Language-oriented Sentiment Analysis based on the Grammar Structure and Improved Self-attention Network

Hien D. Nguyen[1,2,*][a], Tai Huynh[3,4,*], Suong N. Hoang[4][b], Vuong T. Pham[5] and Ivan Zelinka[6][c]

[1]*Faculty of Computer Science, University of Information Technology, Ho Chi Minh City, Vietnam*
[2]*Vietnam National University, Ho Chi Minh City, Vietnam*
[3]*Ton Duc Thang University, Ho Chi Minh City, Vietnam*
[4]*Kyanon Digital, Vietnam*
[5]*Faculty of Information Technology, Sai Gon University, Ho Chi Minh City, Vietnam*
[6]*Technical University of Ostrava (VŠB-TU), Czech Republic*

* *Equal contribution by Hien D. Nguyen and Tai Huynh*

Keywords: Sentiment Analysis, Sentiment Classification, Vietnamese, Self-attention, Transformer, Natural Language Processing.

Abstract: In the businesses, the sentiment analysis makes the brands understanding the sentiment of their customers. They can know what people are saying, how they're saying it, and what they mean. There are many methods for sentiment analysis; however, they are not effective when were applied in Vietnamese language. In this paper, a method for Vietnamese sentiment analysis is studied based on the combining between the structure of Vietnamese language and the technique of natural language processing, self-attention with the Transformer architecture. Based on the analysing of the structure of a sentence, the transformer is used to process the word positions to determine the meaning of that sentence. The experimental results for Vietnamese sentiment analysis of our method is more effectively than others. Its accuracy and F-measure are more than 91% and its results are suitable to apply in practice for business intelligence.

1 INTRODUCTION

Sentiment analysis (SA) is one of the subfields of Computational Linguistics and Natural Language Processing (NLP) (Gamal et al., 2019). In the businesses intelligence, the sentiment analysis makes the brands understanding the sentiment of their customers (Rokade and Kumari, 2019). They can know what people are saying, how they're saying it, and what they mean. The sentiment of customer sentiment can be found in tweets, comments, reviews, or other places where people mention the brands.

In the current era, social network is a popular platform for communication and interaction (Beigi, 2016). Many people found innovative information on social network and due to that social network is the important data source. SA is also used to detect the influencer on the social network for the influencer marketing (Huynh et al, 2019).

Vietnamese is a language isolate (Nguyen et al., 2006). The meaning of a sentence belongs to the way for organizing of its predicates (Clark, 1974). In other words, the information about word positions contribute the sentence meaning and grammatical meaning. The analysing on the Vietnamese sentence has to combine the studying of the grammar structure.

Some machine learning-based approaches have been studied to analysis the sentiment of a Vietnamese sentence.

CountVectorizer (Irfan et al., 2015) and Term Frequency–Inverse Document Frequency (Tf-idf) (Aggarwal, 2011) are used for word representations. However, they cannot analysis the positions of words in a sentence, so their results are not exactly. Support

[a] https://orcid.org/0000-0002-8527-0602
[b] https://orcid.org/0000-0002-3354-013X
[c] https://orcid.org/0000-0002-3858-7340

Nguyen, H., Huynh, T., Hoang, S., Pham, V. and Zelinka, I.
Language-oriented Sentiment Analysis based on the Grammar Structure and Improved Self-attention Network.
DOI: 10.5220/0009358803390346
In *Proceedings of the 15th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE 2020)*, pages 339-346
ISBN: 978-989-758-421-3

Vector Machine (Joachims, 1998) and Naïve Bayes (Irfan et al., 2015) are used as classifiers. However, those methods did not mention to the structure of a sentence, so their results are not suitable in the practice.

In (Krouska et al., 2017, Troussas et al., 2016), authors present five well-known learning-based classifiers (Naïve Bayes, Support Vector Machine, k-Nearest Neighbor, Logistic Regression and C4.5) and a lexicon-based approach (SentiStrength) to analysis the sentiment on Twitter. However, it only studies on English.

Besides, some types of recurrent neural networks (RNNs), such as long short-term memory (LSTM) (Hochreiter, 1997, Cheng et al., 2016), Bi-Directional LSTM (biLSTM) (Schuster and Paliwal, 1997) or gated recurrent unit (GRU) (Chung et al., 2014), are very complex and take a long time to solve the problem about sentiment analysis on Vietnamese.

