsequently, the economically relevant data is stored in the 'BAN TIN' table for further analysis.

3.3.   Clustering and Segmenting

Within the Clustering and Segmenting phase, our system employs sophisticated algorithms to segment news articles retrieved from the 'BAN TIN' table within the database. The objective is to extract segments discussing specific banks within a given topic, enhancing the precision and relevance of subsequent analysis. The workflow of this module unfolds as follows:
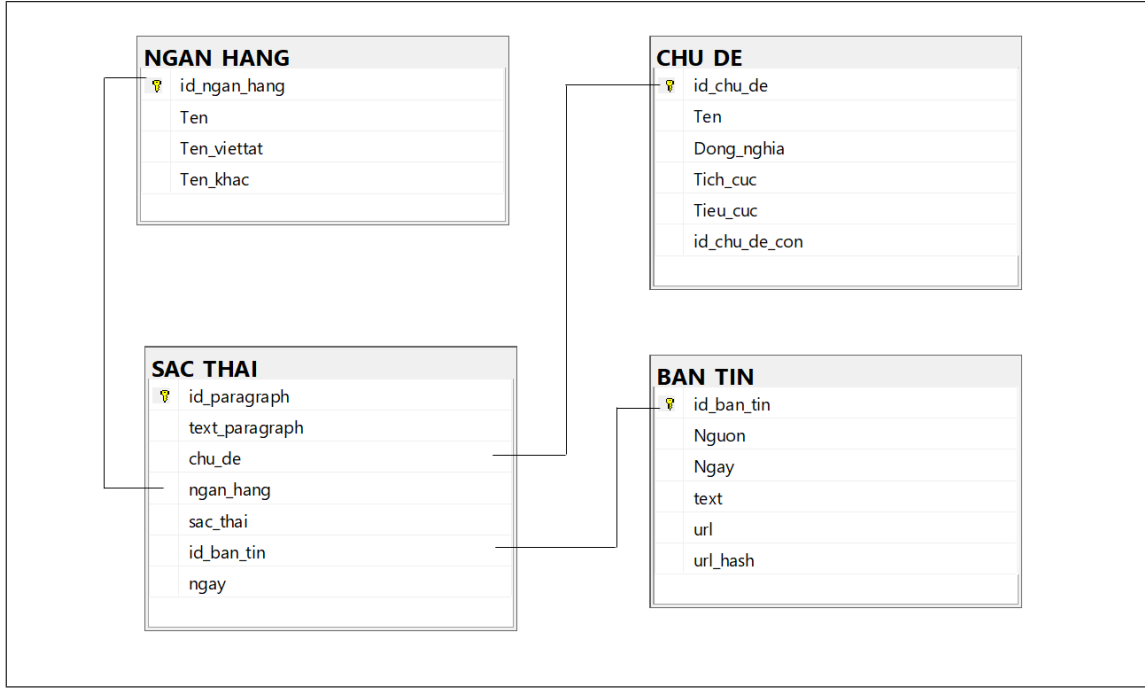
Figure 1: Database diagram

	precision	recall	f1-score	support
Not Economic-Related	0.76	0.80	0.78	188
Economic-Related	0.81	0.78	0.80	212
accuracy			0.79	400
macro avg	0.79	0.79	0.79	400
weighted avg	0.79	0.79	0.79	400

Table 1: Performance of the SVM Classifier with TF-IDF.

3.3.1. Text Segmentation

The module retrieves news articles from the 'BAN TIN' table in the database, then initiates the process of text segmentation. This involves breaking down the article into smaller segments, each representing a coherent and meaningful unit of information. By segmenting the text, the module ensures that fine-grained analysis can be conducted on specific aspects of the content

3.3.2. Topic Identification

Concurrently, the module discerns the overarching topic(s) discussed within each article by leveraging keyword extraction techniques. This process involves extracting relevant keywords from the text, utilizing a predefined list of topics stored in the 'CHU DE' table within the database. By matching keywords extracted from the article with those in the topic repository, the module identifies the primary themes and subtopics addressed in the text. This approach enables efficient topic identification

without relying on complex topic modeling algorithms, ensuring timely and accurate classification of articles based on predefined topic categories

