RESEARCH ARTICLE

EXPLORING THE POTENTIAL OF GRAPH NEURAL NETWORKS FOR VIETNAMESE SENTIMENT ANALYSIS

Dinh Minh Hoa, Nguyen Tran Thanh Nha, Nguyen Van Xanh, Bui Thi Thanh Tu, Tran Khai Thien*

Ho Chi Minh City University of Foreign Languages and Information Technology, Vietnam

hoadm@huflit.edu.vn, 22DH112486@st.huflit.edu.vn, xannv@huflit.edu.vn, tubtt@huflit.edu.vn, thientk@huflit.edu.vn

ABSTRACT — Graph Neural Networks (GNNs) have demonstrated outstanding potential in various natural language processing (NLP) tasks; however, their application to Vietnamese sentiment analysis remains relatively underexplored. This study evaluates the effectiveness of several GNN-based models – including TextGCN, HGAT, BertGCN, and GraphSAGE – in analyzing sentiments in Vietnamese textual data. Experiments were conducted on two benchmark Vietnamese sentiment datasets, UIT-VSFC and Foody. The empirical results compare the performance of GNN models with both traditional machine learning approaches and representative deep learning architectures. Key performance metrics such as accuracy and F1-score were analyzed to highlight the strengths of each method. The findings reveal that GNN-based models exhibit superior capabilities in capturing contextual and semantic relationships within texts, particularly in complex sentiment scenarios. This study aims to investigate the potential of applying GNNs to enhance Vietnamese sentiment analysis, offering a novel perspective compared to traditional and deep learning models. Additionally, the implementation code for the GNN models has been made available on GitHub† to serve as a resource for other research groups interested in this domain.

Keywords — Natural Language Processing, Sentiment Analysis, Vietnamese Sentiment Analysis, Graph Neural Networks, Graph-based Feature Extraction.

I. INTRODUCTION

Sentiment analysis, also known as opinion mining, is a critical task in natural language processing (NLP). It involves the automatic identification and classification of opinions, attitudes, and emotions expressed in textual data. With the explosive growth of user-generated content on social media platforms, e-commerce websites, and online forums, sentiment analysis has become an indispensable tool for businesses and organizations. Notable applications of sentiment analysis include monitoring customer satisfaction [1], predicting market trends [2], and detecting public sentiment toward political or social issues [3]. Despite being a well-researched problem, sentiment analysis still presents considerable challenges – especially for languages with complex linguistic structures such as Vietnamese.

In recent years, the emergence of GNNs has revolutionized the way structured data is processed in machine learning. Unlike traditional deep learning models that primarily operate on sequential or tabular data, GNNs excel at modeling relational data and capturing interdependencies among elements in a graph. These capabilities make GNNs particularly powerful for NLP tasks, where contextual relationships and syntactic structures are crucial. By leveraging graph-based representations such as dependency trees or word co-occurrence graphs, GNNs have shown significant improvements in tasks like machine translation, text classification, and question answering.

In the context of Vietnamese, applying GNNs to sentiment analysis remains in its early stages. Current approaches to Vietnamese sentiment analysis largely rely on traditional machine-learning models or transformer-based architectures such as BERT [4]. While these methods have achieved certain levels of effectiveness, they often struggle to represent the intricate grammatical and semantic relationships in Vietnamese texts. This study focuses on exploring the potential of GNN-based models to address these limitations and provides a comparative analysis with traditional approaches to highlight the strengths of GNNs in Vietnamese sentiment analysis.

II. RELATED WORK

As a critical task within NLP, sentiment analysis has significantly expanded its scope of applications alongside the rapid increase in user-generated content on digital platforms. In the business domain, it enables companies to monitor customer feedback [1], enhance service quality, and develop targeted marketing strategies [5]. Politicians and policymakers employ sentiment analysis to gauge public opinion, understand social trends, and forecast election outcomes. In market research, it assists in evaluating product performance, predicting consumer demand, and understanding customer behavior. Furthermore, sentiment analysis has found

* Coressponding Author

† https://github.com/hoadm-net/TextGraph

applications in mental health care by identifying emotional states in textual data such as social media posts or clinical notes, thereby providing valuable insights for timely interventions [3].

Sentiment analysis techniques have evolved substantially over time, encompassing rule-based methods [6], traditional machine learning approaches [7] and modern deep learning techniques [8]. Rule-based methods utilize sentiment lexicons containing positive and negative terms in conjunction with syntactic rules to determine sentiment polarity. While interpretable and transparent, these approaches often struggle with linguistic complexities such as negation, sarcasm, and idiomatic expressions. Traditional machine learning models like Support Vector Machines (SVM), Naïve Bayes, and logistic regression rely on manually engineered features such as n-grams, part-of-speech (POS) tags, and TF-IDF values. Although more effective than rule-based methods, these models require substantial feature engineering effort and often fail to capture deep semantic relationships within the text.

