HNUE JOURNAL OF SCIENCE
Natural Sciences 2024, Volume 69, Issue 3, pp. 46-56
This paper is available online at http://hnuejs.edu.vn/ns
DOI: 10.18173/2354-1059.2024-0034

EMOTION RECOGNITION IN LEARNERS WITH EMOJI SENTIMENT ACCOMPANIMENT USING THE PHOBERT MODEL

Tran Thi Dung[1], Le Nhat Tung[2], Bui Ngoc Dung[3] and Vu Huan[3,*]

[1]*Division of Information Technology, Campus in Ho Chi Minh City, University of Transport and Communications, Ho Chi Minh city, Vietnam*
[2]*Faculty of Information Technology, Dong Nai Technology University, Dong Nai province, Vietnam*
[3]*Faculty of Information Technology, University of Transport and Communications, Hanoi city, Vietnam*
[*]Corresponding author: Vu Huan, e-mail: huanv@utc.edu.vn

Received October 4, 2024. Revised October 24, 2024. Accepted October 31, 2024.

Abstract. This paper proposes an advanced method for recognizing learners' emotions by incorporating the use of emojis to reflect the modern communication tendencies of learners, typically young individuals. The method is built on the PhoBERT model, a variant of BERT optimized for Vietnamese. Data was collected from opinion surveys of learners at the Ho Chi Minh City campus of the University of Transport and Communications to train and test the model. The system is designed to analyze text and recognize seven basic emotions: enjoyment, trust, hope, sadness, surprise, fear, and others. Corresponding emojis are then assigned to each emotion type to more clearly illustrate the learners' emotional states. Experimental results show that combining PhoBERT and emojis not only enhances the accuracy of emotion recognition but also makes communication more intuitive and vivid. The model achieved an accuracy of 74.1%. The paper also discusses practical applications of this system in the field of education, where teachers can quickly and accurately understand and respond to students' emotions, thereby improving teaching effectiveness.

Keywords: opinion mining, sentiment analysis, emotion recognition, emoji, BERT, PhoBERT.

1. Introduction

Emotion recognition in text is a research field that has been attracting significant interest in the scientific community, particularly in the context of modern communication. This method not only helps in better understanding users' emotional responses but also has applications in various domains such as education, customer service, and psychology. Globally, numerous studies have been conducted to develop emotion recognition models based on different languages using advanced techniques in natural language processing (NLP). One of the prominent models in this field is BERT (Bidirectional Encoder Representations from Transformers), introduced by Devlin et al. [1]. BERT marked a significant breakthrough in NLP with its ability to understand bidirectional context, greatly improving the accuracy of text classification tasks, including emotion recognition. Following BERT's success, Liu et al. developed RoBERTa (Robustly Optimized BERT Pretraining Approach) [2], optimizing the training process and enhancing BERT's performance across various NLP tasks. Subsequent studies have demonstrated the effectiveness of these models in diverse contexts [3], [4].

In Vietnam, research and development of NLP models for the Vietnamese language have been limited. However, Vietnamese scientists have made significant strides in building suitable language models. Notably, D. Q. Nguyen and A. T. Nguyen from the VinAI Research Institute developed the PhoBERT model, an optimized variant of RoBERTa for Vietnamese [5]. PhoBERT has shown outstanding performance in Vietnamese text analysis tasks thanks to training on a large volume of data and the use of advanced tokenization and encoding techniques.

Apart from domestic efforts, many international studies have focused on combining NLP models with emojis to enhance emotion recognition capabilities. Felbo et al. [6] demonstrated that using emojis can significantly improve accuracy in emotion recognition by providing additional visual context. Other studies have also shown that integrating emojis into text analysis can help recognize emotions more accurately and align with modern communication trends [7], [8].

This research presents a sophisticated approach for detecting learners' emotions by integrating emojis, which are commonly used in modern communication, particularly among young people. Currently, most young learners, when communicating on social networking sites, often use accompanying emoticons instead of text, which helps save typing time and makes communication more engaging. Therefore, we also focus on further analysis of emoticons. This technique is based on the PhoBERT model, which we have specifically enhanced for the Vietnamese language. We gathered data from student opinion surveys conducted at the Ho Chi Minh City campus of the University of Transport and Communications to train and evaluate the model, comprising more than 5,000 responses. Surveys were taken from the Department of Testing - Quality Assurance and Inspection at the school. In addition, we also collected additional data from sources such as the UTC2 Fanpage, fan pages of faculties, departments, UTC2 Confessions, etc., from July 2014 to November 2022.

