ICIC Express Letters
Part B: Applications
Volume **14**, Number **11**, November **2023**

ICIC International ©2023 ISSN 2185-2766
pp. 1115–1124

SENTIMENT ANALYSIS FOR A LOW-RESOURCE LANGUAGE: A STUDY ON A VIETNAMESE UNIVERSITY

Quoc-Tuan Vo[1], Phuoc Tran[2,*] and Tram Tran[3]

[1]Faculty of Information Technology
University of Phan Thiet
225 Nguyen Thong Street, Phu Hai Ward, Phan Thiet City 800000, Vietnam
vqtuan@upt.edu.vn

[2]Natural Language Processing and Knowledge Discovery Laboratory
Faculty of Information Technology
Ton Duc Thang University
19 Nguyen Huu Tho Street, Tan Phong Ward, District 7, Ho Chi Minh City 700000, Vietnam
*Corresponding author: tranthanhphuoc@tdtu.edu.vn

[3]Office of Science and Technology
Ho Chi Minh City University of Industry and Trade
140 Le Trong Tan Street, Tay Thanh Ward, Tan Phu District, Ho Chi Minh City 700000, Vietnam
tramtt@hufi.edu.vn

Received February 2023; accepted April 2023

Abstract. *Sentiment analysis for an organization's activities has been widely researched and deployed in recent times. Knowing the client's attitude towards the services that the organization provides will help the organization to further develop the services that customers are satisfied with as well as to limit/remove the bad services. In this paper, we perform a sentiment analysis problem on an organization in the educational field of a low-resource language, that is, social sentiment analysis towards University of Phan Thiet, Vietnamese. We mainly focus on two things: 1) building the sentiment corpus of University of Phan Thiet, which is divided into three classes including positive, neutral and negative; 2) using deep learning algorithms such as LSTM, BERT, DistilBERT and PhoBERT to experiment on this corpus. Experimental results show that PhoBERT gives the highest results with F1-score reaching 89.68%.*
Keywords: Sentiment analysis, Educational data mining, Text classification, BERT, PhoBERT

1. **Introduction.** Universities in general as well as Vietnamese universities in particular actually play the role of a company providing educational services; students/parents/partners are customers using their educational services. Any university that provides good services and satisfies students/parents will attract many excellent domestic and international students to study. With this, the university is growing, the staff in the university will have better income, and stick with the university for a longer time.

To do this, the university needs to know as soon as possible its strengths and weaknesses. From there, the university will further promote the good sides, and at the same time overcome or eliminate the weak sides. Where does this information on strengths/weaknesses come from? Obviously, it does not come from the university's subjective opinion, but must be an objective opinion from the people using the services provided by the university. Those are students, students' parents, and external partners (let us call these objects external users). Therefore, it is essential that the university collects the opinions/attitudes of the external users of its services, and analyzes whether that opinion is good or bad, so that the university can make the next decision. University of Phan

DOI: 10.24507/icicelb.14.11.1115

Thiet (UPT) is a young university, established in 2009, located in Phan Thiet City, Binh Thuan Province, Vietnam. This is a beautiful coastal city, a famous tourist area at home and abroad. However, because it is a young university and is not located in the two largest cities of Vietnam (Ho Chi Minh City and Hanoi City), attracting domestic and foreign students to study at the university is a big challenge for UPT leadership. The best way to promote the university is from students who have studied or are studying at the university as well as the university's partners. The university needs to capture information from students, parents, and partners in order to have an appropriate treatment plan.

UPT has two main channels to collect evaluation information from the external users: the first one is the forms for manual collection, and the second one is online channels for automatic collection. For the first channel, the university will survey, and collect opinions about the quality of training, facilities, training programs, student care, etc. at the end of the semester. In addition, the university also consults students, alumni, businesses, and students' parents on job fairs, graduation holidays, and so on. For the second channel, UPT has a UPT's fan page, faculties' fan pages, and UPT confessions. Through these channels, students will comment on teaching activities, facilities, community activities and other activities. From the information obtained from these two channels, the university leadership will analyze the strengths and weaknesses, and then make the next decision. This method is also good, but it is quite time-consuming and labor-intensive, not keeping up with the reaction of public opinion, especially bad information.