The sentiment analysis for Vietnamese was researched in (Nguyen et al., 2014). This study investigated the task regarding both Support Vector Machine (SVM) model and linguistics feature aspects which is an annotated corpus for sentiment classification extracted from hotel reviews in Vietnamese. However, this method is not designed based on the grammar structure, so some sentences cannot be determined accurately.

Self-attention has been used successfully in a variety of tasks including reading comprehension, abstractive summarization, textual entailment and learning task-independent sentence representations (Zhou et al., 2018). The Transformer (Vaswani et al., 2017) is the transduction model based on self-attention to compute representations of its input and output without using sequence aligned RNNs or convolution. In (Hoang et al., 2019), authors study sentiment analysis of product reviews in Vietnamese by using Self-attention neural networks. However, that study does not mention to the structure of Vietnamese sentence in the analysing, so its results are not exactly and suitable the practical requirements.

In this paper, the method for Vietnamese sentiment analysis is proposed. This method is used to determine the sentiment of a sentiment sentence including positive, negative or neutral. The structures of a Vietnamese sentence are studied. Based on those structures, the meaning of this sentence is analysed by using the self-attention neural network architecture Transformer. Besides, the layer of Squeeze and Excitation (Hu et al., 2018) is also used to recalibrate features in the process. The sentences will be analysed to determine whether they are positive, negative or neutral.

The experimental results show that our method being more effective than other in Vietnamese sentiment analysis. Its accuracy and F-measure are more than 91% and its results are suitable to apply in practice for business intelligence.

The next section presents some techniques of the Transformer. Section 3 presents the method for Vietnamese sentiment analysis. That method uses the improved architecture of self-attention with transformer on the structure of the sentences in Vietnamese to determine their meaning. Section 4 described the experimental results. The last section concludes the main results in this paper.

2 SELF-ATTENTION NETWORK

Scaled Dot-Product Attention: Let s_{i-1} be a query vector q, and h_j is duplicated with one is key vector k_j and the other is value vector v_j (in current NLP work, the key and value vector are frequently the same, there for h_j can be considered as k_j or v_j).

$$c = \sum_{j=1}^{n} a_j v_j \tag{1}$$

$$\text{where } a_j = \frac{\exp(e_j)}{\sum_{k=1}^{n} \exp(e_k)}, \text{ and } e_j = \alpha(q, k_j) = \frac{q.k_j^T}{\sqrt{d_{\text{model}}}} \tag{2}$$

$$(1 \le j \le n)$$

d_{model} is the dimension of input vectors or k vector (q, k, v have the same dimension as input embedding vector)

Self-attention is a mechanism to apply Scaled Dot-Product Attention to every token of the sentence for all others.

For every token in sentence, three vectors *Query*, *Key*, *Value* are created by using a linear feed-forward layer as a transformation, then the attention mechanism is applied to get the context matrix. However, this process is very slow, so we consider three matrices Q, K, V:

- Q is a matrix containing all the query vectors, $Q = [q_1, q_2,..., q_n]$ with q_i is a query vector.
- K is a matrix containing all the key vectors, $K = [k_1, k_2, ..., k_n]$ with k_i is a key vector.
- V is a matrix containing all the key vectors, $V = [v_1, v_2, ..., v_n]$ with v_i is a value vector.

Thus, we have:

$$Attention(Q,K,V) = \text{softmax}\left(\frac{Q.K^T}{\sqrt{d_{\text{model}}}}\right).V \quad (3)$$

Multi-head Attention performs the attention h times with (Q, K, V) matrices of the dimension d_{model}/h. Each head is a time for applying Attention. For each head, the (Q, K, V) matrices are uniquely projected with the dimensions d_{model}/h. Self-attention mechanism is performed to yield an output of the same dimension d_{model}/h. After all, the outputs of h heads are concatenated, and applied a linear projection layer once again. The formula for this process is as follows:

$$MultiHead(Q,K,V) = Concat\left(head_1, head_2, ..., head_h\right).W^O$$

$$\text{where } head_i = \left(Q.W^O, K.W^O, V.W^O\right) \quad (4)$$

3 METHOD FOR VIETNAMESES SENTIMENT ANALYSIS

In this section, the method for analysing the sentiment of a Vietnamese sentence is proposed. The sentences will be analysed to determine whether they are positive, negative or neutral.