The system analyzes text to identify seven basic emotions: enjoyment, trust, hope, sadness, surprise, fear, and others. We then assign corresponding emojis to each emotion to clearly illustrate the learners' emotional states. Experimental results show that combining PhoBERT and emojis not only enhances the accuracy of emotion recognition but also makes communication more intuitive and lively.

The paper also discusses the practical applications of this system in the education sector, where teachers can quickly and accurately understand and respond to students' emotions, thereby improving teaching effectiveness. In conclusion, integrating the PhoBERT language model with emoji sentiment is a promising approach that brings significant value to intelligent systems in education and related fields.

2. Content

2.1. Sentiment analysis using the PhoBERT model

2.1.1. Emotion and perspective recognition

Sentiment analysis is a developing discipline within natural language processing (NLP) that utilizes various NLP methods to gather information about human emotions from spoken or written language. Recently, young people have tended to use many emoticons instead of sentences on social networking platforms (Figure 1). In the process of emotion recognition, if we only process sentences and ignore emoticons, it will not accurately represent the learner's emotions. The goal is to determine whether these feelings are positive or negative. By utilizing social media and NLP methodologies, it is feasible to gather and scrutinize user interest levels regarding shared events and information. This can improve the consolidation of user input for companies and organizations.

Figure 1. Sentiment Analysis from Text

Important concerns that need to be resolved are the following:

- How can machines differentiate between a remark that is based on personal opinion or reflects personal feelings?

- How can machines accurately discern between distinct emotional polarities, such as "Enjoyment", "Sadness", "Hope", "Surprise", "Fear", "Trust", and other emotions?

- How can machines process and respond to subjective emotions?

- How can machines evaluate or assess the perspective or stance of a viewpoint or opinion?

- How can robots assess the magnitude of an emotion?

- How can machines discern between truth and opinion, particularly in situations when language seems positive but is caustic, by employing advanced language processing techniques?

2.1.2. Constructing a dataset and emojis

Based on the six basic emotions identified by psychologist Robert Plutchik [10], we have selected the following emotion labels: happiness ("enjoyment"), trust ("trust"), anticipation ("hope"), sadness ("sadness"), surprise ("surprise"), and fear ("fear"). In addition, sentences that do not express emotions are labeled as “other.” Guided by Ekman's [11] insights into emotional spectra, we annotated Vietnamese texts with seven emotion labels as described below:

- Trust: Assigned to sentences expressing attitudes, viewpoints, or emotions related to the belief that something will certainly happen. Trust encompasses belief in objects or people as well as self-confidence, enhancing strengths, and addressing weaknesses to achieve success.

- Hope: Assigned to sentences depicting a state of anticipation, typically conveyed through verbs indicating expectation and hope, which positively influence life by fostering confidence in future events.

- Enjoyment: Assigned to sentences reflecting feelings of joy, satisfaction, contentment, and a sense of completeness. Happiness is often associated with positive emotions and life satisfaction.

- Surprise: Assigned to sentences indicating a short-term mental and physiological response to unexpected events. Surprise can vary in intensity and can be pleasant, unpleasant, positive, or negative.

- Sadness: Designated for sentences that convey emotional states in reaction to unfavorable circumstances, such as challenges, letdowns, or setbacks. The feeling of sadness typically decreases as time goes on and emotions become more stable.

- Fear: Applied to words that convey intense and innate emotions associated with universal biological responses and personal sentiments. Fear serves as a signal that notifies us of the existence of peril or hazards that have the potential to inflict bodily or mental damage.

- Other: Designated for statements that lack a distinct expression of any particular emotional state.

In this dataset, we have also processed emoticons, commonly referred to as emojis. Figure 2 illustrates several popular emojis, depicting the moods and expressions of the writer through symbolic characters. Initially, these symbols were developed from ASCII art and later expanded to Shift JIS art and Unicode art. Recently, graphic symbols, both static and animated, have been incorporated along with traditional text-based emoticons, often referred to as emojis. The emojis used in this study are primarily those popular on social media platforms, especially favored by the younger generation. Combined with the data collected on learner opinions, the research team filtered out the emojis most frequently used to analyze learner emotions.

