International Journal of Social Media and Online Communities
Volume 12 • Issue 1 • January-June 2020

Implementation of Recurrent Network for Emotion Recognition of Twitter Data

Anu Kiruthika M., Anna University, Chennai, India
Angelin Gladston, Anna University, Chennai, India

ABSTRACT

A new generation of emoticons, called emojis, is being largely used for both mobile and social media communications. Emojis are considered a graphic expression of emotions, and users have been widely used to express their emotions in social media. Emojis are graphic unicode symbols used to express perceptions, views, and ideas as a shorthand. Unlike the small number of well-known emoticons carrying clear emotional content, hundreds of emojis are being used in different social networks. The task of emoji emotion recognition is to predict the original emoji in a tweet. Recurrent neural network is used for building emoji emotion recognition system. Glove is a word-embedding method used for obtaining vector representation of words and are used for training the recurrent neural network. This is achieved by mapping words into a meaningful space where the distance between words is related to semantic similarity. Based on the word embedding in the Twitter dataset, recurrent neural network builds the model and finally predicts the emoji associated with the tweets with an accuracy of 83%.

KEYWORDS

Deep Learning, Emoji, Emoji Classification, Recurrent Neural Network, Text Mining, Twitter, Word Embedding

1. INTRODUCTION

Developing social network platforms has given people a new way of generating and consuming a lot of web-based information (Dixit et. al., 2017). People used to obtain information from portal websites in the past. A large number of websites today provide information on a long list of subjects that vary from politics to entertainment. These traditional online information sources are always useful, but are less efficient since they often contain redundant information. Due to the arrival of online social network platforms, people tend to get information from them in a faster pace and is more efficient. These platforms are available for users to choose the information source they are interested in. To mention, large number of social network platforms such as, Twitter, Google+, and Facebook provide information to users (Geetha et. al., 2019; Xiong et. al., 2018).

The most popular platform for microblogging is Twitter. It is one of the fastest growing social network platforms and has a dominant microblogging position. Every day, more than 500 million registered users post 340 million twitter messages (Dixit et. al., 2017; Mohammad et. al., 2015), sharing their views and activities every day. Twitter posts are much shorter than those on regular microblogging platforms (Pak et. al., 2010; Pennington et. al., 2014). Only 140 characters or less can be posted in one twitter message (Dixit et. al., 2017). This feature makes twitter easier and keeps it

DOI: 10.4018/IJSMOC.2020010101

distinct from the massive amount of information available online for people to get the main point. In twitter, communication is made through messages commonly referred as the tweets. In this social website, people are allowed to make posts about different things, thus enabling people to get their required information from the massive amount of information available.

Twitter users can follow whatever people and source of information they prefer, depending on the users' needs. Twitter has therefore become a powerful platform with many kinds of information from worldwide breaking news to buying products at home, with all the benefits mentioned above. The information streams on twitter have experienced an incredible increase in the popularity of social network over the past few years. Users have a huge amount of information on various aspects (Unnisa et. al., 2016). Not all the information is useful to users however, and each user has their own interests and preferences. There is urgency for users to have personalized services. Nowadays, more and more personalized services are provided to benefit the users. People need this personalized service to make their fast-paced lives more efficient.

Every day, users are publishing a large amount of information on the twitter platform. Twitter data is related to the behaviour of the user and therefore many research studies focus on twitter and its collection of data. One of the twitter based research is user modelling. Researchers started to explore rankings and recommendations of twitter-referenced web resources to provide a personalized service. Based on their published tweets, a large amount of research focuses on modelling users and interests. Microblogs such as Twitter and SinaWeibo are a kind of popular social media (Pennington et. al., 2014) in which millions of people express their feelings, emotions, and attitudes. Because a large number of microblog posts are generated on a daily basis, the mining of feelings from this data source helps to perform research on various topics, such as analysing brand reputation, predicting the stock market, and detecting abnormal events.

It is therefore crucial to improve the performance of the tasks of sentiment analysis in microblog environments (Li et. al., 2017). In recent years, microblog sentiment analysis (Al-Halah et. al., 2019) has been a hot research area and several important issues have been studied, such as identifying whether a post is subjective or objective, called subjectivity classification, identifying whether a post is positive or negative, i.e. polarity classification, and recognizing emotion in a particular post. Supervised machine learning techniques have been widely adopted for analysing microblog feelings and have proved effective. Different features, such as sentiment lexicons, part-of-speech tags, and microblogging features, were used to reinforce the classifiers. However, due to the large vocabulary adopted by microblog users, the manual labelling of sufficient training posts is extremely labor intensive. Fortunately, different emoticons are often adopted in microblog environments and are usually posted along with emotional words (Xia et. al., 2017). In addition, in many microblog platforms, graphical emoticons, which are more accurate than those composed of punctuation marks, have been introduced. Emoticons can thus serve as an efficient source of emotional signals, enabling tasks for the classification of feelings to be performed without or with a small number of manually labelled posts (Bulut et. al., 2019).

Recognizing the emotion of the user is a major challenge for people and machines alike. On one hand, at certain times, people may not be able to recognize or state their own emotions. On the other hand, machines need precise ground truth for modelling emotions, as well as advanced algorithms for machine learning to develop emotion models. Sensors provide data sources, such as audio, gestures, eye gazes and brain signals that may be relevant to emotion recognition in hard sensing methods. Additional sensors may be attached to the user to provide personal physiological cues such as heart rate sensors, however these wearable sensors are not applicable in practical and natural settings, since they can be obtrusive to the user. Soft sensing methods, on the other hand, extract information from software that already exists with the user, on their phone. Events of positive and negative feelings and overall impressions have been studied in order to understand the nature and usage of such characters (Liu, 2012). Public opinions about global happenings can be analysed as the main resource to investigate the effects and usage of Emoji characters (Colneriĉ & Demsar, 2018) on social network sentiments.