A Classifier Ensemble Approach to Detect Emotions Polarity in Social Media

Isidoros Perikos and Ioannis Hatzilygeroudis
Department of Computer Engineering and Informatics, University of Patras, Patras, Greece

Keywords: Sentiment Analysis, Classifier Ensemble, Social Media, Emotion Detection, Text Mining.

Abstract: The advent of social media has changed completely the role of the users and has transformed them from simple passive information seekers to active producers. The user generated textual data in social media and microblogging platforms are rich in emotions, opinions and attitudes and necessitate automated methods to analyse and extract knowledge from them. In this paper, we present a classifier ensemble approach to detect emotional content in social media and examine its performance under bagging and boosting combination methods. The classifier ensemble aims to take advantage of the base classifiers' benefits and constitutes a promising approach to detect sentiments in social media. Our classifier ensemble combines a knowledge based tool that performs deep analysis of the natural language and two machine learning classifiers, a Naïve Bayes and a Maximum Entropy which are trained on ISEAR and Affective text datasets. The evaluation study conducted revealed quite promising results and indicates that the ensemble classifier approach can improve the performance of sole classifiers on emotion detection in Twitter and that the boosting seems to be more suitable and to perform better than bagging.

1 INTRODUCTION

Over the last years, social media became a new means that connects people all over the globe with information, news and events in real time and has changed completely the way of human communication. Social media and microblogging platforms are constantly becoming an important aspect of everyday life providing various opportunities for social interaction, informing on news and events, expression of opinions and sharing of thoughts and attitudes. With the advent of Web 2.0 and social media platforms, people became more eager to express their opinions and share their experiences on web regarding almost all aspects of their day-to-day activities and global issues as well (Ravi and Ravi, 2015). Indeed, social media appeals to people of all ages because it provides opportunities for personal sharing of experiences and feelings, expressing opinions and attitudes and also offering reflections on a variety of social issues. Social media and microblogging platforms like Twitter have transformed people from passive information consumers to active producers. Every day, a vast amount of articles and messages are posted in various sites, blogs, news portals, social networks and forums which is rich in emotional content, opinions, attitudes and necessitates automated methods to analyse and extract knowledge from it (Shaheen et al., 2014).

A vital piece of information that could be extracted from user generated data in social media concerns the underlying emotional content expressed. Emotions can provide very indicative aspects of the personality of a person, his/her status and behaviour. The detection of emotional content can considerably enhance our understanding of users' states (Wang and Pal, 2015) and also to understand the public emotional attitude and views towards various events. From a user centric scope, analysing the text messages of a specific person can provide very indicative factors of the person's emotional situation, his/her behaviour and also provide deeper clues for determining his/her personality (Qiu et al., 2012). Furthermore, regarding events and user comments on them, from a topic centric perspective, the analysis of users' comments on a specific topic can provide very meaningful information about public stance, feelings and attitude towards various topics and events. In this line, emotion models can be employed to specify how people feel about a given entity such as a topic, an event and other (Wang and Pal, 2015).

The sentiment analysis and the recognition of

Perikos, I. and Hatzilygeroudis, I.
A Classifier Ensemble Approach to Detect Emotions Polarity in Social Media.
In *Proceedings of the 12th International Conference on Web Information Systems and Technologies (WEBIST 2016) - Volume 1*, pages 363-370
ISBN: 978-989-758-186-1

emotion in text is a hard problem on its own and when it comes to the analysis of user generated data in social media things can get even harder (Augustyniak et al., 2014). In the context of this work we present an ensemble classifier approach to detect emotional presence in tweets and specify their emotional polarity. The ensemble classifier relays on a knowledge based tool that performs deep analysis of the natural language and two machine learning approaches which are a Naïve Bayes and a Maximum Entropy learner. The knowledge based tool tries to analyze the sentence structure, spot words that convey emotional content and based on the word's dependencies, specify the overall emotional content of the sentence. The ensemble classifier schema combines the base learners under bagging and boosting methods with the aim to take advantages of their benefits and minimize their drawbacks. We examine the performance of the ensemble learning on user generated content on Twitter and assess its performance based on both combination methods. The evaluation study conducted on annotated tweets revealed very promising results regarding the ensemble classifier's performance to detect emotional content in tweets and specifying their emotional polarity.

2 BACKGROUND TOPICS

The detection of emotional presence in social media and the recognition of its emotional polarity are important for sensing and monitoring public stance and people feelings towards various events all over the globe. It could provide very indicative aspects of both individual behaviour and public attitude and also can assist in identifying emerging topics and trends. Applications and systems that determine the underlying emotional polarity could present an efficient and effective evaluation of people stance, thoughts and attitudes in real time and can assist a wide range of interest bodies such as governments, marketing agencies and other. The analysis can shed light into people behavioral tendencies and also present opportunities to learn about their feelings and perceptions in real time. The detection of sentiments and feelings in user data in social media also offer an unprecedented opportunity for marketing intelligence. Public sentiment as expressed in large-scale collections of Twitter posts can provide factors of social and economic attitudes and even be utilized to even predict stock market exchanges (Bollen et al., 2011).