After identifying the list of emojis, our research team annotated them with labels. The labels include happiness ("enjoyment"), trust ("trust"), anticipation ("hope"), sadness ("sadness"), surprise ("surprise"), and fear ("fear"). Emojis that do not convey an emotion are temporarily labeled as "other."

Emoji	Name	Meaning	Lable
😂	Smiley face with tears	Extremely happy	Enjoyment
❤	Red heart	Love	Enjoyment
😍	Smiling face with heart-shaped eyes	Love or admiration	Enjoyment
😰	Face broke out in cold sweat	Sad, disappointed	Fear
😕	Confused face	Concerned or disappointed	Fear
🙏	Two hands touching each other	Praying, thanking, sometimes means high fives	Hope
💪	Flex your arms, build up your muscles	Show strength or physical strength	Hope
👌	Hands make the OK sign	Agree or want to say understood	Trust
✌	Hands make a V sign	Victory or success	Trust
🙌	The person raised both hands in celebration	Congratulations	Trust

Figure 2. List of emojis

Overall, the process of constructing the dataset involves three stages as follows:

**** Data collection stage***

In the data collection stage, the team employed various methods to gather information from different domains, including learners, workers, parents, and employment businesses, with a particular focus on those who are or have been directly studying at the university. The more opinions and perspectives from individuals who have studied at the University of Transport and Communications Campus in Ho Chi Minh City (UTC2) are collected, the richer the dataset becomes. The survey system collected feedback from undergraduate and postgraduate students at the end of each term and gathered opinions from the UTC2 Fanpage, departmental fan pages, UTC2 Confessions, evaluations, and comments on forums where UTC2 students voiced their opinions from July 2013 to November 2022, as well as from postgraduate students and alumni. We surveyed student courses K61, K62, K63, and K64 at UTC2. The number of surveys exceeded 10,000 samples. Figure 3 is an example of the data sample we collected.

Time	Question
7/7/2013 6:23:51	ad cho e hoi la hien nay tren bang diem sinh vien da co diem tich luy vay diem do da tinh chinh xac chua?va dua vao do de tinh diem ra truong co duoc ko?(vi e thay hinh nhu nha truong tinh ca diem giao duc quoc phong voi giao duc the chat,vi e tu tinh bang tay thi dc co 1,6 tich luy nhung xem trong bang diem cua nha truong thi duoc 1,7)
7/7/2013 16:24:47	các thầy cho e hỏi chút ạ. Bằng tốt nghiệp có được nhờ người nhận thay ko ạ.
7/7/2013 23:07:50	thầy cô ơi cho em hỏi ngành kinh tế xây dựng cần có những chứng chỉ gì ạ?e cảm ơn
7/8/2013 20:23:28	AD cho e hỏi về việc đăng kí môn tự chọn cho học kì I của năm học 2013-2014 Trong TKB mới của lớp e không có môn tự chọn vậy có nghĩa là không phải đăng kí hay là phải lên Website thì mới có môn để đăng kí học E cảm ơn
7/9/2013 20:14:02	Cho em hỏi tại sao hệ thống cập nhập thông tin trường mình lâu quá.Em đăng kí nộp tiền ở KTX từ hôm mùng 5 mà đến giờ nhà trường vẫn chưa xuất gửi ngân hàng,mà ngày mai mùng 10 là hết hạn rồi.
7/10/2013 14:38:00	thưa các thầy cô, nhà trường hứa hẹn là 08 là công bố lịch học kì phụ tháng 6 năm 2013, hi vọng nhà trường đăng sớm và cách lịch học thời gian để bọn em về hè còn sắp xếp lịch vào kip ạ, em xin cảm ơn
7/10/2013 17:15:55	Tại sao điểm tính trên bảng điểm online không chính xác mà nhà trường lại đưa lên nhỉ, sinh viên dựa vào đó mà tính đến khi tính lại xét ra trường thì điểm lại khác đi, đúng là bất cập quá đi mà.