Firstly, the structures of a Vietnamese sentence are studied. Because the scope of this study is the evaluation comments for a product on the social network, there are two kinds of declarative sentence were mentioned: positive and negative sentence.

Secondly, based on those structures, the meaning of this sentence is analysed by using the self-attention neural network architecture Transformer. Because the meaning of a Vietnamese sentence belongs to the positions of words, our method is added the layer determining the word positions into the processing the transformer. Besides, the layer of Squeeze and Excitation (Hu et al., 2018) is also used to recalibrate features in the process.

3.1 Structure of a Vietnamese Sentence

Vietnamese is a language isolate. The structure of a normal sentence of Vietnamese includes subjectum (or thema) and praedicatum (or rhema). Subjectum is the direct factor of a sentence describing the scope of thing which is mentioned in the second direct factor - praedicatum (Cao, 2017).

There are three frequent sentence types: declarative, interrogative, and imperative. The declarative is subject to judgments of truth and falsehood (Cao, 2017). The interrogative elicits a verbal response from the addressee. The imperative indicates the speaker's desire to influence future events. In the problem about sentiment analysis, we only need to determine whether a sentence is positive, negative or neutral; thus, in the scope of this paper, we only mention to the declarative sentence type.

The structure of a single declarative sentence in Vietnamese is shown in Fig.1:

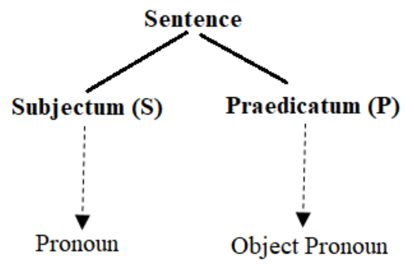

Figure 1: Structure of a single declarative sentence in Vietnamese.

Definition 1: *Kinds of the structure of a positive sentence*

A single positive declarative sentence in Vietnamese has the foundation structure:

<Sentence> = <S> <P>

It is classified as Table 1.

Table 1: Kinds of the structure of a positive sentence.

Kinds	Variants
P is <noun>: <Sentence> = <S><noun>	<S> "là" <noun>
	<S> <quantity> <noun>
	<S> <comparative word> <noun>
	<S> <word of kind> <noun>
P is <verb>: <Sentence> = <S> <verb>	<S><verb><object pronoun>
	<S_1><verb><S_2><P_2>
	<S_1><verb><P_2><object>
P is <adjective>: <Sentence> = <S> <adj.>	<S> "thì" <adj.>
P is <noun><adj.> with <noun> belongs to <S> <Sentence> = <S><noun><adj.>	

Definition 2: *Kinds of the structure of a negative sentence*

A single negative declarative sentence in Vietnamese has the foundation structure:

<Sentence> = <S><negative word><P>

It is classified as Table 2.

Table 2: Kinds of the structure of a negative sentence.

Kinds	Variants
P is <noun>: <Sentence> = <S> <negative word> <noun>	<S><negative word> "là" <noun>
	<S> <negative word> <quantity> <noun>
	<S> <negative word> <comparative word> <noun>
	<S> <negative word> <word of kind> <noun>
P is <verb>: <Sentence> = <S> <negative word><verb>	<S><negative word> <verb> <object pronoun>
	<S_1><negative word> <verb> <S_2><P_2>
	<S_1> <negative word> <verb> <P_2><object>
P is <adjective>: <Sentence> = <S> <negative word> <adj.>	
P is <noun><adj.> with <noun> belongs to <S> <Sentence> = <S><noun><adj.>	

In a Vietnamese declarative sentence, each word has to been appeared orderly. Although two sentences have the same referent, "same referent" means they both describes an objectivity fact, they are not identity about the meaning. The meaning of a sentence belongs to the way for organizing of its predicates. In other words, the information about word positions contribute the sentence meaning and grammatical meaning.

