

# SENTIMENT ANALYSIS IN MYANMAR LANGUAGE USING CONVOLUTIONAL LSTM NEURAL NETWORK

Nwet Yin Tun Thein and Khin Mar Soe

Department of Computer Science, University of Computer Studies, Yangon, Myanmar

## **ABSTRACT**

*In recent years, there has been an increasing use of social media among people in Myanmar and writing review on social media pages about the product, movie, and trip are also popular among people. Moreover, most of the people are going to find the review pages about the product they want to buy before deciding whether they should buy it or not. Extracting and receiving useful reviews over interesting products is very important and time consuming for people. Sentiment analysis is one of the important processes for extracting useful reviews of the products. In this paper, the Convolutional LSTM neural network architecture is proposed to analyse the sentiment classification of cosmetic reviews written in Myanmar Language. The paper also intends to build the cosmetic reviews dataset for deep learning and sentiment lexicon in Myanmar Language.*

## **KEYWORDS**

*Social Media, Sentiment Analysis, Convolutional LSTM*

## **1. INTRODUCTION**

Analysing social media contents from Facebook, Twitter, and YouTube are important research area in Natural Language Processing. There has been increasing use of Sentiment analysis to analyse the social media contents. Spotting the sentiment of mentions (especially negative) is extremely helpful in marketing areas as reputation management and customer feedback. From the customer viewpoint, customers' reviews and feedbacks about the products are very useful for their determination whether to buy their interesting products or not. Moreover, deep learning is very popular in sentiment analysis and text classification. In this paper, convolutional LSTM deep learning approach is proposed for sentiment analysis on product reviews written in Myanmar language.

## **2. RELATED WORK**

In the recent years, Deep Learning approach was very popular and achieved the significant results in computer vision and speech recognition [1]. They are very common to use in NLP applications from learning word vector representations through building the neural language models. Moreover, they also perform the composition over the learned word vectors for text classification and solve the data sparsity problem [2].

Word embedding is the neural representation of a word and is a real vector. Word embedding allows us to measure similarity between words by simply using the distance between two embedded

vectors [2][3]. Recently, researchers observed that is not necessary for deep neural network to perform at word level. As long as the document represented as one-hot vector, the

model could work without any change, regardless of if each one-hot vector corresponds to a word. Character sequence proposed as an alternative to the one-hot vector. Similar ideas also applied to dependency parsing [4].

Deep CNN for NLP [5] is composed numerous of layers of convolutional and max pooling, it is identical to the convolutional architecture in the computer vision [6]. CNN was initially designed for computer vision and exploits layers with convolving filters that are applied to local features. CNN reached an outstanding result in computer vision where handcrafted features were used, e.g., scale invariant features transform (SIFT) followed by a classifier. The main idea is to consider features extractors and classifier as one jointly trained task [7]. CNN models for NLP achieved excellent results in semantic parsing [7], sentence modelling [8], search query retrieval [9], and other NLP tasks.

In this work Convolutional LSTM neural network architecture will be implemented for sentiment classification of Myanmar Language Cosmetic Review.

### 3. NEURAL NETWORK ARCHITECTURE

The proposed neural network architecture will be implemented by the combination of convolutional neural network with LSTM model.

#### 3.1. Convolutional Neural Network (CNN)

A convolutional neural network (CNN) is a neural network that applies convolutional layers to local features. When the input sentence is received, the CNN can be used to classify the sentiment of sentence is positive, negative, or neutral? The CNN architecture is shown in Figure 1.

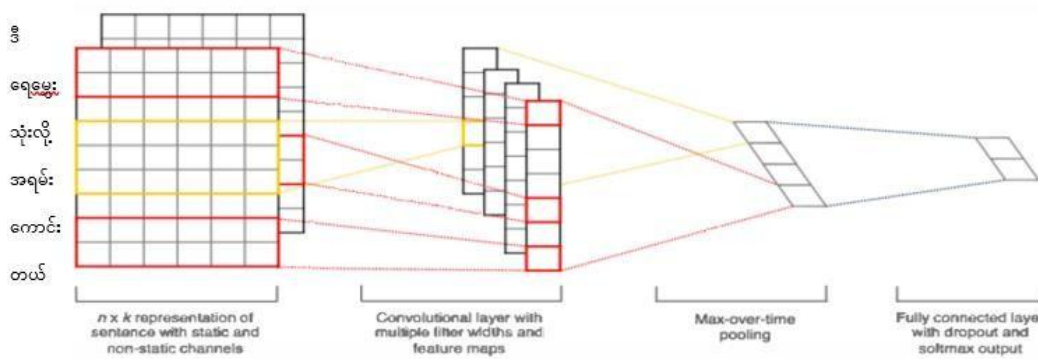


Figure 1. CNN Model Architecture

As shown in the figure, the CNN model is created with four layers.

- (i) Word Embedding layers
- (ii) Convolution layers
- (iii) Max-over time pooling layers and
- (iv) Fully connected layer.

### 3.2. Recurrent Neural Network (RNN)

The objective of the RNN is to make use of sequential information, based on the previous computation results. RNN has a memory that capture information in arbitrary long sequences. In this model, word by word analysis is taken over the input and then the semantic of all the previous text are preserved in a fixed-sized hidden layer. However, recurrent networks are biased model because recent words are more significant than earlier words. Therefore, the efficiency is degraded when the semantic of whole document is captured. To overcome this problem, Long short-term memory (LSTM) is applied.

### 3.3. Long Short-Term Memory (LSTM)

LSTM is more complicated function that learns to control the flow of information, to prevent the vanishing gradient and to allow the recurrent layer to capture long-term dependencies more easily. The memory cell consists of four main components: input, output, forget gates and candidate memory cell. Firstly, the value of input gate ( $i_t$ ) and candidate ( $c_t^{\sim}$ ) are computed according to the following equations.