In this paper, we propose to use advanced machine learning models to analyze the sentiment of external users towards UPT. To do this, we first collect the sentiment corpus of the external users, mainly from the two channels above. This corpus will be manually labeled, consisting of three labels including positive, negative, and neutral. Then, we use popular sentiment analysis models to conduct experiments on this corpus, including five models, namely LSTM, BERT-base, BERT-RCNN, DistilBERT, and PhoBERT. From the experimental results, we will choose the best model to build an application to analyze sentiment of external users for UPT (we will do it in the next study).

The rest of the paper is presented as follows. Section 2 presents some background knowledge related to the problem of sentiment analysis. Steps to analyze external users' sentiments for UPT will be presented in Section 3. Section 4 shows and discusses the results of experiments. Finally, Section 5 summarizes our work and gives main conclusions.

2. **Basic Knowledge.** In this section, we will present some background knowledge, including the concept of sentiment analysis, close related works of sentiment analysis, and models for sentiment analysis problems.

2.1. **Sentiment analysis.** Sentiment analysis is a type of text classification problem to evaluate user emotions. The following three cases are examples of users' feeling, comment, and opinion about UPT.

- Feeling: "Vị trí của trường thuận lợi cho sinh viên đi làm thêm tại các resort, khách sạn 5 sao, giúp đáp ứng và phát triển tiếng Anh" (The location of the school is convenient for students to work part-time at resorts and 5-star hotels, helping to meet and develop English) → Sentiment analysis: belongs to positive class.
- Comment: "máy chiếu quá cũ làm không rõ nét" (Projector too dim in computer room 101) → The projector is too old, so it is not clear.
- Opinion: "Chúng tôi không có ý kiến" (We have no opinion) → Sentiment analysis: belongs to neutral class.

The input of the sentiment analysis problem is a sentence or a short paragraph, and the output is the probabilities of many sentiment classes that we need to determine. In this study, we choose the type of sentiment analysis problem with three classes, namely positive, negative, and neutral.

2.2. **Related works.** In [1], Liu analyzed emotions at document level, sentence level and feature level. Recently, Méndez et al. [2] studied the satisfaction of public transport users in Santiago City, Chile through Twitter. The authors used techniques of text mining, opinion analysis and topic modeling to assess the satisfaction level of public transport users, especially buses. Regarding sentiment analysis in Vietnamese context, Trinh et al. [3] combined dictionary-based opinion analysis methods and machine learning methods to evaluate customer opinions. The authors used features including emotional signs and values of emotions to analyze customer emotions. In [4], Nguyen et al. used emotion dictionaries in specific fields to improve the accuracy of emotion analysis in Vietnamese.

Tran et al. [5] used the term feature selection approach to analyze opinions for Vietnamese texts. This approach mainly uses three classical algorithms, such as Naïve Bayes, decision tree and support vector machine. In [6], Le et al. performed sentiment analysis for low-resource languages using 4,000 manually labeled tweets, including positive, negative, and neutral, calculating 73.2% accuracy with LSTM without normalizer. Regarding the problem of sentiment analysis for the educational field, the first work that can be mentioned is [7]. The authors used the feedback tag to create learning resources for programming courses. In [8], Baradwaj and Pal mined educational data to assess student performance. Using user opinions on the social network Twitter to rate universities was also conducted by Abdelrazq et al. [9]. Particularly for Menaha et al. [10], the authors used emotional analysis models to build a system to exploit student feedback. In [11], Sharma and Jain analyzed student feedback and it also helped create feedback summaries. In [12], Dake and Gyimah analyzed student emotions on a qualitative feedback text after a semester-based course at the College of Education, Winneba. The authors used Naïve Bayes, Support Vector Machine (SVM), J48 Decision Tree, and Random Forest algorithms, in which SVM achieved the highest result at 63.79%.

For the sentiment analysis of Vietnamese education, the first work is of Vo et al. [13]. The authors used two approaches including topic classification and sentiment analysis for the Vietnamese education survey system. The feedback from internship places of students, the information about the class, the quality of the thesis, etc. are classified as positive or negative. Nguyen et al. [14] built a corpus of more than 16,000 sentences on student feedback for universities. These feedbacks were also labeled as positive, negative, or neutral with an accuracy of 87.94%. Then, Nguyen et al. [15] relied on the corpus in [14] and used algorithms such as Naïve Bayes, LSTM, Bi-LSTM for their experiments. The authors compared these algorithms on the same dataset, resulting in Bi-LSTM giving the highest results with F1-score of 89.6%.