Some characteristics of an isolate language, especially Vietnamese, for learning context are as follows:

- In linguistic activities, words do not change their morphemes. Grammatical meanings are not included in words.
- Formal word, word position and word order clarify the grammatical relationship as well as the grammatical meaning of words and sentences. Example: Add the formal words "sẽ" (will) or "đang" (_ing) before "học" (study) will change the tense of the action. Another example of reversing words also changes the meaning of grammar, for example "chân bàn" (leg of table) and "bàn chân" (foot).
- The lines between syllables, morphemes and words are not clear. Example: In Vietnamese "nhà" is a morpheme, and also is a word.

The main point of this research is around the importance of word position information to contribute sentence meanings and grammatical meanings.

3.2 Pre-processing Method

Datasets will be gone through a pre-processing pipeline of the text documents. Some available research, such as sentence segmentation, normalize the text, word segmentation and noise cleaning, were mainly used to do this pipeline automatically.

Sentence segmentation is a procedure to split a paragraph into sentences. Then, each sentence will be text normalized.

In the text normalization, the input will be low cased. Next, all the links, phone numbers and email addresses were replaced by "*urlObj*", "*phonenumObj*" and "*mailObj*", respectively.
Finally, words tokenizer from Underthesea (2019) for Vietnamese was also applied. The input text will be split into words, phrases, or other meaningful parts, namely tokens.

3.3 Word Embedding

The fastText (2019) is used for word embedding. In many cases, users may type a wrong word accidentally or intentionally. fastText deals with this problem very well by encoding at the character level. In case having a wrong word, very rare words or out-of-vocabulary words, fastText still can represent them with an embedding vector that most similar to word met in trained sentences.

There had been no fastText pre-trained model for Vietnamese spoken language. Therefore, we trained fastText for Vietnamese vocabulary as embedding pre-trained weights from a corpus over 70,000 documents of multi-products reviews crawled from e-commerce sites mentioned above with no label. Rare words that occur less than 5 times in the vocabulary were removed. Embedding size is 384. After training, we have 5,534 vocabularies in total.

3.4 Sentiment Analysis in Vietnamese

In original architecture of Transformer, the position encoding for a word is summed with Context encoding from pre-trained fastText model (with same dimensions of features). After this process, the outputs were applied a linear projection to create three vectors *Q* (query), *K* (key), *V* (value) as input for Multi-head Attention layer:

$$A = W_A.(C + P) \quad (5)$$

where *A* is one of the three vectors *Q, K*, or *V*, as inputs of Multi-head Attention, which were mentioned in Section 2. *C* is the context encoding with d_{model} dimension, and *P* is the position encoding

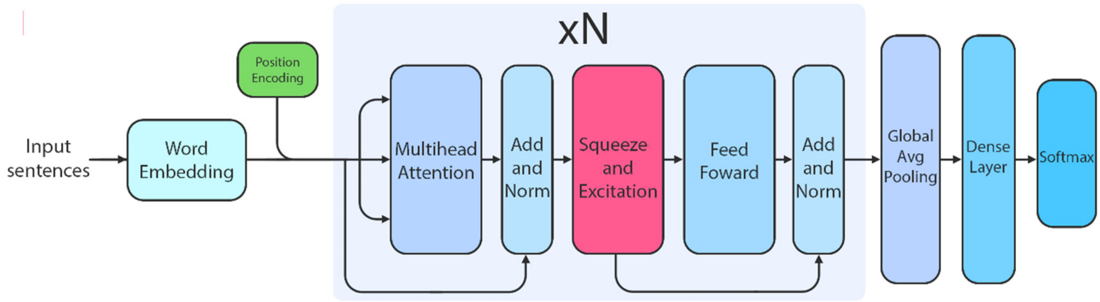

Figure 2: The process of Vietnamese sentiment analysis.

with d_{model} dimension too. W_A is a matrix of trainable weights with the size $d_{model} \times d_{model}$.

From (5), we have:

$$A = W_A.C + W_A.P \tag{6}$$

That means, Context information and Position Information both play the same role to create semantic meaning of a word in the sentence. Nonetheless, in Vietnamese, the meaning of a sentence belongs to the information of word positions. Thus, the concatenate operator is used to combine the information of word positions and the inputs of Multi-head Attention layer. That makes the context information and position information having the different weights during the transformation process.

$$A = W_A'.X \tag{7}$$

$$X = Concatenate(C, P) \tag{8}$$

The dimension of X is $2*d_{model}$, then W_A' is a matrix of trainable weights with the size $2*d_{model} \times 2*d_{model}$; thus, P and C are not the same weight as (6). It means: if the meaning belongs to the word order, the weight of the position (P) will be larger; else if it is the formal word, the weight of the context (C) will be larger.