Vietnamese presents unique challenges to sentiment analysis. As an isolating language, Vietnamese lacks explicit word boundaries, making word segmentation a crucial preprocessing step. Ambiguity poses another major obstacle, as many Vietnamese words possess multiple meanings depending on context. For example, the word "đắt" may mean "expensive" or "popular and in high demand" depending on usage. Additionally, sentiment in Vietnamese often hinges on tone or contextual nuance, which written text cannot directly convey. The use of sarcasm, idioms, and regional dialect variations further adds to the complexity.

Addressing these challenges necessitates innovative approaches capable of capturing the intricate linguistic structures and relationships inherent in Vietnamese texts. Graph-based models, such as GNNs, offer a promising direction by representing textual data as graphs that encode both contextual and relational information. These models hold strong potential to overcome traditional limitations and pave new avenues for Vietnamese sentiment analysis.

III. GRAPH NEURAL NETWORK MODELS FOR VIETNAMESE SENTIMENT ANALYSIS

A. OVERVIEW OF GRAPH NEURAL NETWORKS

GNNs represent a significant advancement in deep learning, specifically designed to operate directly on graph-structured data. Unlike traditional models that focus on grid-like data such as images or text sequences, GNNs are tailored to process non-Euclidean data, where relationships among elements are naturally represented as nodes and edges. In a graph, nodes correspond to entities, while edges capture the relationships or interactions between those entities. GNNs leverage this structure to aggregate and propagate information across nodes, enabling the model to learn both local and global patterns within the graph.

The core mechanism of GNNs is message passing – a process in which information is exchanged between neighboring nodes through edges. In each computational layer, a node aggregates messages from its neighboring nodes, updates its representation based on the aggregated information, and forwards the updated representation to the next layer. For instance, a node v with a set of neighbors $N(v)$ computes its new representation h'_v by aggregating messages $m_{u\to v}$ from each neighbor $u \in N(v)$, followed by applying a transformation function, typically parameterized by learnable weights (see Figure 1). This iterative process allows GNNs to capture complex dependencies, even among distant nodes, by stacking multiple message-passing layers.

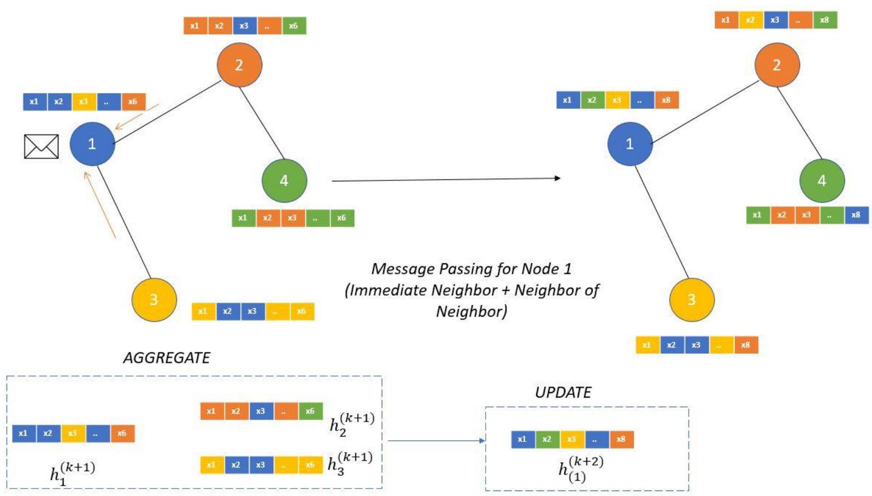

Figure 1. *Message passing mechanism between nodes*

GNNs have found widespread applications across various domains where graph-structured data is prevalent. In social network analysis [9], GNNs are employed to model user interactions and predict relationships, such as recommending new connections or identifying influential users. In molecular biology [10], GNNs have shown promise in drug discovery by predicting molecular properties based on the graph representation of chemical compounds. They also support recommender systems by modeling user–item interactions as bipartite graphs. In finance, GNNs are utilized for fraud detection by analyzing payment networks [11]. Additionally, GNNs are applied in domains such as physical simulation, traffic forecasting, and even computer vision tasks, where spatial relationships are critical.

In natural language processing, GNNs have proven effective in tasks that require a deep understanding of both the contextual and structural aspects of textual data. Text information can be represented as graphs, where words or sentences serve as nodes, and the relationships between them – such as syntactic dependencies or semantic similarity – are represented by edges. GNNs capture these relationships, offering an advantage over traditional sequence-based models by incorporating global structural information. In sentiment analysis, GNNs outperform conventional methods by modeling the dependencies among words or sentences, enabling the system to better interpret complex emotional expressions that may be overlooked by other approaches.

Despite their advantages, GNNs also face notable limitations. One of their greatest strengths – the ability to model complex relationships and dependencies – comes at the cost of high computational complexity, particularly when applied to large-scale graphs with millions of nodes and edges. Training GNNs requires substantial memory resources, as the message-passing process involves maintaining representations for all nodes and edges. Furthermore, the phenomenon of over-smoothing – where node representations become indistinguishable after multiple layers of message passing – can degrade performance, especially in deeper architectures.

