Journal of King Saud University – Computer and Information Sciences (2016) **28**, 2–12

King Saud University

Journal of King Saud University – Computer and Information Sciences

www.ksu.edu.sa
www.sciencedirect.com

A comparative performance evaluation of neural network based approach for sentiment classification of online reviews

G. Vinodhini *, **R.M. Chandrasekaran**

Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar 608002, India

Received 27 July 2013; revised 5 January 2014; accepted 13 March 2014
Available online 31 October 2015

KEYWORDS

Artificial neural networks;
Sentiment analysis;
Text classification;
Opinion mining;
Support vector machine

Abstract The aim of sentiment classification is to efficiently identify the emotions expressed in the form of text messages. Machine learning methods for sentiment classification have been extensively studied, due to their predominant classification performance. Recent studies suggest that ensemble based machine learning methods provide better performance in classification. Artificial neural networks (ANNs) are rarely being investigated in the literature of sentiment classification. This paper compares neural network based sentiment classification methods (back propagation neural network (BPN), probabilistic neural network (PNN) & homogeneous ensemble of PNN (HEN)) using varying levels of word granularity as features for feature level sentiment classification. They are validated using a dataset of product reviews collected from the Amazon reviews website. An empirical analysis is done to compare results of ANN based methods with two statistical individual methods. The methods are evaluated using five different quality measures and results show that the homogeneous ensemble of the neural network method provides better performance. Among the two neural network approaches used, probabilistic neural networks (PNNs) outperform in classifying the sentiment of the product reviews. The integration of neural network based sentiment classification methods with principal component analysis (PCA) as a feature reduction technique provides superior performance in terms of training time also.

1. Introduction

Sentiment analysis is an interdisciplinary area which comprises of natural language processing, text analysis and computational linguistics to identify the text sentiment. Web has been a rapidly growing platform for online users to express their sentiment and emotion in the form of text messages. As the opinionated texts are often too many for people to wade through to make a decision, an automatic sentiment classification method is necessary

* Corresponding author. Tel.: +91 9626885482.
E-mail address: g.t.vino@gmail.com (G. Vinodhini).
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.jksuci.2014.03.024

to classify text messages into different sentiment orientation categories (e.g. positive/negative). Researchers explored various interesting approaches on the development of sentiment classification models for the classification of review sentiments, but work is still in progress. Many researchers investigated using various machine learning methods on English text sentiment classification (Prabowo and Thelwall, 2009; Chaovalit and Zhou, 2005; Turney, 2002; Pang et al., 2002) and a few studies on Chinese text sentiment classification (Ye et al., 2005, 2009; Wang et al., 2007; Tan and Zhang, 2008).

Neural networks have seen a rapid growth over the years, and are being applied successfully in various application domains for the classification problems. But the state of the art technique for neural network based text sentiment classification are found to be rare from the literature (Zhu et al., 2010; Chen et al., 2011; Sharma and Dey, 2012; Moraes et al., 2013). In recent years, there has been a growing interest in using ensemble learning techniques, which combine the outputs of several base classification techniques to enhance classification accuracy. However, compared with other research domains, related work about neural network based ensemble methods contributing to sentiment classification are still limited and more extensive experimental work is needed in this area (Wilson et al., 2006; Tsutsumi et al., 2007; Abbasi et al., 2008; Lu and Tsou, 2010; Whitehead and Yaeger, 2010; Xia et al., 2011; Su et al., 2013; Li et al., 2012). To fill this research gap, this paper makes a comparative study of the effectiveness of neural network based ensemble learning for sentiment classification. Feature selection is a crucial problem in the text classification. The aim of feature selection methods is to reduce the original feature set by removing irrelevant features for sentiment classification in order to improve accuracy of classification models (Wang et al., 2007; Tan and Zhang, 2008; Ahmed et al., 2008). PCA is used in this study in order to extract mathematically the most common features in all the models.

In this study, three different neural network models such as PNN, BPN and a homogeneous ensemble of PNN (HEN) are compared. The investigation is carried out for predicting the sentiment of product reviews with product attributes as features. Results of the evaluation of three different ANN based models are compared with that of the two different statistical methods (support vector machine and linear discriminant analysis) by computing five different quality attributes. Further, to analyze the relationship more clearly three different feature vector models are developed in each method. Model I is created using only unigram features. Model II is created using unigram and bigram features. Model III is obtained using unigram, bigram and trigram features.

1.1. Motivation and contribution

Existing works on the effectiveness of neural network based classification methods have been mainly conducted on text based topic classification (Ghiassi et al., 2012). There is lack of a comparative study on the effectiveness of neural networks based methods in text sentiment classification. The emerging interest and importance of text sentiment classification in the real world applications, motivates us to perform a comparative study of neural network based methods in sentiment classification. This study will greatly benefit application developers as well as researchers in the areas related to sentiment analysis.

Specifically, in this paper, we study the effectiveness of the neural networks based methods in sentiment classification as the interest of this study for three reasons.

- First, neural network based models has been very successfully applied to text classification and many other supervised learning tasks (Ur-Rahman and Harding, 2012; Ghiassi et al., 2012).
- The deep architectures of neural networks with layers (hidden) represents intelligent behavior more efficiently than "shallow architectures" like support vector machines (SVMs).
- The major features of neural networks such as adaptive learning, parallelism, fault tolerance, and generalization provide superior performance.

In spite of the above mentioned features of neural network methods, a few of the present research work on sentiment classification addressed the importance of integrating classification results provided by multiple classifiers (Xia et al., 2011). In addition, not much investigation has been carried out in sentiment classification to evaluate the benefits of combining neural network algorithms in order to increase the accuracy. Moreover, most existing studies in sentiment classification used the traditional measures for performance evaluation. A recent study (Kanmani et al., 2007), however, showed that various quality measures can be proposed to evaluate the accuracy of classification models in another domain like software fault prediction. This further motivates this study to evaluate the various performance evaluation metrics.

This work distinguishes itself from existing works in the following ways: In this work, probabilistic neural network and an ensemble of PNN are used for sentiment prediction which is not considered so far in sentiment analysis literature. This paper also provides a comparative study of existing neural network methods for sentiment classification through extensive experiments on a review dataset. Though proposing new techniques for sentiment prediction is not the main focus, we developed a homogeneous ensemble of PNN for classification which is not done so far in sentiment classification literature (Xia et al., 2013). Most of the earlier studies used various feature reduction methods but we attempted to use a hybrid combination of PCA and neural network (Cambria et al., 2013). In order to evaluate the prediction models in addition to traditional measures, five different quality parameters are used to capture the various quality aspects of the classification model. Training time is measured to show the superiority of feature reduction with the neural network based approach.

This paper outline follows the paper model of Kanmani et al. (2007). Section 2 is for the methodology used to develop the models. The data source used is reported in Section 3. The methods used to model the classification are introduced in Section 4. Section 5 lists out the various evaluation measures used. The findings from the experiments are discussed in Section 6. Section 7 describes the related work and Section 8 concludes the work done and proposes future works.

2. Model methodology

The methodology of the work is summarized below for developing and validating the classification models.