Figure 3. Data sample

**** Data preprocessing stage***

In this stage, the research team processes the input data to ensure the best results when running the model, as the initial data collected contains numerous unnecessary symbols, abbreviations, HTML tags, etc. (see Table 1). To obtain the best data for training and prediction, the team follows these processing steps:

Table 1. Processing steps data

Step	Processing steps data
1	Remove HTML tags.
2	Remove repeated non-alphanumeric characters and excess whitespace.
3	Correct sentences with multiple spelling errors and abbreviations.
4	Replace characters "_" with spaces.
5	Remove bullet points and numbering list markers.
6	Convert uppercase letters to lowercase.
7	Label emojis with corresponding emotions.
8	Normalize text by filtering out Vietnamese stopwords.
9	Remove numbers from the text.
10	Delete website links in the comments.

In Vietnamese, segmenting sentences into words during the preprocessing stage is essential because many phrases can have different meanings when broken down into individual words. For example, the phrase "đi học" (go to school) can change meaning if split into separate words. To address this issue, the team uses the VnCoreNLP [5] library for word segmentation in the input data.

**** Data labeling stage***

During the labeling stage, we assign seven types of labels corresponding to human emotions as analyzed above, including happiness ("enjoyment"), trust ("trust"), anticipation ("hope"), sadness ("sadness"), surprise ("surprise"), fear ("fear"), and other ("other"). The labeling process is based on the content of learner feedback in both text and emoji forms. This stage is crucial and time-consuming, as it relies on subjective interpretation. Therefore, the research team collaborated to label the data consistently, ensuring that each sentence accurately corresponds to the appropriate human emotion.

2.1.3. PhoBert model

The PhoBERT model is a pre-trained Vietnamese language model based on the RoBERTa architecture, introduced in March 2020. There are two variants of PhoBERT: PhoBERT_Base and PhoBERT_Large. The PhoBERT model is trained on a 20GB dataset, which consists of 1GB of data collected from the Vietnamese Wikipedia corpus and 19GB of data obtained and processed from a raw 50GB dataset.

PhoBERT uses the RDRSegmenter tool from VnCoreNLP to segment the input data into words before encoding it using the Byte-Pair Encoding (BPE) encoder. Derived from the RoBERTa architecture, PhoBERT removes the sentence prediction requirement and exclusively relies on masking. PhoBERT is a user-friendly tool specifically built to seamlessly interact with libraries like Facebook's FAIRSeq and Hugging Face's Transformers. This integration enables easy accessibility to the BERT model, even for the Vietnamese language.

In this study, the research team uses the PhoBERT model to experiment on the self-constructed dataset. The goal is to recognize the emotions of learners through their comments. Figure 4 illustrates the processing workflow of PhoBERT for a given sentence.

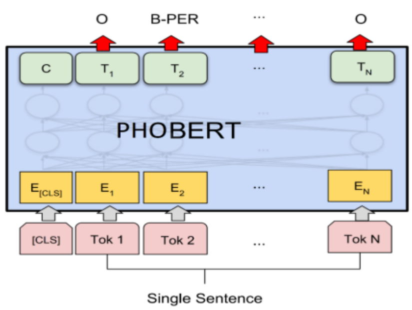

Figure 4. Illustration of the PhoBERT Fine-tuning Model [9]

Single Sentence: This is a sentence that has been tokenized into individual tokens such as [CLS], Tok1, Tok2,… TokN. The [CLS] token is a special token at the beginning of the sentence, usually used in sentence-level classification tasks or for summarizing the sentence.

Embeddings: E_{CLS}: The embedding vector for the [CLS] token. E_1, E_2, ..., E_N: The embedding vectors for the other tokens in the sentence (Tok 1, Tok 2,… Tok N).

PhoBERT: PhoBERT is a transformer model specifically designed for the Vietnamese language. It takes the embeddings of the tokens as input and learns the contextual relationships between the tokens in the sentence. After passing through PhoBERT, each token will have a new context-aware vector representation, capturing the token's meaning based on the surrounding context.

Outputs: After processing by PhoBERT, each token (T_1, T_2,… T_N) is assigned a predicted label. O: This label is for words that do not belong to any named entity (Outside). B-PER: This label is for the first token of a person's name (Beginning of Person). Similarly, the model can predict other labels such as B-LOC (location), B-ORG (organization), or I-PER (inside a person entity).