To address these challenges and fully harness the potential of GNNs for NLP tasks such as sentiment analysis, several specialized architectures have been developed. For instance, TextGCN [12] explicitly models the relationships between words and documents in a text corpus as a graph. HGAT [13] (Heterogeneous Graph Attention Networks) incorporates attention mechanisms to focus on important relationships within heterogeneous graphs that contain multiple types of nodes and edges. GraphSAGE [14](Graph Sample and Aggregate) introduces efficient sampling techniques to scale GNNs to large graphs while maintaining performance. BertGCN [15] combines the strengths of GNNs with transformer-based models like BERT, enabling the system to capture both local dependencies and contextual information from pre-trained language models.

B. TextGCN

TextGCN (Text Graph Convolutional Network) is a graph-based framework designed for tasks such as text classification and sentiment analysis. This method transforms a text corpus into a heterogeneous graph, where nodes represent both words and documents, and edges encode relationships such as word co-occurrence or the relevance between words and documents. Figure 2 illustrates the architecture of the TextGCN model, while Equation 1 defines the method for computing the edge weight between node i and node j.

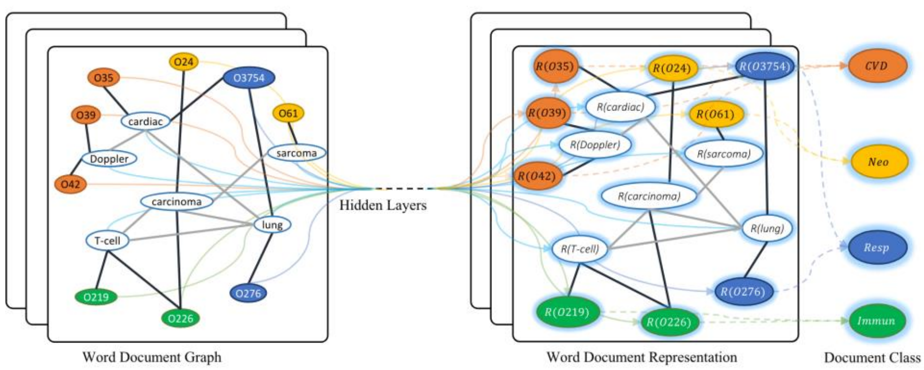

Figure 2. *Architecture of the TextGCN model*

$$A_{ij}\begin{cases} \mathbf{PMI(i,j)} \text{ If both i and j are word nodes} \\ \mathbf{TF\text{-}IDF_{ij}} \text{ If i is a document node and j is a word node} \\ \mathbf{1} \text{ If i=j} \\ \mathbf{0} \text{ Otherwise} \end{cases} \tag{1}$$

The PMI (Pointwise Mutual Information) values are calculated using Equations (2), (3), and (4).

$$PMI(i,j) = log\frac{p(i,j)}{p(i)p(j)} \quad (2)$$

$$p(i,j) = \frac{\#W(i,j)}{\#W} \quad (3)$$

$$p(i) = \frac{\#W(i)}{\#W} \quad (4)$$

Here, $\#W(i)$ denotes the number of sliding windows in the dataset that contains word i, $\#W(i,j)$ represents the number of windows containing both words i, j. And $\#W$ is the total number of sliding windows in the entire corpus.

The core mechanism of TextGCN lies in its use of Graph Convolutional Networks (GCNs) to propagate information through the constructed graph. Word nodes are initialized using pre-trained word embeddings, while document nodes are typically initialized randomly. Through multiple GCN layers, each node aggregates features from its neighboring nodes and updates its representation accordingly. This process allows TextGCN to capture both local dependencies within documents and global structural information across the corpus. A key strength of TextGCN is its ability to model latent relationships between documents, even in the absence of shared vocabulary. By integrating both word–word and word–document relationships, this approach is particularly effective in identifying complex sentiment patterns and long-range dependencies that sequential models often overlook. Moreover, the graph-based structure enables better generalization across datasets with rich relational structures.

However, TextGCN faces challenges in scalability, as constructing and processing large graphs incur high computational and memory costs. Its reliance on a fixed graph structure can also limit adaptability during training. Additionally, the model may fall short in capturing nuanced contextual information that transformer-based architectures handle more effectively. Nevertheless, TextGCN remains a foundational method that has inspired the development of more advanced GNN-based models.

C. HGAT

HGAT (Heterogeneous Graph Attention Networks) introduces an innovative approach to addressing challenges associated with short text classification by leveraging the capabilities of heterogeneous GNNs. Short texts—such as social media posts, search queries, or user reviews—are inherently sparse and often lack sufficient contextual information, making them particularly difficult to classify accurately. Traditional machine learning methods and homogeneous graph-based models often fail to effectively capture the complex relationships and contextual dependencies present in such texts.

