INFORMATION
Volume 26, Number 3, pp.117-162
ISSN 1343-4500 eISSN 1344-8994

Sentiment Analysis using a CNN-BiLSTM Deep Model Based on Attention Classification

WANG Yue
Beijing Academy of Artificial Intelligence (BAAI), Beijing, China

Lei LI
Graduate School of Science and Engineering, Hosei University, Koganei, Tokyo 184-8584 Japan
E-mail. lilei@hosei.ac.jp

Abstract

With the rapid development of the Internet, the number of social media and e-commerce platforms increased dramatically. Users from all over world share their comments and sentiments on the Internet become a new tradition. Applying natural language processing technology to analyze the text on the Internet for mining the emotional tendencies has become the main way in the social public opinion monitoring and the after-sale feedback of manufactory. Thus, the study on text sentiment analysis has shown important social significance and commercial value. Sentiment analysis is a hot research topic in the field of natural language processing and data mining in recent ten years. The paper starts with the topic of "Sentiment Analysis using a CNN-BiLSTM deep model based on attention mechanism classification". First, it conducts an in-depth investigation on the current research status and commonly used algorithms at home and abroad, and briefly introduces and analyzes the current mainstream sentiment analysis methods. As a direction of machine learning, deep learning has become a hot research topic in emotion classification in the field of natural language processing. This paper uses deep learning models to study the sentiment classification problem of short and long text sentiment classification tasks. The main research contents are as follows. Firstly, Traditional neural network based short text classification algorithms for sentiment classification is easy to find the errors. The feature dimension is too high, and the feature information of the pool layer is lost, which leads to the loss of the details of the emotion vocabulary. To solve this problem, the Word Vector Model (Word2vec), Bidirectional Long-term and Short-term Memory networks (BiLSTM) and convolutional neural network (CNN) are combined in Quora dataset. The experiment shows that the accuracy of CNN-BiLSTM model associated with Word2vec word embedding achieved 91.48%. This proves that the hybrid network model performs better than the single structure neural network in short text. Convolutional neural network (CNN) models use convolutional layers and maximum pooling or max-overtime pooling layers to extract higher-level features, while LSTM models can capture long- term dependencies between words hence are better used for text classification. However, even with the hybrid approach that leverages the powers of these two deep-learning models, the number of features to remember for classification remains huge, hence hindering the training process. Secondly, we propose an attention based CNN-BiLSTM hybrid model that capitalize on the

advantages of LSTM and CNN with an additional attention mechanism in IMDB movie reviews dataset. In the experiment, under the control of single variable of Data volume and Epoch, the proposed hybrid model was compared with the results of various indicators including recall, precision, F1 score and accuracy of CNN, LSTM and CNN-LSTM in long text. When the data size was 13 k, the proposed model had the highest accuracy at 0.908, and the F1 score also showed the highest performance at 0.883. When the epoch value for obtaining the optimal accuracy of each model was 10 for CNN, 14 for LSTM, 5 for MLP and 15 epochs for CNN-LSTM, which took the longest learning time. The F1 score also showed the best performance of the proposed model at 0.906, and accuracy of the proposed model was the highest at 0.929. Finally, the experimental results show that the bidirectional long- and short-term memory convolutional neural network (BiLSTM-CNN) model based on attention mechanism can effectively improve the performance of sentiment classification of data sets when processing long-text sentiment classification tasks.

Keywords: sentiment analysis, CNN, BiLSTM, attention mechanism, text classification

1. Introduction

1.1 Background

With the rapid development of the Internet, the number of Internet users increases sharply. When people are surfing the Internet, a large amount of information is generated, which reflects people's views and attitudes, and contains great commercial and social value [1]. Text is still an important way for people to produce and obtain information. The emotional classification of the short text is conducive to complete the user's push service, improve user experience better.

In recent years, the analysis of text affective tendency has attracted the attention of many scholars and has become a hot topic. With the rise of neural network, its related algorithms show a higher classification effect in text classification. The emotional detection and classification of text words, sentences and documents to a group of emotions by using psychological models. Therefore, a great practical significance to analyze the emotional information contained in the text and classify the text emotionally.

Sentiment analysis [2] is a set of linguistic operations belonging to the automatic processing of natural language. It's objective to identify the sentiment expressed in the text and to predict its polarity (positive or negative) towards a given subject [3].