BERT (Bidirectional Encoder Representations from Transformers) [16] was born, marking a new break-through in the field of Natural Language Processing (NLP). A series of studies based on BERT for sentiment analysis problems have been emerging, and the most recent work of Nguyen et al. [18] is a typical case. The authors proposed a model combining BERT with CNN, LSTM, and RCNN. They experimented the model on a Vietnamese dataset consisting of two positive and negative labels. The result achieved with F1-score is 91.15%. Giang et al. [17] conducted sentiment analysis based on a dataset of 5,000 Vietnamese sentences with student feedback sent back to the school at the end of the semester, the data was manually labeled: positive, neutral and negative, and the results on the vector machine algorithm support are very high of 91.36%.

We used PhoBERT – a version of BERT for Vietnamese – to experiment on UPT corpus (presented in Section 3.2). In addition, we also used some models of [18] and experimented on the UPT corpus to compare and evaluate the models for the best results, as a foundation for further applications.

2.3. **Sentiment analysis models – PhoBERT.** PhoBERT [19] is a pre-training model for Vietnamese language based on RoBERTa architecture introduced in March 2020.

PhoBERT also has two versions including PhoBERT_base with 12 transformers block and PhoBERT_large with 24 transformers block. PhoBERT was trained on approximately 20GB of data including 1GB of Vietnamese Wikipedia corpus and 19GB collected and processed from a 50GB raw dataset[1]. PhoBERT uses VnCoreNLP's RDRSegmenter to segment word for the input data before going through the BPE (Byte Pair Encoding) encoder. PhoBERT eliminated the task of the next sentence prediction and used only the masked language model.

3. Sentiment Analysis Model.

3.1. General model.

Figure 1 shows the general model for the UPT sentiment analysis problem. The model consists of two main processes, namely the training process and the testing process.

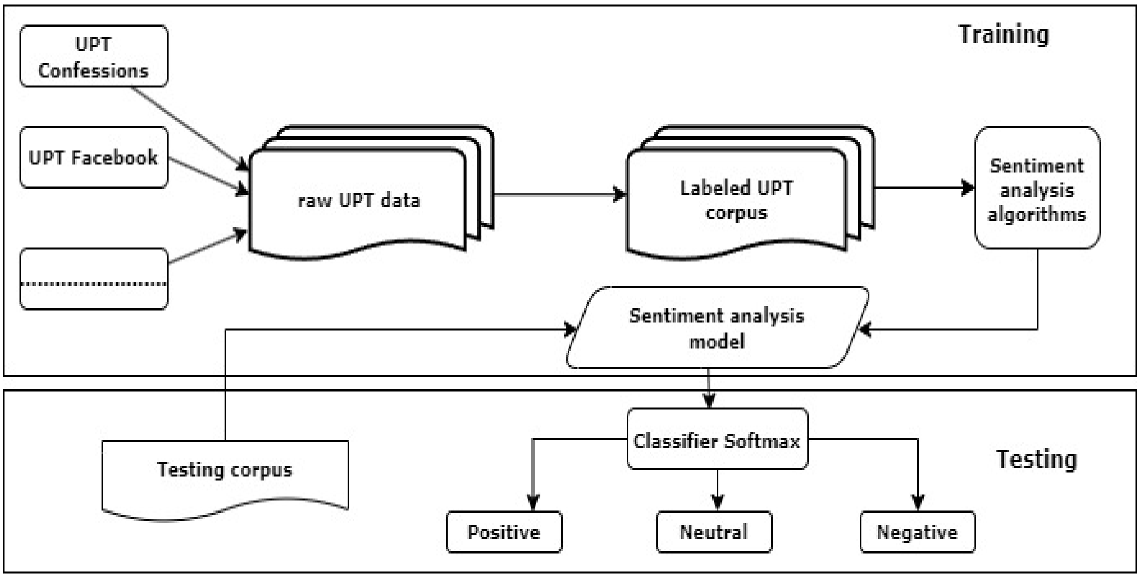

FIGURE 1. Sentiment analysis model for UPT

- Raw UPT data: The raw data set is collected from sources: facebook UPT, facebook faculty, confessions, student feedback on the system when viewing their subject results at the end of each semester, etc.
- Labeled UPT corpus: Processed data includes remove special characters, acronyms, assign positive, negative and neutral labels, etc.