3.3.3.   Bank Entity Recognition

In this process, the module employs name-based recognition techniques, drawing upon data from the 'NGAN HANG' table stored in the database. By extracting bank names, abbreviations, and other aliases from the database, the module accurately identifies mentions of specific banks within each segmented article. This approach ensures precise recognition of bank entities in the text, facilitating comprehensive analysis of banking-related discussions without the need for complex models

3.3.4.   Segment Filtering

Subsequently, the module filters the segmented text to retain only those segments that pertain to specific banks within the identified topic(s). This filtering process ensures that subsequent analysis is focused on relevant discussions related to the banking sector. The retained segments are then passed on to the next module for sentiment analysis, as discussed in subsequent sections. This sequential workflow ensures that only pertinent segments contribute to the sentiment analysis process, enhancing the accuracy and relevance of the insights generated.

3.4.   Sentiment Analysis

Our sentiment analysis module leverages two advanced models: Vistral and PhoBERT, known for their exceptional accuracy and robustness in processing textual sentiment, particularly in Vietnamese.

- **Vistral**: Vistral is a multi-turn conversational large language model for Vietnamese. Vistral is extended from the Mistral 7B model using diverse data for continual pre-training and instruction tuning.

- **PhoBERT**: PhoBERT is built upon the powerful BERT (Bidirectional Encoder Representations from Transformers) framework, is the first public large-scale monolingual language models pre-trained for Vietnamese. It boasts remarkable capabilities in multiple Vietnamese-specific NLP tasks including Part-of-speech tagging, Dependency parsing, Named-entity recognition and Natural language inference.

Details on the performance of these models are provided in the following sections.

After sentiment analysis, the processed paragraphs are stored in the 'SAC THAI' table along with their corresponding sentiments.

By integrating Vistral and PhoBERT into our sentiment analysis module and storing the results in the database, our system ensures accurate sentiment classification, enabling comprehensive analysis of textual data within the financial domain.

3.5.   Visualization

Our system employs sophisticated data visualization techniques to transform processed sentiment data into interactive and insightful charts and graphs. The visualization process begins by extracting relevant data from the 'SAC THAI' table in the database. This data is then filtered based on the desired criteria, such as specific banks, topics, and time periods. Once filtered, the sentiment scores

Figure 2: Distribution of sentiment (positive, negative, neutral) across various banks during a specific quarter, highlighting the public perception of each bank's performance.

(positive, negative, and neutral) are aggregated for each bank, topic, and time period to facilitate comprehensive trend analysis.

To create these visualizations, we utilize the advanced data visualization library Matplotlib[5]. Our system supports various types of charts, including bar charts and pie charts, Our system supports many different chart types, including bar charts and pie charts, in this report, stacked bar charts will be taken as an example. This chart effectively illustrates the distribution of sentiment across different categories. As shown in Figure 2, the stacked bar chart includes the following elements: the X-axis represents categories such as banks, topics, or time periods, while the Y-axis shows the percentage of sentiment (positive, negative, neutral). Each stacked bar represents the proportion of each sentiment category, with a legend indicating the different sentiment types. Tooltips are incorporated to provide detailed information about each data point, enhancing user understanding and engagement.

The system is designed to be highly interactive, allowing users to filter the data by selecting specific banks, topics, or time periods. Users can also drill down to view sentiment distribution at a more granular level, such as individual articles or specific dates. Hover effects are implemented to display additional information when users hover over data points, offering deeper insights into the sentiment data.

Customization options are available to users, enabling them to adjust the appearance of the

	train	test
Bank Capital	846	168
Credit Growth - Money Supply	1198	212
Role of Regulatory Tools	292	67
Savings Deposits	793	142
Liquidity	949	196
Banking Services	1495	311
State Bank Policy Implementation	255	60
Management Quality	390	72
Non-Performing Loans	1381	286
Digitalization of Banking Operations	859	146
Loan Products	554	124
Business Efficiency	635	134
System Safety	353	82

Table 2: Distribution of data across different topics in the dataset.

charts, including colors, fonts, and styles. Additionally, the system provides options to export the charts in various formats, such as PNG, JPEG, or PDF, for reporting and presentation purposes. Users can also export the underlying data in Excel (xlsx) format, allowing for further analysis using tools like Excel or PowerBI.