2.1.4. Utilizing the PhoBERT model in sentiment analysis

In this stage, the research team applies the pre-trained Transformer model. The attention mechanism of the Transformer allows for the simultaneous processing of all words in a sentence, irrespective of their order. Consequently, the Transformer is considered a bidirectional model, though it is more accurately described as non-directional. This characteristic enables the model to learn the context of a word based on all surrounding words, including those to the left and right.

Figure 5 below illustrates the operational process of the PhoBERT architecture combined with a linear function and the pre-trained Transformer model.

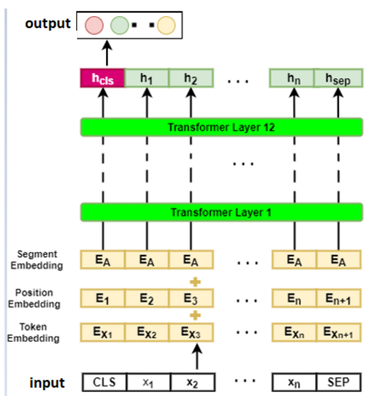

Figure 5. PhoBERT Architecture for Sentiment Analysis [12]

The model, as illustrated in Figure 5, consists of three main components:

Input: Each input profile name is preprocessed as X with n tokens, represented as X_1: n = $x_1, x_2, \ldots, x_n$ where x_i denotes the i-th position in the input sequence. The tokens are divided into vocabulary and encoded numerically using the vocabulary dictionary (Embedding Layers) developed by the PhoBERT model [9]. In addition, the PhoBERT model also takes into account the position of each token as input. The input length for sentences in the training set is determined by selecting the sequence with the largest length. Sentences with shorter lengths are then padded automatically using the padding token.

*PhoBERT Encoding***:** This study utilizes the PhoBERT_Base architecture, which consists of 12 Transformer blocks and 12 self-attention layers, to extract features from input sequences with a maximum length of 512 tokens. The result of this model is a concealed layer H = $\{h_1, h_2, \ldots, h_n\}$ that corresponds to the input sequence. To categorize input sequences based on their emotional content, the team obtains the feature representation from the [CLS] token. This feature vector represents the entire input sequence.

Output: Using the vector representation of the input sequence, the team employs a classifier with a softmax activation function to compute the probability distribution values for each classification label according to the emotion categories.

2.2. Experiments

The research team constructed a dataset comprising 13,590 sentences reflecting learner opinions, labeled with seven emotion tags: happiness ("enjoyment"), trust ("trust"), anticipation ("hope"), sadness ("sadness"), surprise ("surprise"), fear ("fear"), and other. The dataset is divided into subsets: a training set with 10,860 sentences, a test set with 1,365 sentences, and a validation set with 1,365 sentences.

The distribution of emotion labels in the dataset is uneven, as the survey data contains a significant number of sentences expressing happiness. The distribution of emotions is illustrated in Figure 6.

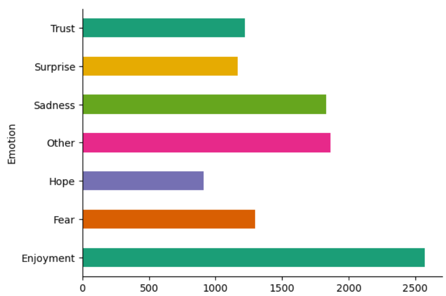

Figure 6. Emotion Label Distribution of the Texts

```
[[215   2   2  32   8  12  26]
 [  1 129   1   6   7   2   0]
 [  1   1  87   5   1   0   3]
 [ 42  13   5 132  41  11  14]
 [ 13   7   5  30 198   2   1]
 [  8   0   0  10   6 123   0]
 [ 28   0   2   4   0   2 128]]
Accuracy: 0.7408491947291361
F1 - micro: 0.7408491947291361
F1 - macro: 0.7643684881228493
```

Figure 7. Results of PhoBERT Model Execution. The results indicate that the model achieved an accuracy of 74.1%