HGAT addresses this issue by modeling data as a heterogeneous graph and incorporating attention mechanisms to selectively focus on the most relevant information. In this framework, the graph consists of diverse node types—such as words, documents, and topics—and edges encode various relationships, including word co-occurrence, word-document links, and semantic similarity. The heterogeneity of the graph allows the model to integrate multiple types of nodes and edges, which is essential for capturing the diverse relationships in short texts. By employing multi-head attention mechanisms, HGAT dynamically assigns varying importance weights to edges and nodes based on their relevance to the classification task. This enables the model to effectively aggregate information from both local context (e.g., word co-occurrence) and global context (e.g., semantic or topical relevance). Figure 3 provides a detailed illustration of the HGAT architecture.

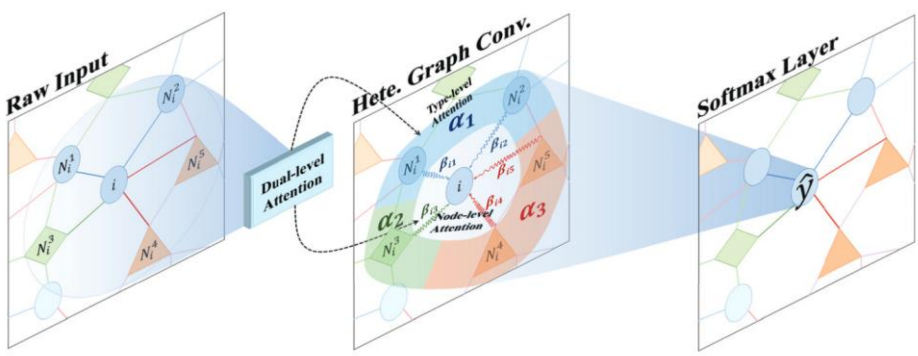

Figure 3. *HGAT model architecture*

Empirical studies have demonstrated the effectiveness of HGAT on several benchmark datasets for short text classification, outperforming state-of-the-art models in terms of both accuracy and robustness to noise. One of the key findings is the model's ability to mitigate issues related to sparsity and noise, which are common in short texts. The attention mechanism in HGAT enables it to selectively amplify meaningful signals while suppressing irrelevant or noisy data, resulting in more accurate and interpretable predictions. Furthermore, the semi-supervised learning framework adopted in HGAT allows the model to better leverage unlabeled data, which is particularly valuable in scenarios where labeled data is scarce.

D. GraphSAGE

GraphSAGE (Graph Sample and Aggregate) is an advanced inductive framework designed for learning representations on large-scale graphs. Traditional GNN architectures typically rely on a transductive learning approach, which assumes that the full graph is available during training. This limitation makes them unsuitable for scenarios involving dynamic or evolving graph structures. GraphSAGE overcomes this challenge by adopting an inductive learning paradigm, allowing the model to generalize to unseen nodes or entirely new graphs by learning an aggregation function.

This approach is particularly useful in applications where the graph is continually growing or where loading the entire graph into memory is computationally infeasible. GraphSAGE operates by iteratively sampling a fixed-size neighborhood for each node and aggregating information from these neighbors to compute the node's representation. The aggregation process is parameterized using neural architectures such as mean aggregators, pooling operations, or even LSTM-based aggregators.

This flexibility enables GraphSAGE to effectively capture both local features and structural patterns while maintaining scalability. Unlike transductive methods, which compute embeddings for a fixed set of nodes, GraphSAGE leverages the learned aggregation functions to generate embeddings for new nodes at inference time, making it highly adaptable in real-world settings. The computation steps of GraphSAGE's message-passing process are illustrated in Figure 4.

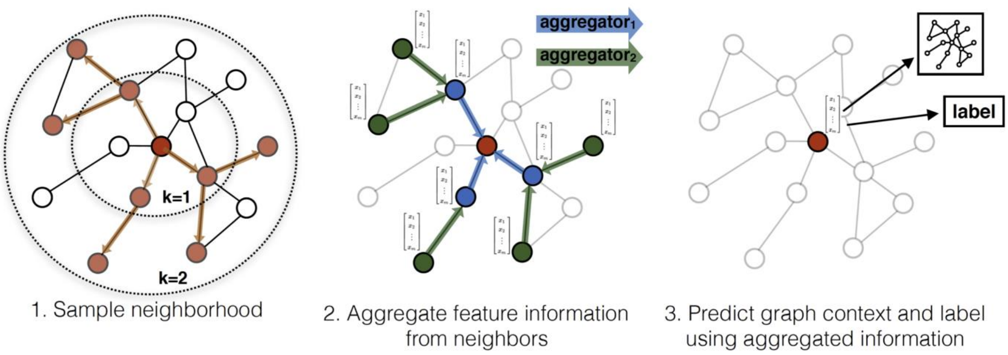

Figure 4. *Illustration of the sampling and aggregation process in the GraphSAGE model*

One of the major advantages of GraphSAGE is its ability to strike a balance between computational efficiency and representational power. By sampling a subset of neighbors rather than using all connected nodes, the model achieves significant computational gains while still retaining the capacity to capture essential graph properties. Moreover, the use of shared aggregation functions ensures that the model can be applied to different graph structures without requiring retraining.

