A Study on Long-Short Term Memory Networks with Attention Mechanism

SU HANG

School of Knowledge Science,
Japan Advanced Institute of Science and Technology
March 2019

Keywords: Natural language processing, sentiment analysis, text data, machine learning, deep learning, Long short term memory, LSTM, Bi-LSTM, attention mechanism

With the rapid development of e-commerce websites, more and more people choose online shopping, which generates a large number of product review information. Sentiment analysis of text information not only tap customers' preference for products and provide shopping suggestions to potential customers, but also can help providers to improve products and services in a timely manner and improve business value.

Therefore, it is very necessary to classify product reviews by sentiment. There are two traditional research methods of sentiment classification: (1) the method based on lexicon; (2) methods based on machine learning. The former requires creating the lexicon manually, which is laborious. The latter is usually classified by naive bayes (NB), maximum entropy (ME), support vector machine (SVM), etc. These methods are easy to lose the syntactic and semantic information of the text and difficult to effectively capture the sentiment in the text.

With the application of deep neural networks in the field of natural language processing, Bengio et al. used neural networks to train word vectors to represent texts in 2003. The word vector can not only obtain semantic information effectively, but also avoid the problem of data sparsity. The word vector is used to represent the text, and by using the deep learning model, such as multi-layer neural network, convolutional neural network (CNN), recurrent neural network (RNN), etc., is adopted for sentiment classification, which can achieve better results than traditional machine learning methods.

Considering that the text data has strongly dependent on the context when classifying product reviews, and the standard neural network model cannot well solve the problem, so in this study, I adopted the Bidirectional Long Short term Memory neural network (Bi-LSTM) for sentiment analysis. In addition,

considering the different contributions of different words to the text, the Attention mechanism is introduced. Based on this, this study proposes a Bi-LSTM model based on the Attention mechanism to classify product reviews.

In order to verify the validity of the model, this study used the dataset of SemEval(International Workshop on Semantic Evaluation) 2014 Task 4 which is composed of reviews in Restaurant and Laptop. to test the model. Finally the experimental results show that the model is effective.