The process for the Vietnamese sentiment analysis is shown in Fig.2. The proposed model for sentiment analysis in Vietnamese is based on Self-attention with Transformer architecture. In this work, the "concatenate" operation is used to incorporate position information with word context information as the input to Multi-head Attention layer as (7)(8). The input sentence is transformed to context embedding by using self-attention mechanism.

Moreover, in Fig.2, the layer of Squeeze and Excitation (Hu et al., 2018) is added between Multi-head Attention and Feed Forward layer to recalibrate features. It uses global information from the context matrix, which was the result of the Multi-head Attention layer, to select important features and suppress less useful ones before performing a transformation with feed forward network layer. It helps network to learn more important features efficiently for the task of sentiment analysis.

The gating mechanism of the Squeeze and Excitation (SE) layer is performed by stacking a GobalAveragePooling1D layer then forming a bottleneck with two dense layers. The first layer is a dimensionality-reduction layer with reduction ratio r with a non-linear activation. The second layer is the dimensionality-increasing layer to return the result of a sigmoid activation function with the dimension d_{model}.

$$X_i = \frac{1}{N}\sum_j^N f_{ij} \tag{9}$$

where, f_i is the i_{th} feature of the context matrix F, j is the j_{th} token of the sequence.

The output of this process is the squeeze global information of a feature of sequence into a feature channel descriptor. Then, every feature in the context matrix are represented by a value of this descriptor.

After that, a bottle-neck dense net is used to select the important features for sentiment task efficiently.

$$S = \sigma(W_{f_{c2}}.\delta(W_{f_{c1}}.X)) \tag{10}$$

where, δ refers to the ReLU function, S is the feature channel descriptor, σ is a sigmoid activation function, and W_{fc2}, W_{fc1} are trainable weights of the network.

The final output of this layer is obtained by the feature-wise multiplication between the scalar S and the context matrix:

$$O = M_{scale}(F, S) \tag{11}$$

where, O is the recalibration of the context matrix, and $M_{scale}(F, S)$ refers to the feature-wise multiplication between the scalar S and the context matrix F.

The work of the SE layer is shown in Fig. 3:

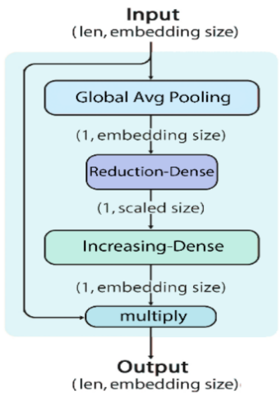

Figure 3: The layer of Squeeze and Excitation.

4 EXPERIMENTAL RESULS

4.1 Dataset

Dataset is set of comments of electronic products which were crawled from Vietnamese e-commerce websites, such as Tiki, Lazada, Shopee, Sendo, Adayroi, Dienmayxanh, Thegioididong, FPTShop, Vatgia. It includes 32,953 documents in labelled corpus: 22,335 positives documents and 10,618 of negatives documents.

For making the dataset balanced, some short negative documents are duplicated and segmented the longer ones. In the final result we have over 43, 500 documents in corpus with 22, 335 positives and 21, 236 negatives. Using for training models, we splitted corpus into 3 sets as following:

- Training set: 27, 489 documents.
- Validation set: 6, 873 documents.
- Test set: 8, 591 documents.

4.2 Evaluation Measures

Four measures which have been used in this study are based on the confusion matrix output. They are True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN).

Precision (P) = TP/(TP + FP)
Recall (R) = TP/(TP + FN)
Accuracy (A) = (TP + TN)/(TP + TN + FP + FN)
F-measure = 2.(P.R)/(P + R)

4.3 Results of Testing

We compare our model with four base line RNN models such as Long-Short Term Memory (LSTM), Gated Recurrent Units (GRU), stacked bidirectional LSTM and stacked bidirectional GRU with the following configurations. All models were evaluated on 8,591 documents.

- LSTM and GRU: 1 layer with 1,200 units.
- Stacked bidirectional model of LSTM and GRU: 2 stacked layers with 1,200 units in forward and 1,200 units in backward for each layer.