In the context of sentiment analysis, GraphSAGE offers strong potential by effectively modeling complex dependencies between words, sentences, and documents. For instance, in Vietnamese sentiment analysis—where contextual nuances and subtle relational cues often play a critical role—GraphSAGE can aggregate both local linguistic features and global semantic patterns, thereby enhancing the model's classification accuracy. Its inductive nature also ensures adaptability, allowing it to process newly added documents or vocabulary without extensive retraining.

Despite its many benefits, GraphSAGE is not without limitations. The sampling process, while computationally efficient, may result in the loss of important structural information if the sample size is not chosen appropriately. Additionally, the model's performance heavily depends on the design of the aggregation function, which may require careful tuning for specific tasks. Nonetheless, GraphSAGE represents a significant advancement in graph-based learning, offering a scalable and generalizable solution for large and dynamic graphs, making it a promising candidate for advancing sentiment analysis—particularly in linguistically complex languages like Vietnamese.

E. BertGCN

BertGCN is an exceptionally innovative model that integrates the strengths of BERT and GCNs for text classification and sentiment analysis tasks. While BERT has become a standard in many NLP applications due to its contextual understanding and pre-trained embedding capabilities, it primarily operates on linear text sequences. This approach can overlook global relational information across the entire dataset, which is crucial for capturing inter-document or inter-sentence dependencies. In contrast, GCNs excel at modeling structural relationships within graph-structured data, making them ideal for capturing dependencies such as word co-occurrence or document similarity.

BertGCN bridges these two approaches by combining the contextual richness of BERT with the structural learning capabilities of GCNs. In the BertGCN architecture, BERT is first used to encode input texts into embeddings that capture deep semantic and contextual information. These embeddings are then used as node features in a graph, where nodes represent documents and edges encode relationships such as document similarity or keyword co-occurrence. A graph convolutional network is subsequently applied to propagate information across nodes, enabling the model to integrate both contextual knowledge and global structural patterns. The transductive nature of the GCN ensures that even unlabeled nodes (documents) can benefit from labeled data via connected edges and embeddings, thereby enhancing classification performance in semi-supervised settings.

BertGCN offers distinct advantages over both traditional methods and standalone deep learning models. By combining BERT's contextual embeddings with the relational modeling strength of GCNs, the model proves particularly effective in handling datasets with sparse or limited labeled data. This is especially valuable in scenarios like Vietnamese sentiment analysis, where nuanced language expressions and contextual meanings require a deep understanding of both document structure and inter-document relationships. Furthermore, the hybrid design of BertGCN enables the model to capture both local and global patterns. Despite its higher computational demands, BertGCN has been shown to achieve state-of-the-art performance in text classification tasks, outperforming traditional methods in terms of accuracy.

IV. EXPERIMENTAL RESULTS

A. EXPERIMENTAL DATASETS

For the task of Vietnamese sentiment analysis, the experiments in this study are based on two benchmark datasets: the Foody dataset[‡] and the UIT-VSFC dataset[§] [16], both of which are commonly used by research groups working on Vietnamese sentiment analysis. Detailed statistics of these two datasets are presented in Table 1.

Table 1. *Descriptive Statistics of the Foody and UIT-VSFC Datasets*

	Foody	UIT-VSFC
Number of Samples	38.279	16.175
Number of Classes	2	3
Shortest Sample Length	4	1
Longest Sample Length	1.336	123
Average Sample Length	80	9
Class Imbalance	No	Yes

The Foody dataset was collected, processed, and released by the NLP research group at Ho Chi Minh City University of Technology. It was gathered from the website foody.vn and contains customer reviews of locations, restaurants, and dishes. In contrast, the UIT-VSFC dataset was developed by the NLP research group at the University of Information Technology (UIT), Ho Chi Minh City. This dataset consists of student evaluations

‡ https://tinyurl.com/FoodyDataset

§ https://nlp.uit.edu.vn/datasets#h.p_4Brw8L-cbfTe covering various topics such as facilities, instructors, and academic programs at UIT, and is annotated with three sentiment labels.

B. BASELINE METHODS

To explore the potential of GNNs in Vietnamese sentiment analysis in comparison with traditional models, we conducted experiments using four GNN-based models – TextGCN, HGAT, GraphSAGE, and BertGCN—as introduced in Section III. In addition, we implemented traditional machine learning models, including SVM, Naïve Bayes (NB), Logistic Regression (LR), and Random Forests (RF), using TF-IDF for feature extraction. Furthermore, we evaluated several deep learning architectures such as CNN, RNN, and LSTM, followed by fine-tuning the PhoBERT model as a strong baseline for comparison against the GNN-based approaches.

