A Simple Supervised Learning Approach to Sentiment Classification at VLSP 2016

Hy Nguyen*, Tung Le*, Viet-Thang Luong† and Dien Dinh‡
Faculty of Information Technology
University of Science VNU-HCM
* Email: {1212166, 1212494}@student.hcmus.edu.vn
† Email: vietthang.hcmus@gmail.com
‡ Email: ddien@fit.hcmus.edu.vn

***Abstract*—This paper describes our system which participates in opinion mining task of VLSP evaluation campaign. The system uses multilayer neural network classifier with unigram and bigram TF-IDF feature. Our system achieves a comparative result with 69.4% F1-score on test data.**

I. Introduction

Opinion mining and sentiment analysis are the tasks to determine users opinion about products, movie, etc. In this topic, sentiment classification is a major task to classify opinion of a sentence or document into three categories: positive, negative and neutral. The prediction is extremely important because the users opinion becomes more and more value. The public interest is the main factor to affect the profit of products like movies, books, etc. Therefore, this problem is the interest of both researchers and companies.

In the context of VLSP competition, we build the system to determine the users opinion of a document into three labels: positive, negative and neutral. To make more practical, our system uses the simple features and classifier. However, our systems accuracy is really promising and acceptable with the low complexity.

The remainder of this paper is organized as follows. Section 2 provides a detail of our system. Section 3 describes the experimental setup and results. Section 4 concludes the paper and points to avenues for future work.

II. System description

Figure 1 illustrates some processes of sentiment classification systems which determine the sentiment label for a sentence or a document from the input. After preprocessing data by removing low-frequency words, there are two approaches to build a sentiment classification system: feature extraction and deep learning. In feature extraction approach, we extract sentence or document feature vector like TF-IDF, VietSentiWordnet. Then these feature vector is input to a classifier such as Support Vector Machine (SVM) or Multilayer Neural Network (MLNN) to determine sentiment label of sentence or document. Besides, in deep learning approach, we just use Long Short Term Memory (LSTM) to determine the sentiment label of sentence or document. Section II-A and II-B demonstrate more details about feature used in feature extraction approach and about classifier (SVM, MLNN, LSTM) in both approaches.

A. Features

This section briefly presents features which we try to use in feature extraction approach. The features extracted from each sentence or each document include:

*a) TF-IDF (Term Frequency * Inverse Document Frequency):* It usually used in information retrieval to determine which words are importance. This feature has solved the local and global information problem in feature extraction approach through TF and IDF score. In our experiment, we use TF-IDF score of unigram and bigram to extract the feature from a sentence or a document.

b) VietSentiWordnet: SentiWordNet is an important lexical resource supporting sentiment analysis in opinion mining applications [1]. A review is represented by a vector which length is the same as SentiWordNet vocabulary size. For each dimension, we compute the objectivity score corresponding to the term which it represents. The equation to compute the objectivity score is described by equation 1

$$score = 1 - (positivescore + negativescore) \quad (1)$$

c) TFIDF-VietSentiWordNet: this feature vector is created by concatenate both feature vectors above.

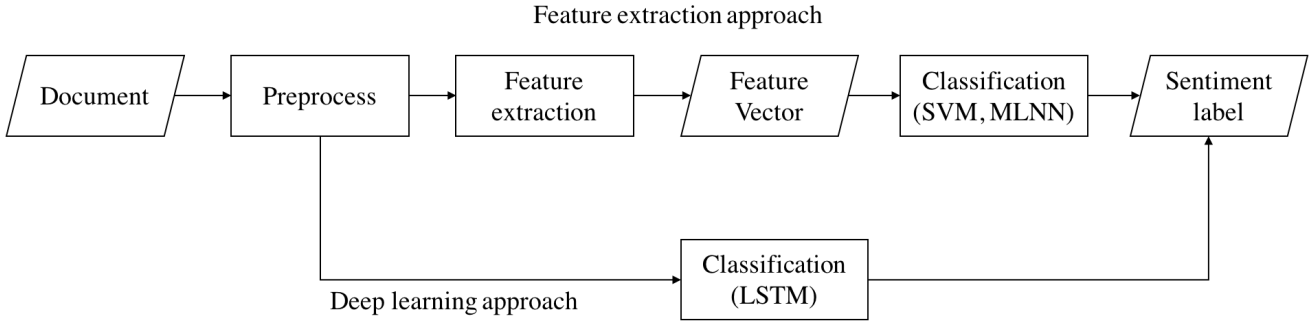
B. Classification

Classifiers are divided into 2 groups based on the input: Firstly, Support Vector Machine and Multilayer Neural Network take a feature vector as input which is gotten after feature extraction stage. Secondly, the input of LSTM is a list of word vector which is generated randomly and updated by training.