4.   DATASET DEVELOPMENT

To effectively analyze sentiment in the banking sector, we developed a comprehensive dataset using the aforementioned system. The dataset consists of 10,000 training samples and 2,000 test samples, all collected through our end-to-end system that scrapes relevant online data. The data was carefully labeled by human annotators, with each sample categorized into one of three sentiment labels: positive, negative or neutral.

The dataset covers a wide range of topics, spanning 13 key aspects of the banking industry. These topics include: Bank Capital, Credit Growth - Money Supply, Role of Regulatory Tools, Savings Deposits, Liquidity, Banking Services, State Bank Policy Implementation, Management Quality, Non-Performing Loans (NPLs), Digitalization of Banking Operations, Loan Products, Business Efficiency and System Safety.

Given the diverse nature of these topics, the dataset provides a well-rounded view of sentiment across different facets of banking. The data is entirely in Vietnamese, reflecting the language used in the source material. The distribution of data across these 13 topics is summarized in the table 2.

This dataset is meticulously designed for sentiment analysis within the banking sector, offering deep insights into public opinion across a broad spectrum of banking topics. By covering 13 distinct aspects of the banking industry, the dataset provides a highly practical and comprehensive foundation for training and evaluating sentiment analysis models. Its extensive coverage ensures that the models are well-aligned with the diverse and nuanced realities of the financial domain, making the dataset exceptionally relevant for real-world applications. Some examples of the dataset are shown in Appendix 1.

5. EXPERIMENTAL SETUP

To evaluate the effectiveness of sentiment analysis in the banking domain, we conducted a series of experiments using both large language models (LLMs) and smaller, fine-tuned models. The LLMs, specifically LLama 3.1 and Vistral, were used exclusively for inference without any fine-tuning. We explored their performance across a range of prompts designed to reflect various aspects of sentiment analysis.

The inference was performed using the following types of prompts:

- **Zero-shot without External Knowledge**: Prompts were provided without any additional context or definitions related to what constitutes positive, negative, or neutral sentiment.

- **Zero-shot with External Knowledge**: Prompts included additional context, explicitly defining what constitutes positive, negative, and neutral sentiment to guide the models.

- **Few-shot without External Knowledge**: Prompts included 1 to 5 examples but did not provide additional context or definitions.

- **Few-shot with External Knowledge**: Prompts included 1 to 5 examples along with definitions of sentiment categories, offering more guidance to the models.

For the few-shot prompts, we retrieved relevant documents from the training set that belonged to the same topic using a vector database implemented with Faiss[3].

In contrast, smaller models such as BARTPho, PhoBERT, and ViT5 were fine-tuned on our domain-specific dataset to optimize their performance in sentiment analysis. These models underwent training on the dataset with the goal of capturing the nuanced sentiment expressed across the 13 topics within the banking sector.

All models, both LLMs and smaller models, were trained and inferred using a V100 16GB GPU. This setup allowed us to compare the out-of-the-box performance of large language models with the fine-tuned performance of smaller models, providing insights into the impact of model size, training, and the inclusion of external knowledge on sentiment analysis accuracy.

6. RESULTS AND DISCUSSION

The performance of the sentiment analysis models was evaluated using accuracy as the primary metric, alongside considerations of GPU usage and inference time. Our experiments reveal several noteworthy trends, as detailed below.

First, the fine-tuned smaller models BARTPho[2], PhoBERT[3], and ViT5[4] consistently outperformed the larger language models (LLMs) such as LLama 3.1 and Vistral when it came to accuracy. Despite their smaller size, these models not only achieved higher accuracy but also required significantly less computational resources and ran much faster, demonstrating the efficiency and effectiveness of fine-tuning in specialized domains like banking.