3.2. UPT corpus collection.

3.2.1. *Resources for collection of UPT corpus.* UPT corpus is collected from the following resources:
- UPT fan page, fan pages of faculties, UPT Confessions.
- Evaluations and comments of students from 2016-2019.
- The opinions of business partners in the annual job fair.
- The comments of the student's parents.
- Surveys via Google docs.

3.2.2. *Data.* Data collected from fan pages related to UPT contains a lot of slang terms, symbols, acronyms, etc. So we clean the data as follows.
- Remove HTML tags, excess space, emoji expressions, bullet characters, and repeated characters when they are not alphanumeric.

[1]https://github.com/binhvq/news-corpus

- Correct spelling errors and abbreviations. Here is an example of a sentence with many abbreviations: "Nguyễn Diệp Trân Trân ra nhìn xe mà k pk ns j". We do both automatic and manual editing for cases like this. This sentence is edited to "Nguyễn Diệp Trân Trân ra nhìn xe mà không biết nói sao" ("Nguyen Diep Tran Tran looked at the car without knowing what to say").
- Replace the characters "_" with a space character.

3.2.3. *Word segmentation.* Unlike western languages (such as English), spaces in Vietnamese do not define word boundaries. Therefore, word segmentation is often performed before conducting other tasks of the NLP problem. For example, the word "sinh viên" (student) is made up of two syllables "sinh" and "viên" ("sinh" in "nơi sinh" (place of birth) and "viên" in "viên thuốc" (tablet), "viên" in "công viên" (park), "viên" in "nhân viên" (employee)), but when these syllables come together, it forms the word "sinh viên" (student), whose meaning is not related to the syllables' meanings that make it up. To solve this problem, we use the VnCoreNLP[2] toolkit to segment words for this corpus. Using VnCoreNLP to segment from "Khuôn viên trường Đại học Phan Thiết đẹp thoát mát" (University of Phan Thiet campus is beautiful and airy) results: "Khuôn_viên trường Đại_học Phan_Thiết đẹp thoáng mát".

3.2.4. *Sentiment labeling.* Data labeling process: my team (including Le Trung Thanh, Luong Quoc Vu and I) participated in data labeling. We learned the labeling rule, each of us labeled 100 sentences several times until consensus reached about 85% or more. Based on the above results, we continued to label another 3,000 sentences, two of our teammates labeled 1,500 sentences each. After the labeling process was completed, we synthesized and evaluated the quality of the dataset according to Cohen's Kappa K[3] consensus measure according to the formula $K = \frac{P_0 - P_1}{1 - P_e}$, where P_0 is the observed consensus of 95.27%, P_e is the expected consensus of 46.19%, and K is the consensus of 91.2%. We label this corpus manually. Table 1 shows three cases corresponding to three different emotional classes.

TABLE 1. Some comment sentences in the UPT corpus

ID	Sentences	English meaning	Sentiment labels
1	trường nằm ở khu_vực du_lịch , thuận_lợi cho sinh_viên học_tập và thực_hành ngành du_lịch	The school is located in a tourist area that is convenient for students to study and practice tourism	Positive
2	Dựa vào nghành_nghề mà em chọn chúng_tôi sẽ tư_vấn phù_hợp	Depending on the profession you choose, we will advise accordingly	Neutral
3	Máy_chiếu quá cũ làm không rõ nét	The projector is too old so it is not clear	Negative

- Positive: The sentences express the external user's satisfaction, praise, encouragement for UPT. For example, the sentence "trường nằm ở khu_vực du_lịch, thuận_lợi cho sinh_viên học_tập và thực_hành ngành du_lịch" in Table 1 is labeled as Positive.
- Neutral: The sentences do not imply any emotion, or incomplete sentences, or unclear in meaning, or general meanings. For example, the sentence "Dựa vào nghành_nghề mà em chọn chúng_tôi sẽ tư_vấn phù_hợp" in Table 1 is labeled as Neutral.
- Negative: The sentences express dissatisfaction of external users for UPT. For example, the sentence "Máy_chiếu quá cũ làm không rõ nét" is labeled as Negative.