Although the accuracy of this research is not high due to limited data and sentences with conflicting emotional content and emojis, it provides a foundational understanding of emotion recognition using PhoBERT. To demonstrate the accuracy of the PhoBERT classification model, we used a confusion matrix to visualize the confusion between actual labels and predicted labels. Figure 8 shows the confusion matrix for the best-performing classification model (PhoBERT) on the constructed dataset. It can be observed that the model performs well in classifying the labels Trust, Sadness, and Surprise, while it confuses the labels Enjoyment and Fear. There are two main reasons for this confusion. Firstly, there is a large variety of emojis associated with Enjoyment, but many of these emojis do not accurately reflect the text, as learners might use contradictory emojis to express sarcasm. Secondly, the number of texts for the Fear emotion is relatively small, and there are not many emojis representing this emotion. The limited data for these labels is a barrier to the model's optimal performance.

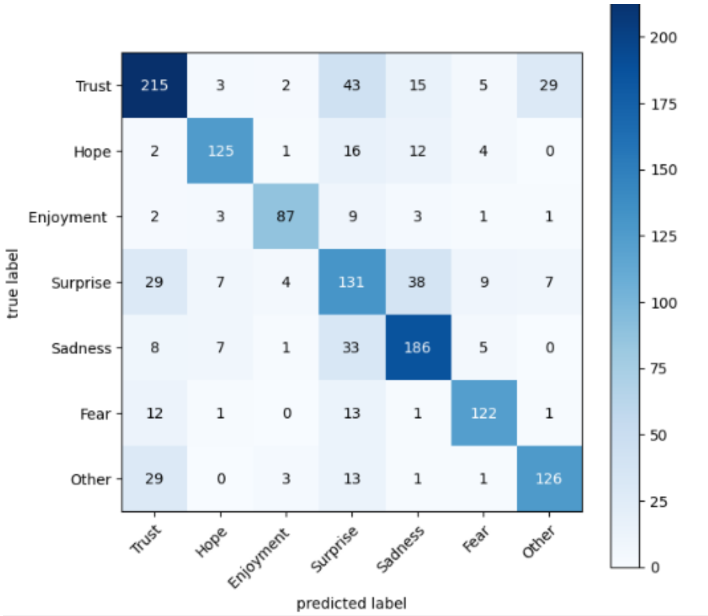

Figure 8. Confusion Matrix of Emotion Labels for the PhoBERT Model

Therefore, the data needs to be reviewed and relabeled by reliable experts to improve label agreement, which will optimize the model's results. Since the dataset for these labels has not yet achieved full prediction accuracy, improving and diversifying the dataset will be a focus for the research team.

The experimental results of the sentences are presented in Table 2. This table includes both correct and incorrect outcomes. The two incorrect cases are sentence 6 and sentence 7, with specific reasons as follows: In case 6, the initial text data yielded the emotion "Enjoyment," but when combined with an emoji representing the emotion "Trust," the model misinterpreted the result. In case 7, the text indicated the emotion "Fear," but the emoji was labeled as "Surprise," leading to confusion in the final output. Additionally, sentence 7 contained the number "86" and the character string "P501C2," which further contributed to the misinterpretation. Example 8 demonstrates the benefits of processing emoji characters. In the sentence "Về Giảng viên UTC2 ❤ ❤" (About UTC2 Lecturers ❤ ❤), if the emotional symbols are not considered, the model will classify the emotion as "other" because, based solely on the text, the sentence does not convey any specific emotion. However, when combined with the ❤ ❤ emojis, the result is classified as "Enjoyment." This highlights the importance of processing emojis to enhance the accuracy of emotion recognition.

Table 2. Experimental results

ID	Sentence (in Vietnamese)	Sentence (in English)	Emotion labels	Result
1	Chất lượng đào tạo rất tốt 👌	Teaching quality very good 👌	Trust	Trust