For the LLMs, accuracy improved as more shots were provided during inference. The highest accuracy was achieved by the quantized version of LLama 3.1 using a 5-shot prompt without external knowledge, reaching an accuracy of 0.6175. However, this came at the cost of higher GPU usage and

---

[2]Our model is available at `https://huggingface.co/iaiuet/bartpho-banking-sentiment`

[3]Our model is available at `https://huggingface.co/iaiuet/PhoBERT-base-sentiment`

[4]Our model is available at `https://huggingface.co/iaiuet/vit5-base_sentiment`

	Prompt Type	**Acc.**	**Inference Time (minutes)**	**GPU Usage (GB)**
LLama 3.1 (quant.)	0-shot, No external knowledge	0.551	24	7.816
LLama 3.1 (quant.)	1-shot, No external knowledge	0.5265	30	8.148
LLama 3.1 (quant.)	2-shot, No external knowledge	0.567	32.6	8.424
LLama 3.1 (quant.)	3-shot, No external knowledge	0.587	35.5	8.844
LLama 3.1 (quant.)	4-shot, No external knowledge	0.595	36.7	9.158
LLama 3.1 (quant.)	5-shot, No external knowledge	**0.6175**	39.5	9.414
LLama 3.1 (quant.)	0-shot, With external knowledge	0.4635	25.5	7.838
LLama 3.1 (quant.)	1-shot, With external knowledge	0.4985	31.6	8.172
LLama 3.1 (quant.)	2-shot, With external knowledge	0.525	37	8.41
LLama 3.1 (quant.)	3-shot, With external knowledge	0.555	37.9	8.902
LLama 3.1 (quant.)	4-shot, With external knowledge	0.5725	40.7	9.256
LLama 3.1 (quant.)	5-shot, With external knowledge	0.5665	41	9.492
Vistral (quant.)	0-shot, No external knowledge	0.5375	28.7	6.738
Vistral (quant.)	1-shot, No external knowledge	0.354	45.8	6.816
Vistral (quant.)	2-shot, No external knowledge	0.354	57	9.053
Vistral (quant.)	3-shot, No external knowledge	0.4	68.8	9.443
Vistral (quant.)	4-shot, No external knowledge	0.4335	81.2	8.33
Vistral (quant.)	5-shot, No external knowledge	0.4845	97	11.97
Vistral (quant.)	0-shot, With external knowledge	0.529	39	7.287
Vistral (quant.)	1-shot, With external knowledge	0.4105	66	7.104
Vistral (quant.)	2-shot, With external knowledge	0.3715	78	10.484
Vistral (quant.)	3-shot, With external knowledge	0.4095	78	9.67
Vistral (quant.)	4-shot, With external knowledge	0.435	92	11.184
Vistral (quant.)	5-shot, With external knowledge	0.48	105.8	12.268
PhoBERT (fine-tuned)		0.689	0.50	0.214
BARTPho (fine-tuned)		0.678	0.77	0.889
ViT5 (fine-tuned)		0.682	1.45	1.482

Table 3: Performance comparison of sentiment analysis models in terms of accuracy, GPU usage, and inference time across different prompt configurations.

longer inference times compared to the smaller models. In term of large language models, LLama 3.1 generally showed higher accuracy, used less GPU memory, and completed inference faster compared to Vistral. This indicates that LLama 3.1 is more efficient and better suited for sentiment analysis in this context, particularly when GPU resources and inference time are constrained.

An unexpected finding was the decrease in accuracy when external knowledge was introduced in the prompts. Contrary to expectations, this additional context led to lower performance across all LLM configurations. This suggests that the inclusion of external knowledge may introduce noise or confusion rather than improving model understanding, a phenomenon that warrants further investigation.

From these results, it is evident that the inherent knowledge of LLMs does not always translate into better performance. This outcome indicates that LLMs' pre-existing knowledge may not be entirely effective or even beneficial in specific sentiment analysis tasks, particularly in the nuanced context of banking.

Table 3 summarizes the accuracy, GPU usage, and inference time for different models and prompt configurations, all conducted on the V100 16GB GPU.

In summary, while LLMs show promise with few-shot learning, our results underscore the advantages of fine-tuning smaller models, particularly in terms of accuracy, resource efficiency, and speed. The unexpected drop in accuracy with external knowledge prompts indicates a need for further research to better understand and optimize the use of LLMs in sentiment analysis.

7. CONCLUSION AND FUTURE WORK

7.1. Summary of Contributions

In this paper, we presented a comprehensive end-to-end system for sentiment analysis within the banking sector, leveraging a combination of large language models (LLMs) and smaller, fine-tuned models. Our system's ability to scrape data, perform sentiment analysis, and visualize results makes it a valuable tool for understanding public sentiment on a wide range of banking topics.