[2]https://github.com/vncorenlp/VnCoreNLP
[3]https://en.wikipedia.org/wiki/Cohen%27s_kappa

There are some cases that make it difficult to label, which are a sentence that has both a negative and a positive opinion; the sentence type often has linking words such as "nhưng" (but), "tuy nhiên" (however), "mặc dù" (although), "dù" (although), "tuy rằng" (although), and so on. In this case, we choose the clause with stronger polarity to label. For example, the sentence "Thư viện, phòng máy, và phòng học cần phải nâng cấp; tuy nhiên, môi trường xung quanh rất ổn" (Library, computer labs, and classrooms need to be upgraded; however, the surroundings are very good) is labeled as negative, even though in this sentence, there is a sub-sentence "môi trường xung quanh rất ổn" (the surroundings are very good) with positive emotion.

3.2.5. *Final UPT corpus.* After the preprocessing step, we get a total of 6,000 labeled sentences. The details of this corpus are described in Table 2.

TABLE 2. Some comment sentences in the UPT corpus

	Negative	**Neutral**	**Positive**	**Total**
Number of sentences	2,128	1,760	2,112	6,000
Percentage	35.47%	29.33%	35.20%	100.00%
Average length of sentences	21.73	10.66	19.62	17.73

3.3. **Sentiment analysis models for UPT.** In this step, we use machine learning models such as LSTM, BERT, DistilBERT and PhoBERT respectively to train UPT corpus, producing UPT sentiment analysis models. Then, for an input sentence, these sentiment analysis models will give positive, negative or neutral results.

Figure 2 illustrates a sentence that is analyzed using the PhoBERTbase model. The steps are similar for other models.

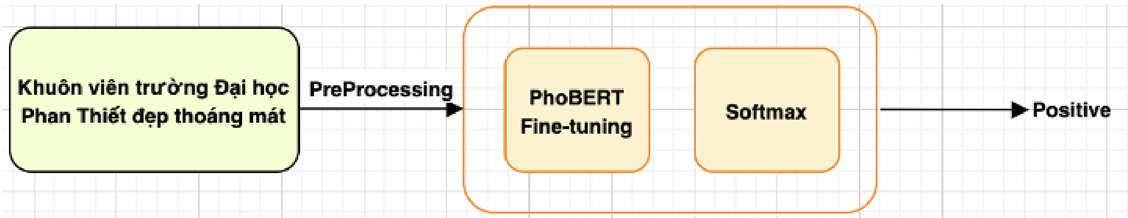

FIGURE 2. Example of sentiment analysis problem

The sentiment analysis process of the sentence in Figure 2 is specifically demonstrated through four steps as follows.

- Step 1: Use the VnCoreNLP library to perform the word segmentation. For example, the sentence "khuôn viên trường Đại học Phan Thiết đẹp thoáng mát" (The campus of University of Phan Thiet is beautiful and cool) is word-segmented as "khuôn_viên trường Đại_học Phan_Thiết đẹp thoáng mát".
- Step 2: Add the token [CLS] to mark the beginning of the sentence and [SEP] to mark the end of the sentence.
- Step 3: Use the BPE algorithm to put the input sentence as a subword and map the subword to the index form in the dictionary.
- Step 4: Add the result of Step 3 to the PhoBERT fine-tuning model. The output is a feature vector, continue to use the softmax function to calculate the output probability, and use argmax function to select the maximum value from softmax function to get the final value.

The formula of the Softmax function is as follows (Equation (1)):

$$a_i = \frac{\exp(z_i)}{\sum_{j=1}^{C} \exp(z_j)}, \ \forall i = 1, 2, \ldots, C \tag{1}$$

where z_i consists of values that are elements of the input vector X, it can be a negative or positive number, so $\exp(z_i)$ returns a result in the range 0 to 1. In Figure 3, the Softmax function returns three classes, including class 0: 0.1; class 1: 0.0; class 2: 0.9. Next, the argmax function finds the maximum value and it is 0.9 (positive label).