- Support Vector Machine (SVM) [2] SVM is the classic supervised machine learning algorithm. The goal of SVM is to determine the hyperplane that has the largest distance to the support vectors. With its effectiveness in classification, SVM is popularly used in many areas as the hand-written recognizer, opinion mining, etc. In this work, we implement SVM from Scikit-Learn[1] with default parameters to classify feature vector.

[1]http://scikit-learn.org/

Fig. 1. Sentiment classification system



- Multilayer Neural Network (MLNN) In neural network classifier, sentence's features are integrated into a multilayer full connected network. The last layer use softmax function classify feature vector. We built a Multilayer Neural Network classifier with simple architecture to avoid the over-fitting problem, it has one hidden layer with 100 neural units inside. Particularly, MLNN was trained by stochastic gradient descent optimizer (learning rate 0.1) and it uses sigmoid function as activation function at hidden layer.
- Long Short Term Memory Network (LSTM) [3] LSTM is a special kind of Recurrent Neural Network proposed in 1997 by Hochreiter. In every step, LSTM determines what information need be saved, so it is capable of learning long-term dependencies through the time. We built LSTM classifier based on the public code from DeepLearning.net[2]. In details, LSTM has the size of word vector and hidden vector are 50 and 100 respectively. For optimizing LSTM, we used Adam[4], a new method for stochastic optimization.

III. Experiments

A. Dataset

To train and test our systems, we use the data provided by VLSP evaluation campaign in opinion mining task. It contains user's reviews about technological device following three categories: "negative", "positive" and "neutral". In details, table I shows the data distribution for opinion mining task.

TABLE I
Data distribution

	Positive	Neutral	Negative	Total
Train	1700	1700	1700	5100
Test	350	350	350	1050

B. Systems setup

At figure 1, feature extraction approach consists of 2 stage extracting feature and classification. There are 3 kind of features: TF-IDF, VietSentiWordNet and TFIDF-VietSentiWordNet and 2 classifiers: SVM and MLNN. Therefore, we tried several systems which use a pair of feature and classifier. Example: SVM + TF-IDF is a system that uses Support Vector Machine classifier and TF-IDF feature. In deep learning approach, we have LSTM system which is described above.

[2]http://deeplearning.net/tutorial/lstm.html

C. Metrics

Since we have to choose only one system which achieves the best performance to submit the result on test data, we compared systems by using F1-score. F1-score was chosen because it represents the balanced score between precision and recall score which are the evaluation metrics of VLSP opinion mining task. Table II shows F1-Score of our systems. Each one is average of F-Score of 10-Fold cross-validation.

D. Results

In table II, the system uses MLNN classifier and TF-IDF feature perform the best result with 69.73% F1-score. Beside that, using VietSentiWordNet feature not only does not obtain good result but also reduce the result when combining with TF-IDF feature. Finally, although LSTM performs the good result, its results is less than system use MNLL and TF-IDF feature 1.83% F1-score. So we chose the system which uses MLNN classifier and TF-IDF feature to compete in VLSP opinion task. Table III illustrates the performance of competitive system on test data.

TABLE II
Our systems result

Classifier	Feature	F1-score
SVM	TF-IDF	68.16
	VietSentiWordNet	50.94
	TF-IDF & VietSentiWordNet	67.00
MLNN	TF-IDF	**69.73**
	VietSentiWordNet	50.97
	TF-IDF & VietSentiWordNet	68.09
LSTM		67.90

TABLE III
PERFORMANCE OF COMPETITIVE SYSTEM

	Precision	Recall	F1-Score
Negative	69.94	69.14	69.54
Neutral	65.80	64.86	65.32
Positive	72.42	74.29	73.34
Average	69.39	69.43	69.40

IV. CONCLUSION AND FUTURE WORDS

In this paper, we present several classification systems for opinion mining task of VLSP evaluation campaign. The systems follow two kinds of approaches: feature extraction and deep learning. In training data, the system use multilayer neural network classifier with unigram and bigram obtains the best result up to 69.63% F1-score through 10 fold cross-validation. In testing data, this system also achieves 69.4% F1-score.

In future, we will try to use different of features such as POS tag, NER, and semantic label and try to combine feature extraction and deep learning approach. It is regarded as the promising development of this work.

REFERENCES

[1] X. Vu and S. Park, "Construction of vietnamese sentiwordnet by using vietnamese dictionary," *CoRR*, vol. abs/1412.8010, 2014. [Online]. Available: http://arxiv.org/abs/1412.8010

[2] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995. [Online]. Available: http://dx.doi.org/10.1023/A:1022627411411

[3] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997. [Online]. Available: http://dx.doi.org/10.1162/neco.1997.9.8.1735

[4] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014. [Online]. Available: http://arxiv.org/abs/1412.6980