The newly developed sentiment analysis dataset, specifically tailored for the banking domain and covering 13 distinct topics, serves as a critical resource for training and evaluating sentiment models. Our experiments demonstrated that smaller, fine-tuned models like BARTPho, PhoBERT, and ViT5 outperform larger LLMs such as LLama 3.1 and Vistral in terms of accuracy, resource efficiency, and inference speed. This outcome highlights the practical advantages of domain-specific fine-tuning over relying solely on the extensive pre-existing knowledge within LLMs.

Interestingly, our results also revealed that providing LLMs with external knowledge did not enhance their performance as expected, sometimes even leading to a decline in accuracy. This suggests that LLMs' inherent knowledge, while extensive, may not always align with the specific nuances of domain-specific tasks like sentiment analysis in banking.

In conclusion, while LLMs offer powerful capabilities, our findings suggest that for specialized tasks within constrained environments, smaller, fine-tuned models may offer more reliable and efficient performance. Future work could explore the factors contributing to the observed inefficacy of external knowledge in LLMs and seek to optimize their application in specialized domains.

7.2. Future Work

Future work will focus on improving the data classification process within the Processing and Storing module. This improvement will involve refining algorithms and methodologies to better categorize

economic content, thereby increasing the relevance and accuracy of sentiment analysis.

APPENDIX

1.   Example of the dataset

Table 4 illustrate the structure and content of our sentiment analysis dataset, which is specifically tailored for the banking sector. The dataset consists of sentences labeled as positive, negative, or neutral, each covering various banking topics such as bank capital, credit growth, digitalization of banking operations...

These examples reflect the diversity and specificity of the topics covered in the dataset, enabling accurate sentiment analysis across a wide range of banking-related issues.

ACKNOWLEDGMENT

REFERENCES

[1] "Trang thông tin điện tử tổng hợp," https://cafef.vn/.

[2] "Tin tức việt nam và quốc tế nóng, nhanh, cập nhật 24h | báo dân trí," https://dantri.com.vn/.

[3] M. Douze, A. Guzhva, C. Deng, J. Johnson, G. Szilvasy, P.-E. Mazaré, M. Lomeli, L. Hosseini, and H. Jégou, "The faiss library," 2024.

[4] A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Yang, A. Fan, A. Goyal, A. Hartshorn, A. Yang, A. Mitra, A. Sravankumar, A. Korenev, A. Hinsvark, A. Rao, A. Zhang, A. Rodriguez, A. Gregerson, A. Spataru, B. Roziere, B. Biron, B. Tang, B. Chern, C. Caucheteux, C. Nayak, C. Bi, C. Marra, C. McConnell, C. Keller, C. Touret, C. Wu, C. Wong, C. C. Ferrer, C. Nikolaidis, D. Allonsius, D. Song, D. Pintz, D. Livshits, D. Esiobu, D. Choudhary, D. Mahajan, D. Garcia-Olano, D. Perino, D. Hupkes, E. Lakomkin, E. AlBadawy, E. Lobanova, E. Dinan, E. M. Smith, F. Radenovic, F. Zhang, G. Synnaeve, G. Lee, G. L. Anderson, G. Nail, G. Mialon, G. Pang, G. Cucurell, H. Nguyen, H. Korevaar, H. Xu, H. Touvron, I. Zarov, I. A. Ibarra, I. Kloumann, I. Misra, I. Evtimov, J. Copet, J. Lee, J. Geffert, J. Vranes, J. Park, J. Mahadeokar, J. Shah, J. van der Linde, J. Billock, J. Hong, J. Lee, J. Fu, J. Chi, J. Huang, J. Liu, J. Wang, J. Yu, J. Bitton, J. Spisak, J. Park, J. Rocca, J. Johnstun, J. Saxe, J. Jia, K. V. Alwala, K. Upasani, K. Plawiak, K. Li, K. Heafield, K. Stone, K. El-Arini, K. Iyer, K. Malik, K. Chiu, K. Bhalla, L. Rantala-Yeary, L. van der Maaten, L. Chen, L. Tan, L. Jenkins, L. Martin, L. Madaan, L. Malo, L. Blecher, L. Landzaat, L. de Oliveira, M. Muzzi, M. Pasupuleti, M. Singh, M. Paluri, M. Kardas, M. Oldham, M. Rita, M. Pavlova, M. Kambadur, M. Lewis, M. Si, M. K. Singh, M. Hassan, N. Goyal, N. Torabi, N. Bashlykov, N. Bogoychev, N. Chatterji, O. Duchenne, O. Çelebi, P. Alrassy, P. Zhang, P. Li, P. Vasic, P. Weng, P. Bhargava, P. Dubal, P. Krishnan, P. S. Koura, P. Xu, Q. He, Q. Dong, R. Srinivasan, R. Ganapathy, R. Calderer, R. S. Cabral, R. Stojnic, R. Raileanu, R. Girdhar, R. Patel, R. Sauvestre, R. Polidoro, R. Sumbaly, R. Taylor, R. Silva,