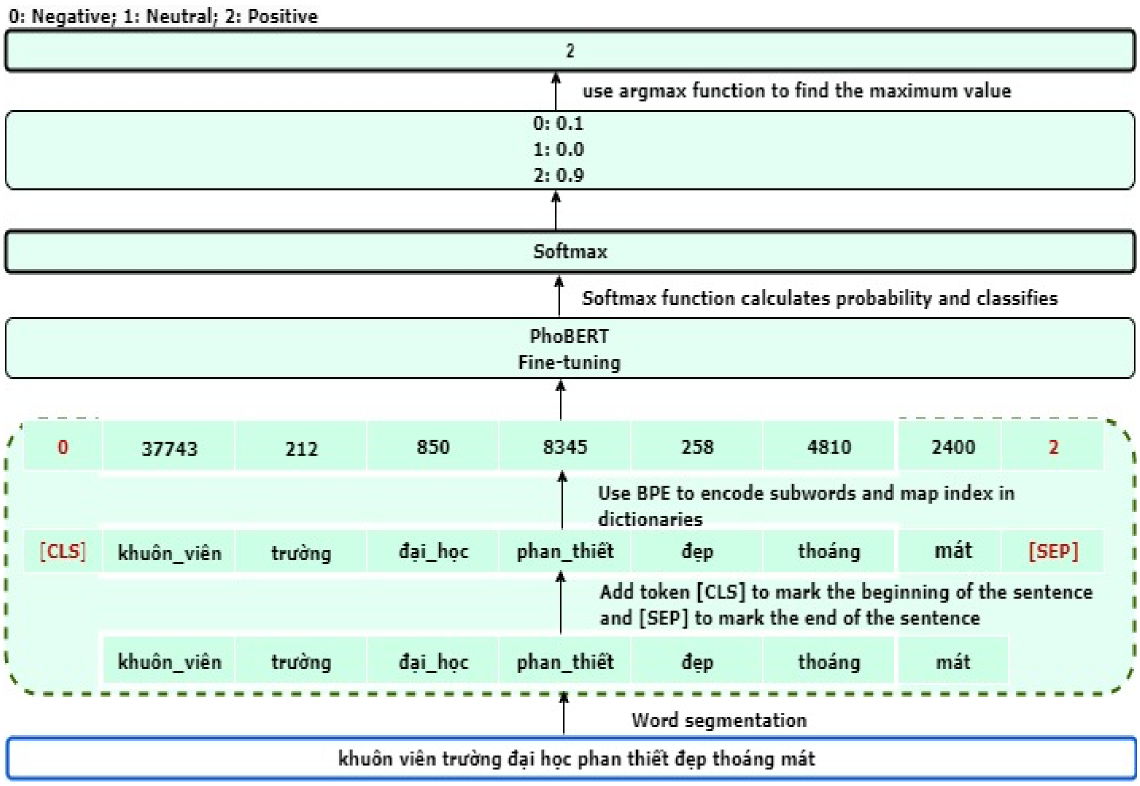

Figure 3. Detailed simulation of the label prediction of the sentence in Figure 2

4. Experiment.

4.1. Experimental corpora.

We divided the UPT corpus of 6,000 sentences into three corpora including training, development and testing corpus. We used 70% of the sentences for training, 10% of the sentences for developing, and the remaining 20% of the sentences for testing.

4.2. Evaluation methods.

We use Precision, Recall, F1-score to evaluate the experimental results.

$$\begin{aligned} \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}} \\ \text{Recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}} \\ \text{F1-score} &= 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \end{aligned} \tag{2}$$

- True Positive (TP): The number of points of the positive class is correctly classified as positive.

- True Negative (TN): The number of points of the negative class is correctly classified as negative.
- False Positive (FP): The number of points of the negative class is mistakenly classified as positive.
- False Negative (FN): The number of points of the positive class is mistakenly classified as negative.

4.3. **Parameters of the models.** We implement our experiments on Google Colab environment with 16GB Tesla V100 GPU, Python programming language, Huggingface library, Pytorch framework. Table 3 presents the parameters of the models.

TABLE 3. Parameters of the models

	Bacth size	**Epoch**	**lr**	**Embedding**
LSTM	32	250	0.001	128
DistilBERT	32	10	5e-5	
PhoBERT	32	10	1e-5	
BERT-base, BERT-RCNN [18]	16	10	2e-5	

4.4. **Experimental results.** Table 4 shows experimental results of the models.

TABLE 4. Experimental results

Models	**Precision**	**Recall**	**F1-score**
LSTM	72.23	72.03	72.11
DistilBERT	82.83	82.28	82.40
BERT-base [18]	86.71	86.18	86.25
BERT-RCNN [18]	87.05	86.87	86.83
PhoBERT	**89.89**	**89.71**	**89.68**

4.5. **Discussion.** From the experimental results in Table 3 and Table 4, we see that the BERT models give better results than the LSTM model, in which, PhoBERT gives the best results. The PhoBERT model is taken from Facebook's RoBERTa but trained on word-segmented Vietnamese data. Therefore, the performance of PhoBERT is higher than that of other BERT models as well as the LSTM model. Table 5 and Table 6 show three examples from the testing corpus.