In the context of GNN-based models, the preprocessing stage plays a pivotal role in transforming raw Vietnamese text into graph structures suitable for learning. This process typically involves Vietnamese word segmentation, normalization, and the removal of stopwords to reduce noise and highlight salient terms. Edges in the resulting graphs are constructed based on statistical co-occurrence measures such as PMI or other linguistic relations, directly shaping the topology of the graph. The effectiveness of these preprocessing choices substantially influences how information is propagated within the network and, consequently, the overall performance of GNN models. Given the unique linguistic characteristics of Vietnamese – particularly the necessity of accurate word segmentation – preprocessing decisions are especially critical for ensuring robust sentiment analysis outcomes. As such, the comparative results among GNN architectures in this study are inherently affected by the adopted preprocessing strategies.

C. EVALUATION METRICS

In sentiment analysis tasks, model performance is evaluated using a set of widely adopted and well-established metrics that quantify different aspects of system effectiveness. Among these, accuracy and F1-score are the most commonly used metrics due to their ability to provide a comprehensive assessment of classification models. These metrics are calculated based on the confusion matrix (Table 2) and are defined by Equations (5), (6), (7), and (8).

Table 2. *Confusion Matrix*

		Predicted class	
		Positive	Negative
Actual class	Positive	TP – True Positive	FN – False Negative
	Negative	FP – False Positive	TN – True Negative

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (7)$$

$$F1 - Score = 2 \times \frac{Precision \ \times Recall}{Precision + Recall} \quad (8)$$

D. RESULTS AND DISCUSSION

In the experimental phase, traditional machine learning models were implemented using the scikit-learn library, while deep learning models were developed with TensorFlow. GNN models were built using the Deep Graph Library (DGL). All experiments were conducted in the Colab-Pro environment equipped with an A100 GPU (40GB RAM) and 84GB of system RAM. The performance results of the models on the two datasets are summarized in Tables 3 and 4, respectively.

Table 3. *Results of the Models on the Foody Dataset*

	Accuracy	F1-Score
SVM	88,57	88,66
NB	85,29	85,27
LR	88,57	87,75
RF	85,66	85,30
CNN	87,64	87,59
RNN	81,30	81,32
LSTM	85,62	85,17
Fine-tuned PhoBERT model	92,39	92,38
TextGCN	88,14	88,13
HGAT	87,10	87,09
GraphSAGE	88,01	88,01
BertGCN	94,16	94,12

Table 4. *Results of the Models on the UIT-VSFC Dataset*

	Accuracy	F1-Score
SVM	88,69	86,99
NB	83,95	82,63
LR	88,31	87,43
RF	87,81	86,91
CNN	89,64	89,29
RNN	86,01	83,85
LSTM	89,58	89,07
Fine-tuned PhoBERT model	94,13	93,84
TextGCN	89,47	88,36
HGAT	87,64	85,75
GraphSAGE	88,86	88,22
BertGCN	94,54	97,19

The experimental results on the two datasets, as presented in Tables 3 and 4, indicate that both deep learning models and GNN-based approaches significantly outperform traditional machine learning methods. Notably, as initially hypothesized, BertGCN achieved the highest accuracy among all models, confirming the advantage of combining pre-trained language models with GNNs.

Based on the experimental results, the following conclusions can be drawn. Traditional machine learning models offer simplicity with acceptable performance, making them suitable for small datasets or systems with limited computational resources. In contrast, deep learning models significantly enhance performance through advanced encoding and learning mechanisms. Notably, the fine-tuning of the PhoBERT model reaffirms the remarkable capability of transformer-based architectures in sentiment analysis tasks, particularly in capturing contextual and structural information. Graph-based models such as GNNs—including TextGCN, GraphSAGE, HGAT, and BertGCN—represent a state-of-the-art approach to sentiment analysis. These models treat textual data as graphs, where nodes represent words, sentences, or documents, and edges encode relationships such as word co-occurrence or document similarity. GNNs excel at capturing both local and global dependencies, making them especially well-suited for tasks involving complex semantic structures, such as sentiment analysis. While TextGCN, GraphSAGE, and HGAT achieve performance comparable to deep learning models, BertGCN outperforms all others on both benchmark datasets.

Although BertGCN demonstrates superior performance in Vietnamese sentiment analysis, it is important to note that this model is significantly more computationally intensive than other GNN-based approaches. The integration of BERT for contextual embedding and GCN for structural learning results in higher memory consumption and longer training times, especially on large datasets. In our experiments, BertGCN required substantially more GPU resources to store and process both the contextual embeddings and the graph structure. These resource demands may pose challenges for deployment in environments with limited computational capacity. Therefore, while BertGCN achieves state-of-the-art results, its practical application should carefully consider the associated computational and hardware requirements.