R. Hou, R. Wang, S. Hosseini, S. Chennabasappa, S. Singh, S. Bell, S. S. Kim, S. Edunov, S. Nie, S. Narang, S. Raparthy, S. Shen, S. Wan, S. Bhosale, S. Zhang, S. Vandenhende, S. Batra, S. Whitman, S. Sootla, S. Collot, S. Gururangan, S. Borodinsky, T. Herman, T. Fowler, T. Sheasha, T. Georgiou, T. Scialom, T. Speckbacher, T. Mihaylov, T. Xiao, U. Karn, V. Goswami, V. Gupta, V. Ramanathan, V. Kerkez, V. Gonguet, V. Do, V. Vogeti, V. Petrovic, W. Chu, W. Xiong, W. Fu, W. Meers, X. Martinet, X. Wang, X. E. Tan, X. Xie, X. Jia, X. Wang, Y. Goldschlag, Y. Gaur, Y. Babaei, Y. Wen, Y. Song, Y. Zhang, Y. Li, Y. Mao, Z. D. Coudert, Z. Yan, Z. Chen, Z. Papakipos, A. Singh, A. Grattafiori, A. Jain, A. Kelsey, A. Shajnfeld, A. Gangidi, A. Victoria, A. Goldstand, A. Menon, A. Sharma, A. Boesenberg, A. Vaughan, A. Baevski, A. Feinstein, A. Kallet, A. Sangani, A. Yunus, A. Lupu, A. Alvarado, A. Caples, A. Gu, A. Ho, A. Poulton, A. Ryan, A. Ramchandani, A. Franco, A. Saraf, A. Chowdhury, A. Gabriel, A. Bharambe, A. Eisenman, A. Yazdan, B. James, B. Maurer, B. Leonhardi, B. Huang, B. Loyd, B. D. Paola, B. Paranjape, B. Liu, B. Wu, B. Ni, B. Hancock, B. Wasti, B. Spence, B. Stojkovic, B. Gamido, B. Montalvo, C. Parker, C. Burton, C. Mejia, C. Wang, C. Kim, C. Zhou, C. Hu, C.-H. Chu, C. Cai, C. Tindal, C. Feichtenhofer, D. Civin, D. Beaty, D. Kreymer, D. Li, D. Wyatt, D. Adkins, D. Xu, D. Testuggine, D. David, D. Parikh, D. Liskovich, D. Foss, D. Wang, D. Le, D. Holland, E. Dowling, E. Jamil, E. Montgomery, E. Presani, E. Hahn, E. Wood, E. Brinkman, E. Arcaute, E. Dunbar, E. Smothers, F. Sun, F. Kreuk, F. Tian, F. Ozgenel, F. Caggioni, F. Guzmán, F. Kanayet, F. Seide, G. M. Florez, G. Schwarz, G. Badeer, G. Swee, G. Halpern, G. Thattai, G. Herman, G. Sizov, Guangyi, Zhang, G. Lakshminarayanan, H. Shojanazeri, H. Zou, H. Wang, H. Zha, H. Habeeb, H. Rudolph, H. Suk, H. Aspegren, H. Goldman, I. Damlaj, I. Molybog, I. Tufanov, I.-E. Veliche, I. Gat, J. Weissman, J. Geboski, J. Kohli, J. Asher, J.