These model are compared with our method: i/ improved Self-Attention without the SE layer, ii/ improved Self-Attention combining the SE layer. The comparison results based on evaluate measures are shown in Table 3.

The method using the improved Self-attention with SE layer is more effective than other for Vietnamese sentiment analysis. Besides, this method is also more useful about the inference time.

From the experimental results, the improved Self-attention has the accuracy and F-measure is better than original methods of GRU and LSTM. Although the precision and F-measure of the improved Self-attention are lower than improved methods of GRU and LSTM (Stacked bi-GRU and Stacked bi-LSTM, its inference time is faster because it worked based on the grammatical structure of Vietnamese sentence. The SE layer helps to select important features and remove less useful ones before performing a transformation with feed forward network layer. It helps network to learn more efficiently for the task of sentiment analysis. From that, the improved Self-

Table 3: The results of the comparison between methods.

	P (%)	R (%)	A (%)	F-measure (%)	Time (s)
GRU	58.9	58.5	58.7	58.7	0.25
Stacked bi-GRU	91.1	90.8	90.9	91	1.05
LSTM	63.7	61.7	61.3	62.7	0.38
Stacked bi-LSTM	89.6	89.2	89.3	89.4	1.63
Improved Self-attention	86.6	85.1	85.3	85.9	0.003
Improved Self-attention combining SE layer	**91.7**	**91.6**	**91.6**	**91.6**	**0.07**

attention combining SE layer is more precise than other methods, it has the accuracy and F-measure being more than 91%. Moreover, the inference time of the combining method is also better than others. Hence, the proposed method can be useful in practice, especially in business intelligence.

5 CONCLUSIONS

In this paper, a method for sentiment analysis in Vietnamese is proposed. This method is studied based on the combination between the structure of a Vietnamese sentence and the technique of NLP, the self-attention with Transformer. The structures of a declarative sentence are studied and applied in the analysing of their meaning. Based on those structures of the sentences, the Self-attention network with the Transformer is used to analysis the sentiment of the sentence. The Self-attention network is improved by two steps:

(1) Adding the layer to determine the word positions by using the formulas (7)(8).

(2) Adding the layer of Squeeze and Excitation between Multi-head Attention and Feed forward layer to recalibrate features.

The experimental results of our method for Vietnamese sentiment analysis has the accuracy more than 91%, it is more effective than other methods. Besides, the inference time of the proposed method is also better than others. The process of this method can be applied in business for analysing the information on social network which serves in the influencer marketing.

In practice, the vast amount of training examples necessary to get satisfactory results is an obstacle to develop the natural language processing. In the future, we will use the method to transform this paper proposes a method for transforming the sentiment of a sentence to the opposite sentiment (Leeftink and Spanakis, 2019). This method can reduce by half the work required in the generation of training examples.

In the real-word, people can show their views in a sarcastic way that is difficult to determine. In the future work, the method need to be developed to classify the sentiment in those cases. That improvement has to analysis deeper in the sentence's structure and the technique of self-attention network. Moreover, for applying in business intelligence, such as the influencer marketing, the sentiment analysis in Vietnamese will be used to design the method for detecting the influencer on the social network, which were presented by the relational model (Do et al., 2018, Nguyen et al., 2015).

ACKNOWLEDGEMENTS

This research is supported by VinTech Fund, a grant for applied research managed by VinTech City, under grant number DA132-15062019.

REFERENCES

Aggarwal, C., Wang, H. 2011. Text Mining in Social Networks. In *Social Network Data Analytics*, Springer, 353-378.

Beigi, G., Hu, X., Maciejewski, R., Liu, H. 2016. An Overview of Sentiment Analysis in Social Media and Its Applications in Disaster Relief. In *Sentiment Analysis and Ontology Engineering: An Environment of Computational Intelligence*. Springer, 313-340.

Cao, H. 2017. *Vietnamese: Review of Functional Grammar*, Publisher of Social Science, 2nd edition. (Vietnamese)

Cheng, J., Dong, L., Lapata, M. 2016. Long short-term memory-networks for machine reading. In *EMNLP 2016, Conference on Empirical Methods in Natural Language Processing, Nov. 2016.*

Chung, J., Gulcehre, C., Cho, K., Bengio, Y. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. In *NIPS 2014 Workshop on Deep Learning, Dec. 2014*.