However, in the literature most of the approaches train, use and rely on sole classifiers to perform the textual classification. In this work, we present an ensemble classifier approach that aims to improve the accuracy of base learners and the performance of sentiment analysis applications in detecting emotional presence in tweets and also determining their emotional polarity. The combination of classifiers is an effective method for improving the performance of a classification system (Li et al., 2007; Perikos and Hatzilygeroudis, 2016). The design and development of effective classifier ensembles requires that the used learner units have some level of diversity. There are many reasons for designing, developing and using classifier ensembles (Dietterich, 2000). From a statistical scope, by constructing an ensemble schema out of trained classifiers, the algorithm can average their votes and reduce the risk of choosing the wrong or underperforming classifier on new data. Even when different classifiers are trained and report a good performance, when just one is chosen, it may not yield the best generalization performance in unseen data. From a computational perspective, many learning algorithms work by performing some form of local search and it is very possible to get stuck at a local optimum. So, an ensemble constructed by running the local search from many different starting points may provide a better approximation to the true unknown function than any of the individual classifiers. Finally, from a representational scope the decision boundaries that separate data from different classes may be too complex and an appropriate combination of classifiers can make it possible to cope with this issue. In this line, given the characteristics of the user generated textual data in social media platforms, the utilization of ensemble classifier methods seems a suitable and efficient approach and the work presented in this paper is a contribution towards examining this direction.

3 RELATED WORK

Over the last years, the domain of sentiment analysis and emotion detection in social media has attracted a lot of interest. There is a huge research interest and several works study the way people express emotions and try to detect emotions in web and in social media (Cambria et al., 2013; Medhat et al., 2014; Liu, 2015). Machine learning supervised methods have been used on sentiment classification and emotion detection and are mainly based on supervised learning relying on manually labelled samples (Pang and Lee, 2008). Authors, in the work presented in (Go et al., 2009), study sentiment classification of tweets and examine the performance of Naïve Bayes, Maximum Entropy and Support Vector Machine algorithms and report performance results up to 82.7% for the Naïve Bayes and max entropy and 82.2 for SVM. Authors tried Unigram, Bigram model in conjunction with parts of speech features and found that the unigram model outperforms others. In (Firmino Alves et al., 2014), authors employ machine learning techniques for sentiment analysis of tweets in Portuguese during the world cup and achieved accuracy of approximately 80% with support vector machines and 73% with Naïve Bayes. The utilization of ensemble classifiers approaches could improve the efficiency of sentiment analysis and emotion detection systems (Devi et al., 2015; Fersini et al., 2014; Whitehead and Yaeger, 2010). In the text mining, ensemble classifiers have been applied successfully in various sub-domains, such as named entity recognition, word sense disambiguation and text classification (Xia et al., 2011). In (da Silva et al., 2014), authors present an ensemble classifier approach for sentiment analysis of tweets consisting of random forest, support vector machine, multinomial naïve Bayes and logistic regression classifiers. In the study, authors report that the classifier ensemble can improve classification accuracy that bag-of-words representation is suitable and can assist classifiers to achieve better accuracy. In (Wang et al., 2014), authors experimented with the performance of an ensemble classifier consisting of five base learners, that is naïve Bayes, maximum entropy, decision tree, k-nearest neighbor and support vector machine combined using random subspace method. Results indicate that ensemble classifier substantially improve the performance of base learners and reports better results than using solely the base learners and so authors suggest that ensemble learning methods can be used as a very viable approach for sentiment classification.

However, the ensemble classifier approaches in the literature mainly rely on machine learning classifiers. Machine learning approaches in general cannot fully leverage semantic and syntactic features of the sentences. On the other hand, the classification methods that are based only on keywords can suffer from the ambiguity in the keyword definitions in the sense that a word can have different meanings according to its usage and context and also the incapability of recognizing emotions within sentences that do not contain emotional keywords (Shaheen et al., 2014). So, an ensemble classifier approach that would combine both machine learning and knowledge-based approaches could be of great interest. In addition, our work presented in this paper is, to the best of our knowledge, one of the first approaches in the sentiment analysis domain to examine this direction and study the performance of an ensemble schema that combines diverse classifiers under different combinations methods.

4 THE ENSEMBLE CLASSIFIER

In this Section, we present the ensemble classifier approach, illustrate its architecture and analyse its functionality. The ensemble classifier combines two statistical machine learning learners and a knowledge based tool that performs deep analysis of the natural language sentences. The machine learning base learners are a naive Bayes and a maximum entropy learner which are trained on sentences from ISEAR, Affective Text and additional annotated tweets. The performance of the ensemble classifier is examined under bagging and boosting combination methods. In the following subsections, the base classifiers, their training and the different combination methods are described in detail.