V. CONCLUSION

This study systematically evaluated the effectiveness of several GNN architectures – including TextGCN, HGAT, BertGCN, and GraphSAGE – for Vietnamese sentiment analysis. Through comprehensive experiments on two benchmark datasets, we demonstrated that GNN-based models consistently outperform traditional machine learning and standard deep learning approaches in capturing both contextual and semantic relationships within Vietnamese texts. Notably, BertGCN achieved the highest accuracy, underscoring the benefit of integrating contextual embeddings with graph-based structural learning.

However, our findings also highlight important practical considerations. While GNNs, especially BertGCN, offer superior performance, they require significant computational resources, particularly for large-scale datasets. This factor may limit their applicability in resource-constrained environments. Additionally, the effectiveness of GNN models is closely tied to the quality of preprocessing and graph construction, which remains a challenging aspect of Vietnamese language processing.

In summary, GNNs represent a promising direction for advancing Vietnamese sentiment analysis by effectively modeling complex linguistic structures. Future research should focus on optimizing graph construction techniques and developing more efficient GNN architectures to address scalability and resource challenges. We hope that our findings and released codebase will serve as a valuable reference for further studies and applications in Vietnamese NLP and related domains.

VI. ACKNOWLEDGMENTS

This research is funded by Ho Chi Minh City University of Foreign Languages – Information Technology under grant number H2024-02.

VII. REFERENCES

[1] N. Pleerux and A. Nardkulpat (2023). “Sentiment analysis of restaurant customer satisfaction during COVID-19 pandemic in Pattaya, Thailand,” *Heliyon*, vol. 9, no. 11, Nov. 2023, doi: 10.1016/j.heliyon.2023.e22193.

[2] W. jun Gu *et al.* (2024). “Predicting stock prices with finbert-lstm: Integrating news sentiment analysis,” in *Proceedings of the 2024 8th International Conference on Cloud and Big Data Computing*, 2024, pp. 67–72.

[3] M. Wankhade, A. C. S. Rao, and C. Kulkarni (2022). "A survey on sentiment analysis methods, applications, and challenges," *Artif. Intell. Rev.*, vol. 55, no. 7, pp. 5731–5780, Oct. 2022, doi: 10.1007/s10462-022-10144-1.

[4] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, J. Burstein, C. Doran, and T. Solorio, Eds., Association for Computational Linguistics, 2019, pp. 4171–4186. doi: 10.18653/V1/N19-1423.

[5] N. Hicham, H. Nassera, and S. Karim (2023). "Strategic framework for leveraging artificial intelligence in future marketing decision-making," *J. Intell. Manag. Decis.*, vol. 2, no. 3, pp. 139–150, 2023.

[6] G. Qiu, X. He, F. Zhang, Y. Shi, J. Bu, and C. Chen (2010). "DASA: Dissatisfaction-oriented Advertising based on Sentiment Analysis," *Expert Syst. Appl.*, vol. 37, no. 9, pp. 6182–6191, 2010, doi: https://doi.org/10.1016/j.eswa.2010.02.109.

[7] R. Moraes, J. F. Valiati, and W. P. G. Neto (2013). "Document-level sentiment classification: An empirical comparison between SVM and ANN," *Expert Syst. Appl.*, vol. 40, no. 2, pp. 621–633, 2013, doi: https://doi.org/10.1016/j.eswa.2012.07.059.

[8] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya (2021). "ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis," *Future Gener. Comput. Syst.*, vol. 115, pp. 279–294, 2021, doi: https://doi.org/10.1016/j.future.2020.08.005.

[9] X. Li, L. Sun, M. Ling, and Y. Peng (2023). "A survey of graph neural network based recommendation in social networks," *Neurocomputing*, vol. 549, p. 126441, 2023, doi: https://doi.org/10.1016/j.neucom.2023.126441.

[10] H.-C. Yi, Z.-H. You, D.-S. Huang, and C. K. Kwoh (2021). "Graph representation learning in bioinformatics: trends, methods and applications," *Brief. Bioinform.*, vol. 23, no. 1, p. bbab340, Sep. 2021, doi: 10.1093/bib/bbab340.

[11] F. Shi, Y. Cao, Y. Shang, Y. Zhou, C. Zhou, and J. Wu (2022). "H2-FDetector: A GNN-based Fraud Detector with Homophilic and Heterophilic Connections," in *Proceedings of the ACM Web Conference 2022*, in WWW '22. New York, NY, USA: Association for Computing Machinery, 2022, pp. 1486–1494. doi: 10.1145/3485447.3512195.

[12] L. Yao, C. Mao, and Y. Luo (2019). "Graph convolutional networks for text classification," in *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence*, in AAAI'19/IAAI'19/EAAI'19. Honolulu, Hawaii, USA: AAAI Press, 2019. doi: 10.1609/aaai.v33i01.33017370.

[13] H. Linmei, T. Yang, C. Shi, H. Ji, and X. Li (2019). "Heterogeneous Graph Attention Networks for Semi-supervised Short Text Classification," 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, 2019, pp. 4820–4829. doi: 10.18653/v1/D19-1488.

[14] W. L. Hamilton, R. Ying, and J. Leskovec (2017). "Inductive representation learning on large graphs," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, in NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, pp. 1025–1035.

