International Journal of Synthetic Emotions
Volume 9 • Issue 1 • January-June 2018

Sentiment Analysis in the Light of LSTM Recurrent Neural Networks

Subarno Pal, Academy of Technology, Hooghly, India
Soumadip Ghosh, Academy of Technology, Hooghly, India
Amitava Nag, Academy of Technology, Hooghly, India

ABSTRACT

Long short-term memory (LSTM) is a special type of recurrent neural network (RNN) architecture that was designed over simple RNNs for modeling temporal sequences and their long-range dependencies more accurately. In this article, the authors work with different types of LSTM architectures for sentiment analysis of movie reviews. It has been showed that LSTM RNNs are more effective than deep neural networks and conventional RNNs for sentiment analysis. Here, the authors explore different architectures associated with LSTM models to study their relative performance on sentiment analysis. A simple LSTM is first constructed and its performance is studied. On subsequent stages, the LSTM layer is stacked one upon another which shows an increase in accuracy. Later the LSTM layers were made bidirectional to convey data both forward and backward in the network. The authors hereby show that a layered deep LSTM with bidirectional connections has better performance in terms of accuracy compared to the simpler versions of LSTM used here.

KEYWORDS

Bidirectional LSTM, Long Short-Term Memory, LSTM, Recurrent Neural Network, Sentiment Analysis

1. INTRODUCTION

Sentiment analysis is a computational method of identifying or categorizing opinions expressed in a text, which is also one of the very active fields of research (Manning et al., 2008). Text obtained from different sources like user reviews and micro blogs express user's view or attitude towards the particular product or event etc. Sentiment analysis of small text is challenging because they are contextually limited. Decisions are taken on the basis of limited number of words used by the user. We deal with Sentiment Analysis as a supervised learning process where each data element (text reviews) are labeled as either 'positive' or 'negative' (Pang, Lee, & Vaithyanathan, 2002). Machine learning models are trained with word embeddings on these datasets and their accuracy is measured on the basis of their performance.

Artificial neuron network, a computational model developed on the basis of the structure and functions of biological neural networks, has achieved huge success over other machine learning techniques in sentiment analysis (Yoon, 2014; Socher, Pennington, Huang, Ng, & Manning, 2011; Xiong, Zhong, & Socher, 2002). Deep neural networks (DNNs) have recently achieved significant performance gains in a variety of NLP tasks such as language modeling (Bengio, Ducharme, Vincent, & Jauvin, 2003), sentiment analysis (Socher et al., 2013), syntactic parsing (Collobert & Weston, 2008), and machine translation (Lee, Cho, & Hofmann, 2016). A recurrent neural network

DOI: 10.4018/IJSE.2018010103

(RNN) is a special type of neural network, where connections are made between units which form a directed cycle, which allows it to exhibit a dynamic temporal behavior for the model. An RNN has an Input layer, variable number of hidden layers and finally one output layer. Basic RNNs are a network of neuron-like nodes, each with a directed (one-way) connection to every other node in which all connection (synapse) has a modifiable real-valued weight. These weights are constantly updated through successive iterations of the neural network. RNN are mostly used in hand writing recognition and speech recognition. For text classification purpose RNN is certainly more effective than any other variations of neural networks in practice.

A special variation of RNN, long short term memory (LSTM) networks is discussed. LSTM showed a striking accuracy in language modeling and speech recognition. We will be varying different forms of LSTM for our text classification purpose. A LSTM network contains LSTM units along with the input and output network layer units. A LSTM unit is capable of remembering values for either long or short time periods (Hochreiter & Schmidhuber, 1997) and it uses no activation function within its recurrent components. The stored value is not iteratively squashed over time, thus solving the vanishing gradient problem. LSTM blocks contain three or four "gates" that control information flow, implemented using the logistic function to compute a value between 0 and 1. An "input" gate controls the extent to which a new value flows into the memory, a "forget" gate controls the extent to which a value remains in memory, and an "output" gate controls the extent to which the value in memory is used to compute the output activation of the block.

Stacked LSTM is a particular model where one LSTM layer is stacked upon another to form a stack of LSTMs in the network. We preferred to form a stack of three LSTM layers to form a deep RNN. The detailed structure of the LSTM models will be described further. We in our study discuss the performance of various types of RNN and LSTMs on a large the IMDB dataset of reviews.

2. RELATED WORKS

Sentiment Analysis has been a favorite chapter for researcher for quite a long time. We discuss some of the important related works in this context.

In Socher et al. (2013b), the authors propose the Recursive Neural Tensor Network (RNTN) architecture, which represents a phrase through word vectors with a parse tree and then compute vectors for higher nodes in the tree using the same function, when trained on the new tree-bank, this model outperformed all previous methods on several metrics.

Santos et al. (n.d.) propose a new deep convolution neural network that exploits from character- to sentence-level information to perform sentiment analysis of short texts. Their approach for two corpora of two different domains: the Stanford Sentiment Tree-bank (SSTb), which contains sentences, from movie reviews; and the Stanford Twitter Sentimentcorpus (STS), which contains Twitter messages.

Mihalcea et al. (2007) make use of English corpora to train sentence-level subjectivity classifiers in Romanian language using two approaches, which they claimed can be applied to any language, and not only Romanian. In the first approach, they use a bilingual dictionary to translate an existing English lexicon to build a target language subjectivity lexicon. In the other one, they generate a subjectivity-annotated corpus in a target language by projecting annotations from an automatically-annotated English corpus. In Zhou et al. (2016) the authors propose an attention-based LSTM network to learn the document presentations of reviews in English and Chinese exploring word vectors as text representation.

Hochrieter and Schmidhuber (1997) pointed out that recurrent backpropagation or simple neural networks are extremely inefficient or fail miserably to learn information that is largely extended over time. long short term memory networks – usually just called "LSTMs" a special kind of RNN, capable of learning long-term dependencies, was proposed by them. LSTMs work tremendously well on a large variety of problems, and are now widely used.