-B. Gaya, J. Marcus, J. Tang, J. Chan, J. Zhen, J. Reizenstein, J. Teboul, J. Zhong, J. Jin, J. Yang, J. Cummings, J. Carvill, J. Shepard, J. McPhie, J. Torres, J. Ginsburg, J. Wang, K. Wu, K. H. U, K. Saxena, K. Prasad, K. Khandelwal, K. Zand, K. Matosich, K. Veeraraghavan, K. Michelena, K. Li, K. Huang, K. Chawla, K. Lakhotia, K. Huang, L. Chen, L. Garg, L. A, L. Silva, L. Bell, L. Zhang, L. Guo, L. Yu, L. Moshkovich, L. Wehrstedt, M. Khabsa, M. Avalani, M. Bhatt, M. Tsimpoukelli, M. Mankus, M. Hasson, M. Lennie, M. Reso, M. Groshev, M. Naumov, M. Lathi, M. Keneally, M. L. Seltzer, M. Valko, M. Restrepo, M. Patel, M. Vyatskov, M. Samvelyan, M. Clark, M. Macey, M. Wang, M. J. Hermoso, M. Metanat, M. Rastegari, M. Bansal, N. Santhanam, N. Parks, N. White, N. Bawa, N. Singhal, N. Egebo, N. Usunier, N. P. Laptev, N. Dong, N. Zhang, N. Cheng, O. Chernoguz, O. Hart, O. Salpekar, O. Kalinli, P. Kent, P. Parekh, P. Saab, P. Balaji, P. Rittner, P. Bontrager, P. Roux, P. Dollar, P. Zvyagina, P. Ratanchandani, P. Yuvraj, Q. Liang, R. Alao, R. Rodriguez, R. Ayub, R. Murthy, R. Nayani, R. Mitra, R. Li, R. Hogan, R. Battey, R. Wang, R. Maheswari, R. Howes, R. Rinott, S. J. Bondu, S. Datta, S. Chugh, S. Hunt, S. Dhillon, S. Sidorov, S. Pan, S. Verma, S. Yamamoto, S. Ramaswamy, S. Lindsay, S. Lindsay, S. Feng, S. Lin, S. C. Zha, S. Shankar, S. Zhang, S. Zhang, S. Wang, S. Agarwal, S. Sajuyigbe, S. Chintala, S. Max, S. Chen, S. Kehoe, S. Satterfield, S. Govindaprasad, S. Gupta, S. Cho, S. Virk, S. Subramanian, S. Choudhury, S. Goldman, T. Remez, T. Glaser, T. Best, T. Kohler, T. Robinson, T. Li, T. Zhang, T. Matthews, T. Chou, T. Shaked, V. Vontimitta, V. Ajayi, V. Montanez, V. Mohan, V. S. Kumar, V. Mangla, V. Albiero, V. Ionescu, V. Poenaru, V. T. Mihailescu, V. Ivanov, W. Li, W. Wang, W. Jiang, W. Bouaziz, W. Constable, X. Tang, X. Wang, X. Wu, X. Wang, X. Xia, X. Wu, X. Gao, Y. Chen, Y. Hu, Y. Jia, Y. Qi, Y. Li, Y. Zhang, Y. Zhang, Y. Adi, Y. Nam, Yu, Wang, Y. Hao, Y. Qian, Y. He, Z. Rait, Z. DeVito, Z. Rosnbrick, Z. Wen, Z. Yang, and Z. Zhao, "The llama 3 herd of models," 2024. [Online]. Available: https://arxiv.org/abs/2407.21783