[15] Y. Lin *et al.* (2021). "BertGCN: Transductive Text Classification by Combining GNN and BERT," in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, Online: Association for Computational Linguistics, 2021, pp. 1456–1462. doi: 10.18653/v1/2021.findings-acl.126.

[16] K. V. Nguyen, V. D. Nguyen, P. X. V. Nguyen, T. T. H. Truong, and N. L.-T. Nguyen (2018). "UIT-VSFC: Vietnamese Students' Feedback Corpus for Sentiment Analysis," in *2018 10th International Conference on Knowledge and Systems Engineering (KSE)*, Ho Chi Minh City: IEEE, Nov. 2018, pp. 19–24. doi: 10.1109/KSE.2018.8573337.

KHÁM PHÁ TIỀM NĂNG CỦA MẠNG NƠ-RON ĐỒ THỊ TRONG PHÂN TÍCH CẢM XÚC TIẾNG VIỆT

Đinh Minh Hòa, Nguyễn Trần Thanh Nhã, Nguyễn Văn Xanh, Bùi Thị Thanh Tú, Trần Khải Thiện

TÓM TẮT – Mạng nơ-ron đồ thị (GNN) đã chứng minh tiềm năng vượt trội trong nhiều nhiệm vụ xử lý ngôn ngữ tự nhiên, tuy nhiên, ứng dụng của chúng trong phân tích cảm xúc tiếng Việt vẫn chưa được khám phá sâu sắc. Nghiên cứu này đánh giá hiệu quả của các mô hình dựa trên GNN, bao gồm TextGCN, HGAT, BertGCN và GraphSAGE, trong việc phân tích cảm xúc từ văn bản tiếng Việt. Chúng tôi thực hiện các thử nghiệm trên hai tập dữ liệu tiếng Việt UIT-VSFC và Foody, là các tập dữ liệu tiêu chuẩn trong phân tích cảm xúc tiếng Việt. Kết quả thực nghiệm so sánh hiệu suất của các mô hình GNNs với các phương pháp học máy truyền thống cũng như các mô hình học sâu điển hình. Chúng tôi phân tích các chỉ số hiệu suất chính như độ chính xác (accuracy) và F1-score để làm nổi bật ưu điểm của từng phương pháp. Kết quả cho thấy, các mô hình dựa trên GNN đạt hiệu suất vượt trội trong việc nắm bắt mối quan hệ ngữ cảnh và ngữ nghĩa trong văn bản, đặc biệt trong các tình huống cảm xúc phức tạp. Nghiên cứu này nhằm khảo sát và khám phá tiềm năng của việc áp dụng GNN để nâng cao phân tích cảm xúc tiếng Việt, cung cấp góc nhìn mới so với các mô hình truyền thống và học sâu, và đồng thời cung cấp mã nguồn triển khai các mô hình GNNs trên Github làm tài liệu cho các nhóm nghiên cứu khác có quan tâm đến chủ đề này.

Từ khóa – Xử lý ngôn ngữ tự nhiên, Phân tích cảm xúc, Phân tích cảm xúc Tiếng Việt, Mạng nơ-ron đồ thị, Trích xuất đặc trưng bằng đồ thị.

Dr. Tran Khai Thien received his Ph.D. in Computer Science from Ho Chi Minh City University of Technology (HCMUT), Vietnam National University, in 2022. He is currently a lecturer at the Faculty of Information Technology, Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), where he also leads the IDPS research group. Dr. Thien serves as a reviewer and has authored over 20 publications in SCIE-indexed journals and prestigious international conferences.

His main research interests include Natural Language Processing and Artificial Intelligence.

Mr. Nguyen Van Xanh received his M.Eng. in Geographic Information Systems (GIS) from Ho Chi Minh City University of Technology, Vietnam National University, in 2005. He is currently a lecturer at the Faculty of Information Technology, Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT).

His main research interests include Intelligent Information Systems and Data Visualization.

Nguyen Tran Thanh Nha is currently an undergraduate student majoring in Software Engineering within the Information Technology program at Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT).

Mr. Dinh Minh Hoa received his M.Sc. in Information Technology from Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT) in 2022. He is currently a lecturer at the Faculty of Information Technology, HUFLIT, and a member of the university's IDPS research group.

His primary research interests include Artificial Intelligence, Natural Language Processing, and Question Answering Systems.

Ms. Bui Thi Thanh Tu received her M.Sc. in Information Technology from the Francophone Institute for Informatics (IFI), Hanoi University of Science and Technology, in 2003. She is currently a lecturer at the Faculty of Information Technology, Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT).

Prior to her academic career, she accumulated over 15 years of industry experience in software engineering, holding various roles such as QC Manager, QA Leader, Project Manager, and Business Analyst across diverse organizations, including software companies, telecommunications firms, and multinational corporations.

Her primary research interests focus on the application of Artificial Intelligence and Natural Language Processing in software engineering