Clark, M. 1974. Passive ane Ergative in Vietnamese. In: Nguyen Dang Liem (ed.), *South-East Asian Linguistic Studies*, 75-80.

Do, N., Nguyen, H., Selamat, A. 2018. Knowledge-Based model of Expert Systems using Rela-model. *International Journal of Software Engineering and Knowledge Engineering (IJSEKE)* 28(8), 1047 – 1090.

fastText. 2019. https://fasttext.cc/

Gamal, D., Alfonse, M., El-Horbaty, E.M., Salem, A.M., 2019. Implementation of Machine Learning Algorithms in Arabic Sentiment Analysis Using N-gram Features. *Procedia Computer Science* 154(2019), 332-340.

Hoang, S., Nguyen, L., Huynh, T., Pham, V. 2019. An Efficient Model for Sentiment Analysis of Electronic Product Reviews in Vietnamese. In *FDSE 2019, 6th International Conference on Future Data and Security Engineering*. Springer. LNCS 11814, 132–142.

Hochreiter, S., Schmidhuber, J. 1997. Long short-term memory. In *Neural Computing* 9(8), 1735–1780.

Hu, J., Shen, L., Sun, G. 2018. Squeeze-and-Excitation Networks. In *CVPR 2018, Conference on Computer Vision and Pattern Recognition, June 2018.*

Huynh, T., Zelinka, I., Pham, Nguyen, H. 2019. Some measures to Detect the Influencer on Social Network Based on Information Propagation. In *WIMS 2019, 9th International Conference on Web Intelligence, Mining and Semantics, June 2019.*

Irfan, R., et al. 2015. A survey on text mining in social networks. In *The Knowledge Engineering Review* 30(2), 157-170.

Joachims, T. 1998. Text Categorization with Support Vector Machines: Learning with Many Relevant Features. In *ECML 1998, 10th European Conference on Machine Learning Chemnitz*, April 1998. Springer, 137-142.

Krouska, A., Troussas, C., Virvou, M. 2017. Comparative Evaluation of Algorithms for Sentiment Analysis over Social Networking Services. *Journal of Universal Computer Science* 23(8), 755-768.

Leeftink, W., Spanakis, G. 2019. Towards Controlled Transformation of Sentiment in Sentences. In *ICAART 2019, 11th International Conference on Agents and Artificial Intelligence, Feb. 2019*.

Nguyen, B., Nguyen, H., Romary, L., Vu, L. 2006. Lexical descriptions for Vietnamese language processing. In *Language Resources and Evaluation* 40(3/4), *Asian Language Processing: State-of-the-Art Resources and Processing*, 291-309.

Nguyen, D., Ngo, B., Tu, P. 2014. An empirical study on sentiment analysis for Vietnamese. In *ATC 2014, 7th International Conference on Advanced Technologies for Communications, Oct. 2014.*

Nguyen, H., Pham, V., Le, T., Tran, D. 2015. A Mathematical Approach for Representation Knowledge about Relations and Its Application. In *KSE 2015, 6th International Conference on Knowledge and Systems Engineering*, *Oct. 2015*.

Rokade, P., Kumari, D.A. 2019. Business intelligence analytics using sentiment analysis - a survey. *International Journal of Electrical and Computer Engineering (IJECE)* 9(1), 613-620.

Schuster, M., Paliwal, K. 1997. Bidirectional recurrent neural networks. *IEEE Trans. Sig. Process.* 45(11), 2673–2681.

Troussas, C., Krouska, A., Virvou, M. 2016. Evaluation of ensemble-based sentiment classifiers for Twitter data. In *IISA 2016, 7th International Conference on Information, Intelligence, Systems & Applications*, *July 2016*.

Underthesea. 2019. http://undertheseanlp.com/#!/

Vaswani, A., et al. 2017. Attention is all you need. In *NIPS 2017, 31st Conference on Neural Information Processing Systems*, *Dec. 2017*.

Zhou, Y., Zhou, J., Liu, L., Feng, J., Peng, H., Zheng, X. 2018. RNN-based sequence-preserved attention for dependency parsing. In *AAAI 2018, 32nd AAAI Conference on Artificial Intelligence, Feb. 2018*.