[5] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[6] "Cơ quan ngôn luận của ubnd tp hà nội," https://kinhtedothi.vn/.

[7] D. Kouzis-Loukas, *Learning Scrapy*.   Packt Publishing Ltd, 2016.

[8] C. V. Nguyen, T. Nguyen, Q. Nguyen, H. Nguyen, B. Plüster, N. Pham, H. Nguyen, P. Schramowski, and T. Nguyen, "Vistral-7b-chat - towards a state-of-the-art large language model for vietnamese," 2023.

[9] D. Q. Nguyen and A. T. Nguyen, "PhoBERT: Pre-trained language models for Vietnamese," in *Findings of the Association for Computational Linguistics: EMNLP 2020*, 2020, pp. 1037–1042.

[10] "Tạp chí Điện tử của hiệp hội doanh nghiệp Đtnn," https://nhadautu.vn/.

[11] L. Phan, H. Tran, H. Nguyen, and T. H. Trinh, "ViT5: Pretrained text-to-text transformer for Vietnamese language generation," in *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Student Research Workshop*.   Association for Computational Linguistics, 2022, pp. 136–142. [Online]. Available: https://aclanthology.org/2022.naacl-srw.18

[12] V. C. Storey, "Relational database design based on the entity-relationship model," pp. 47–83, 1991. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0169023X9190033T

[13] N. L. Tran, D. M. Le, and D. Q. Nguyen, "BARTpho: Pre-trained Sequence-to-Sequence Models for Vietnamese," *arXiv preprint*, vol. arXiv:2109.09701, 2021.

[14] "Trang thông tin điện tử tổng hợp," https://vietnambiz.vn/.

[15] "Báo vietnamnet - tin tức mới nhất, tin nhanh việt nam và thế giới," https://vietnamnet.vn/.

[16] "Vietstock - tin chứng khoán, bất động sản, kinh tế đầu tư, tài chính và tiêu dùng," https://vietstock.vn/.

[17] "Nhịp sống kinh tế việt nam & thế giới," https://vneconomy.vn/.

[18] "Báo vnexpress - báo tiếng việt nhiều người xem nhất," https://vnexpress.net/.

Topic	Text	Label
Bank Capital	Tăng vốn điều lệ, từng bước nâng tầm vị thế VPB đã có một nhịp tăng ấn tượng bắt đầu từ tháng 04/2020. Trong đó, điểm nhấn là giai đoạn bắt đầu từ tháng 02/2021 đến tháng 06/2021 khi cổ phiếu này liên tục phá đỉnh, tăng hơn 150% từ mức giá 28,000 lên hơn 70,000 sau khi đón nhận nhiều thông tin tích cực từ kết quả kinh doanh, bán 49% cổ phần FE Credit, tăng vốn điều lệ "khủng" (*Increasing charter capital, progressively elevating the status VPB has experienced an impressive growth trajectory starting from April 2020. Notably, from February 2021 to June 2021, the stock consistently reached new highs, increasing more than 150% from a price of 28,000 to over 70,000 after receiving positive news including business results, the sale of a 49% stake in FE Credit, and a significant increase in charter capital.*).	Positive
Credit Growth - Money Supply	Các doanh nghiệp phi tài chính như thép, phân bón, dầu khí, bán lẻ, xây dựng, hóa chất, du lịch... đều tăng trưởng âm, thậm chí thua lỗ nặng trong quý cuối năm 2022. Kết quả kinh doanh của nhóm bất động sản cũng không mấy khả quan trước khó khăn bởi 2 kênh huy động vốn chủ yếu là tín dụng ngân hàng và trái phiếu bị thắt chặt. (*Non-financial businesses such as steel, fertilizers, oil and gas, retail, construction, chemicals, and tourism all experienced negative growth, and even incurred heavy losses in the last quarter of 2022. The financial performance of the real estate sector was also not very promising due to difficulties, as the two main funding channels—bank credit and bonds—were tightened.*)	Negative
Banking Services	Ở giai đoạn 2 đến hết ngày 11/08/2024, với hai kỳ xét giải, chủ thẻ tín dụng Sacombank Visa hạng Infinite/Signature/Platinum có cơ hội được hoàn tiền đến 2 triệu VND . Theo đó, 1,000 khách hàng có doanh số giao dịch thẻ tại nước ngoài cao nhất (tối thiểu từ 4 triệu VND) sẽ được hoàn 1 triệu VND/kỳ. (*In Phase 2, until August 11, 2024, with two evaluation periods, Sacombank Visa Infinite/Signature/Platinum credit cardholders have the opportunity to receive a cashback of up to 2 million VND. Accordingly, the top 1,000 customers with the highest transaction volume abroad (minimum of 4 million VND) will receive a cashback of 1 million VND per period.*)	Neutral

Table 4: Examples from the sentiment analysis dataset, showcasing various banking topics with their corresponding sentiment labels.