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (1)$$

$$c_t^{\sim} = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (2)$$

Then the value of forget gate ( $f_t$ ) is computed using the Eq. (3).

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (3)$$

After getting the input gate activation ( $i_t$ ), forget gate activation ( $f_t$ ) and the candidate state value ( $c_t^{\sim}$ ), the candidate for new memory cell is computed using the Eq. (4).

$$c_t = (i_t * c_t^{\sim} + f_t * c_{t-1}) \quad (4)$$

With the new state of the memory cells, the value of output gate is computed using the Eq.(5) and (6).

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + V_o c_t + b_o) \quad (5)$$

$$h_t = (o_t * \tanh(c_t)) \quad (6)$$

Where  $x_t$  is the input to the memory cell layer at time t.  $W_i, W_f, W_c, W_o, U_i, U_f, U_o$  and are weighted matrices and  $b_i, b_f, b_c, b_o$  are bias vectors.

## 4. PROPOSED CONVOLUTIONAL LSTM MODEL FOR SENTIMENT ANALYSIS

In the proposed model, a recurrent layer is added as the substitution of pooling layer to reduce the number of convolution layers and capture the long-term dependencies. Therefore, convolutional, and recurrent layers are considered as the single model in the proposed model. Moreover, rectified linear (ReLU) was used for nonlinearity, padding was set to zero. All elements that would fall outside the matrix are taken to be zero. To reduce overfitting, we applied dropout 0.5 only before the recurrent layer. The Proposed model architecture is shown in Figure. 2.

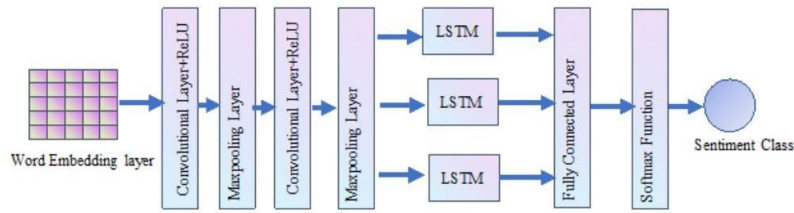


Figure 2. Proposed Model Architecture

## 5. EXPERIMENTAL SETUP AND ANALYSIS

The performance of the proposed model will be evaluated over the cosmetic review dataset with 70,000 binary labelled reviews. The sample reviews are shown in Table 1. The reviews are divided into 50:50 training and testing sets. The average length of each document is 320 tokens, with standard deviation of 205.8 tokens; the maximum length of a document is 2,840 words. In compared with other classification models, the proposed model outperforms the sentiment classification results as shown in the Table 2.

Table 1. Sample Review Text with Result Class

No.	Review Text	Class
1.	ဒီရေမွှေးသုံးလို့အရမ်းကောင်းတယ်။	Pos
2.	ဒီဆပ်ပြာကအက်ဆစ်များတယ်မကောင်းဘူး။	Neg
3.	ဒီမိတ်ကပ်ကကိန်းရိုးယားထုတ်ကုန်တစ်ခုပါပဲ။	Neu

Table 2. Accuracy of Sentiment Classification Models.

Models	Accuracy
Naïve Bayes	80.3%
SVM	75%
RNN	85%
CNN	87%
Convolutional LSTM (Proposed Model)	93.4%

## 6. CONCLUSION

In this paper, the combination of convolutional and recurrent layer into single neural network model was proposed. The proposed model was validated on the cosmetic review text collected from social media cosmetic pages written in Myanmar language. It achieved comparable results with less number of convolutional layers compared to the convolutional only architecture. The proposed model will also be applied other NLP such as semantic web search and spam filtering applications.

## REFERENCES

- [1] Graves, A, A.-r. Mohamed, and G. Hinton. Speech recognition with deep recurrent neural networks. in 2013 IEEE international conference on acoustics, speech and signal processing. 2013. IEEE.
- [2] Collobert, R., et al., Natural language processing (almost) from scratch. Journal of Machine Learning Research, 2011. 12(Aug): p.2493-2537 acoustics, speech and signal processing. 2013. IEEE.

- [3] Mikolov, T., et al. Distributed representations of words and phrases and their compositionality. in Advances in neural information processing systems. 2013.
- [4] Ling, W., et al., Finding function in form: Compositional character models for open vocabulary word representation. arXiv preprint arXiv:1508.02096, 2015.
- [5] Conneau, A, et al., Very Deep Convolutional Networks for Natural Language Processing. arXiv preprint arXiv:1606.01781, 2016.
- [6] Yih, W.-t., X. He, and C. Meek. Semantic Parsing for Single-Relation Question Answering. in ACL(2). 2014. Citeseer.
- [7] Kalchbrenner, N., E. Grefenstette, and P. Blunsom, A convolutional neural network for modelling sentences. arXiv preprint arXiv:1404.2188, 2014.
- [8] Sundermeyer, M., H. Ney, and R. Schlüter, From feedforward to recurrent LSTM neural networks for language modeling. IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), 2015.23(3): p. 517-529.

## AUTHORS

**Ms. Nwet Yin Tun Thein** is a Ph.D candidate of University of Computer Studies, Yangon. Her research interest is Natural Language Processing, especially in Sentiment analysis. Currently, she is doing research on sentiment analysis on Myanmar Language



**Dr. Khin Mar Soe** is a professor of University of Computer Studies, Yangon. Her research interest are Natural Language Processing, Machine Learning, and Deep Learning. She has been supervising Master thesis and Ph.D thesis on Natural language processing such as Information Retrieval, Morphological Analysis, Summarization, Parsing, Machine Translation and sentiment analysis.