From the three examples in Table 5 and Table 6, we see that the PhoBERT model gives exactly two cases. Although the LSTM model has the lowest final results in all evaluation criteria (Table 4), in these three cases, the LSTM gives correct results in two cases. Both LSTM and PhoBERT models give correct results in two cases, but their correct results

TABLE 5. Three sentences in the testing corpus

ID	**Label**	**Vietnamese**	**English**
1	2	các hoạt động để tích điểm công ích dễ dàng, không gây khó khăn quá cho sinh viên	activities for accumulating public points are easy, not too difficult for students
2	2	tận tâm gọi những sinh viên không tập trung để trả lời câu hỏi	conscientiously call on unfocused students to answer the question
3	1	về việc thầy có các cách kỷ luật hay đối với sinh viên đi trễ, vắng buổi	about the teacher has good disciplinary methods for students who are late or absent

Table 6. Experimental results of the three sentences in Table 5

Models	Sentence 1	Sentence 2	Sentence 3
LSTM	0	**2**	**1**
DistilBERT	0	1	**1**
BERT-base	0	0	0
BERT-RCNN	0	0	2
PhoBERT	**2**	**2**	**2**
Reference labels	**2**	**2**	**1**

are in different sentences. PhoBERT is correct in cases 1 and 2; LSTM is correct in cases 2 and 3. The remaining models are correct in only one case.

The reason for the confusion of the models in the labeling for the three sentences in Table 5 is because in all these three sentences, there are both positive, negative and neutral words. For example, in sentence 1, even though it is properly labeled as positive, it contains words/phrases with negative meanings such as "khó khăn" (difficult), "khó khăn quá" (too difficult), so all models except PhoBERT assign negative labels to this sentence. For sentence 2, its correct label is positive. However, there is a word "không tập trung" (unfocused) that has negative meaning, so some models give negative or neutral results.

The last case is the difficult one. Because the correct manual labeling for this case is also controversial, some consider this statement neutral, while others consider it to be positive. In fact, if we omitted the word "về việc" (about) from the sentence "về việc thầy có các cách kỷ luật hay đối với sinh viên đi trễ, vắng buổi", this sentence would certainly be positive. Because there is the word "về việc" at the beginning of the sentence, it makes this sentence no longer mean to express emotions, but it becomes a sentence referring to the "discipline" of the teacher, not specific emotions. Therefore, when manually labeling, we label this sentence as neutral. Because there are words/phrases in this sentence that have both positive meanings ("các cách kỹ luật hay": good disciplinary methods) and negative meanings ("đi trễ": late, "vắng buổi": absent), many models label this sentence as either positive (PhoBERT, BERT-CNN) or negative (BERT-base). The LSTM and DistilBERT models give correct results in this case.

5. **Conclusion.** In this paper, we have initially performed the sentiment analysis problem on a specific field of a low-resource language, which is the education field of University of Phan Thiet, Vietnam. To do this, we must first perform corpus construction for the sentiment analysis problem. We have constructed the corpus semi-automatically: The corpus is automatically extracted from UPT's fanpages, cleaned and extracted automatically; however, data labeling is done manually. The corpus of 6,000 sentences is experimented on many machine learning models such as LSTM, BERT, DistilBERT, and PhoBERT. The experimental results show that the BERT models have superior results compared to other models, in which PhoBERT has the highest performance. This result is the premise for us to continue to develop the sentiment analysis system for UPT in the next work.

In addition to the results mentioned above, the work still has limitations that need to be overcome, for example, the corpus with only 6,000 sentences is not much, it needs to be developed more. Resources for data collection are currently poor, mainly from UPT forums. It is necessary to survey and collect from many other resources such as Websites of Phan Thiet Province, the Ministry of Education and Training, and comments from users on those websites. Once more data is available, more experimentation, we will discover many interesting results to serve other studies.

REFERENCES

[1] B. Liu, Sentiment analysis and subjectivity, *Handbook of Natural Language Processing*, pp.627-666, 2010.