2	Các lớp học ở các tầng cao cần trang bị thêm rèm ở cửa sổ khi trời nắng 🙏	Classrooms on high floors need to be equipped with curtains on the windows when it is sunny 🙏	Hope	Hope
3	Các phòng học trong dãy e6 đang gặp khó khăn về kỹ thuật và có phòng bị dột mưa 😟 😟	The classrooms in block e6 are having technical difficulties and some rooms are leaking from rain 😟 😟	Sadness	Sadness
4	Các phòng học và khu vực giải trí của trường này thật sự rất tiện nghi	This school's classrooms and recreation areas are truly comfortable	Surprise	Surprise
5	Chất lượng kiến thức phù hợp với ngành nghề học 🙌	The quality of knowledge is appropriate to the field of study 🙌	Trust	Trust
6	Các giảng viên, cố vấn học tập, nhân viên rất tuyệt vời ạ 👌 👌 👌	The lecturers, academic advisors, and staff are wonderful 👌 👌 👌	Enjoyment	Trust
7	86 người chen vào phòng P501C2 thì thực sự là dã man. 😱	86 people squeezing into room P501C2 is truly barbaric 😱	Fear	Surprise
8	Về giảng viên UTC2 ❤ ❤	About UTC2 lecturers ❤ ❤	Enjoyment	Enjoyment

Through these experimental results, it is evident that properly aligning text with emotional symbols is crucial. Sentences with sarcastic tones accompanied by conflicting emotional symbols can also lead to errors in the model's predictions. Furthermore, numerical values and extraneous characters in the text should be handled appropriately to improve the model's accuracy.

3. Conclusions

This research presents an effective methodology for the recognition of various learner emotions through the integration of the PhoBERT model and emojis. The findings demonstrate that this approach not only enhances the accuracy of emotion detection but also augments the visual and interactive aspects of communication. Implementing this method within educational environments holds significant potential, facilitating educators in understanding and responding to student emotions with greater speed and precision. Consequently, this contributes to the creation of a more positive and effective learning atmosphere. The amalgamation of PhoBERT and emojis represents a promising advancement in educational technology and has the potential to inspire numerous new applications in intelligent systems focused on emotion analysis and communication.

Acknowledgment. This research is funded by the University of Transport and Communications (UTC) under grant number T2024-CN-KDN-001.

REFERENCES

[1] Devlin J, Chang MW, Lee K & Toutanova K, (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.*

[2] Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D & Stoyanov V, (2019). RoB-ERTa: A Robustly Optimized BERT Pretraining Approach. arXiv preprint arXiv:1907.11692.

[3] Sun C, Huang L & Qiu X, (2019). Utilizing BERT for Aspect-Based Sentiment Analysis via Constructing Auxiliary Sentence. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.*

[4] Yang Z, Dai Z, Yang Y, Carbonell J, Salakhutdinov R & Le QV, (2019). XLNet: Generalized Autoregressive Pretraining for Language Understanding. *Advances in Neural Information Processing Systems,* 33, 1-11.

[5] Nguyen DQ & Nguyen AT, (2020). PhoBERT: Pre-trained language models for Vietnamese. *Proceedings of the 2020 Conference on Empirical Methods, Natural Language Processing*: *Findings.*

[6] Felbo B, Mislove A, Søgaard A, Rahwan I & Lehmann S, (2017). Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion, and sarcasm. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing.*

[7] Barbieri F, Kruszewski G, Ronzano F & Saggion H, (2016). How Cosmopolitan Are Emojis?: Exploring Emojis Usage and Meaning over Different Languages with Distribution-al Semantics. *Proceedings of the 2016 ACM on Multimedia Conference*.

[8] Eisner B, Rocktäschel T, Augenstein I, Bosnjak M & Riedel S, (2016). emoji2vec: Learning Emoji Representations from their Description. *Proceedings of The Fourth International Workshop on Natural Language Processing for Social Media.*

[9] Pham QH, Nguyen VA, Doan LB, Tran NN & Thanh TM, 2020. From Universal Language Model to Downstream Task: Improving RoBERTa-Based Vietnamese Hate Speech Detection. *12th International Conference on Knowledge and Systems Engineering* (KSE), 37-42.

[10] Navyasree G, Saka N, Narayana DRVS, 2019. Using Hashtags to Capture Fine Emotion Categories from Tweets. *International Journal for Innovative Engineering & Management Research*, 8(5), 288-296, SSRN.

[11] Ekman P, Ekman E & Lama D, 2018. *The Ekmans' Atlas of Emotion*. Paul Ekman Group.

[12] Sau TNT & Toanh TQ, 2021. Application of Bert Architecture for Storage Time of Record Classification Problem. *TNU Journal of Science and Technology*, 267(7), 41-49, https://doi.org/10.34238/tnu-jst.3990.