[2] J. T. Méndez, H. Lobel, D. Parra and J. C. Herrera, Using Twitter to infer user satisfaction with public transport: The case of Santiago, Chile, *IEEE Access*, vol.7, pp.60255-60263, 2019.

[3] S. Trinh, L. Nguyen and M. Vo, Combining lexicon-based and learning-based methods for sentiment analysis for product reviews in Vietnamese language, *Computer and Information Science*, pp.57-75, 2018.

[4] H. N. Nguyen, T. V. Le, H. S. Le and T. V. Pham, Domain specific sentiment dictionary for opinion mining of Vietnamese text, *MIWAI 2014*, pp.136-148, 2014.

[5] B. Tran, C. Haruechaiyasak and V. Sornlertlamvanich, Vietnamese sentiment analysis based on term feature selection approach, *The 10th International Conference on Knowledge, Information and Creativity Support Systems (KICSS2015)*, 2015.

[6] T. A. Le, D. Moeljadi, Y. Miura and T. Ohkuma, Sentiment analysis for low resource languages: A study on informal Indonesian tweets, *Proc. of the 12th Workshop on Asian Language Resources (ALR12)*, pp.123-131, 2016.

[7] S. Cummins, L. Burd and A. Hatch, Using feedback tags and sentiment analysis to generate sharable learning resources investigating automated sentiment analysis of feedback tags in a programming course, *2010 10th IEEE International Conference on Advanced Learning Technologies*, pp.653-657, 2010.

[8] B. K. Baradwaj and S. Pal, Mining educational data to analyze students' performance, *International Journal of Advanced Computer Science and Applications*, pp.36-69, 2011.

[9] A. Abdelrazq, S. Jeschke, A. Richert, C. Tummel and D. Janssen, Sentiment analysis of social media for evaluating universities, *The 2nd International Conference on Digital Information Processing, Data Mining, and Wireless Communications (DIPDMWC2015)*, pp.49-62, 2015.

[10] R. Menaha, R. Dhanaranjani, T. Rajalakshmi and R. Yogarubini, Student feedback mining system using sentiment analysis, *International Journal of Computer Applications Technology and Research*, vol.6, pp.51-55, 2017.

[11] N. Sharma and V. Jain, Evaluation and summarization of student feedback using sentiment analysis, *Advanced Machine Learning Technologies and Applications*, vol.1141, pp.385-396, DOI: 10.1007/978-981-15-3383-9_35, 2021.

[12] D. K. Dake and E. Gyimah, Using sentiment analysis to evaluate qualitative students' responses, *Education and Information Technologies*, vol.28, pp.4629-4647, 2023.

[13] H. T. Vo, H. C. Lam, D. D. Nguyen and N. H. Tuong, Topic classification and sentiment analysis for Vietnamese education survey system, *Asian Journal of Computer Science and Information Technology*, vol.6, no.3, pp.27-34, 2016.

[14] K. V. Nguyen, V. D. Nguyen, P. X. V. Nguyen, T. T. H. Truong and N. L. T. Nguyen, UIT-VSFC: Vietnamese students' feedback corpus for sentiment analysis, *2018 10th International Conference on Knowledge and Systems Engineering (KSE)*, pp.19-24, 2018.

[15] P. X. V. Nguyen, T. T. T. Hong, K. V. Nguyen and N. L. T. Nguyen, Deep learning versus traditional classifiers on Vietnamese students' feedback corpus, *2018 5th NAFOSTED Conference on Information and Computer Science (NICS)*, pp.75-80, 2018.

[16] J. Devlin, M. W. Chang, K. Lee and K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, *Proc. of NAACL-HLT 2019*, pp.4171-4186, 2019.

[17] N. T. P. Giang, T. T. Dien and T. T. M. Khoa, Sentiment analysis for university students' feedback, *Advances in Information and Communication: Proc. of the 2020 Future of Information and Communication Conference (FICC)*, pp.55-66, 2020.

[18] Q. T. Nguyen, T. L. Nguyen, N. H. Luong and Q. H. Ngo, Fine-tuning BERT for sentiment analysis of Vietnamese reviews, *2020 7th NAFOSTED Conference on Information and Computer Science (NICS)*, pp.302-307, 2020.

[19] D. Q. Nguyen and A. T. Nguyen, PhoBERT: Pre-trained language models for Vietnamese, *arXiv.org*, arXiv: 2003.00744, 2020.