4.1 Base Classifiers

4.1.1 Naïve Bayes

The Naïve Bayes classifier is a simple and commonly used model for classification which can achieve good performance in text categorization. It is based on Bayes theorem and is a probability based classification approach that assumes that documented words are generated through a probability mechanism. In general, the lexical units of a corpus are labelled with a particular category or category set and are processed computationally. During this processing, each document is treated as a bag-of-words, so the document is assumed to have no internal structure, and no relationships between the words exist and the position of the words in the document is ignored. A universal feature of Naïve Bayes classification is the conditional independence assumption. Naïve Bayes assumes that words are mutually independent and so, each individual word is assumed to be an indication of the assigned emotion. The Bayesian formula calculates the probability of a defined class, based on document`s features and is calculated as:

$$\mathrm{P(c|s)} = \frac{\mathrm{P(c)P(s|c)}}{\mathrm{P(s)}} \tag{1}$$

where $P(c)$ is the probability that a sentence belongs to category c, $P(s)$ is the probability of sentence s occurrence, $P(s|c)$ is the probability that the sentence s belongs to category c and $P(c|s)$ is the probability that given the sentence s it belongs to category c. The term $P(s|c)$ can be computed taking into consideration the conditional probabilities of occurrences of sentence's words given the category c, as follows:

$$P(s|c) = \prod_{1 \le k \le n} P(s_k|c) \qquad (2)$$

where $P(s_k|c)$ represents the probability that term (word) s_k occurs given the category c and n represents the length of sentence s.

4.1.2 Maximum Entropy

The Maximum entropy classifiers are feature based models that prefer the most uniform models that satisfy a given constraint. The aim is to find a model that can satisfy all the problem's constraints having also maximum entropy. The labelled data in training phase are used to derive the constraints for the model that characterize the class. In contrast to Naïve Bayes, the Maximum Entropy classifier does not make independence assumption for its features. So, it is possible to add features to a Maximum Entropy classifier like words unigrams, bigrams and N-grams in general, without worrying about the overlapping of the features. Maximum Entropy classifiers can achieve very difficult classification tasks and indicate good performance in various natural language processing tasks such as sentence segmentation, language modelling and named entity recognition (Nigam et al., 1999). MaxEnt classifier can also be used when we can't assume the conditional independence of the features, something that is particularly true in text mining and sentiment analysis problems, where features such as words are not independent. In general, the Max Entropy classifier requires more time to be trained comparing to Naïve Bayes, mainly due to the optimization problem that needs to be solved in order to estimate the parameters of the model. The classifiers use the bag of words representation technique, where a sentence is considered to be an unordered collection of words, whereas the position of words in the document bears no importance. It is used in combination with removal of stop-words and stemming of useful words.

4.1.3 Knowledge based Tool

The knowledge-based tool analyses and extracts knowledge from each sentence in order to specify its sentimental status (Perikos and Hatzilygeroudis, 2013). The architecture of the tool is depicted in Figure 1.

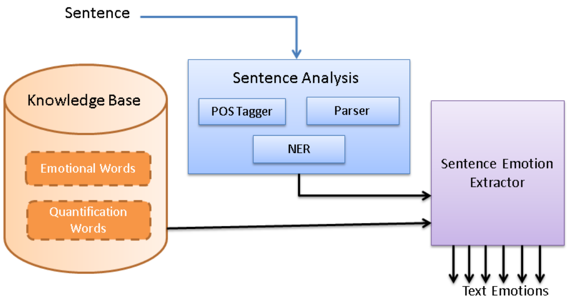

Figure 1: The architecture of the tool.

The Knowledge Base (KB) of the tool stores emotional words that convey emotions. It utilizes the WordNet Affect lexicon which is a widely used extension of the WorldNet and which was also extended by additional emotional words. The Stanford parser is used to analyse the structure of a sentence, specify the relationships between the sentence's words and determine the corresponding dependencies and the sentence dependency tree. The dependency tree represents the grammatical relations between the sentence's words in a tree based approach. Those relationships are presented as triplets consisting of the name of the relation, the governor and the dependent respectively. Dependencies indicate the way that words are connected and interact with each other. Named entity recognizer methods are utilized to detect proper names and named entities that appear in the sentence aiming to assist the sentence analysis and the specification of the way that emotional parts are associated with sentence's entities, such as persons. Words known to convey emotions are spotted using the lexical resources of the knowledge base and each emotional word detected is further analysed by the tool and its relations and the way it interacts with the sentence's words are determined. Based on the words' relationships, the tool identifies specific types of emotional word's interactions with quantification words, in order to specify its emotional strength. Finally, the emotion extractor unit specifies the sentence's overall emotional status based on the sentence emotional parts.

4.2 Training Data

The base learners were trained using annotated sentences from the ISEAR (Scherer and Wallbott, 1994) and the Affective Text (Strapparava and Mihalcea, 2007) datasets and also additional annotated Tweet. These datasets consist of sentences that have been emotionally annotated by experts. The ISEAR dataset consists of 7,660 sentences associated with 7 categories of emotions that are anger, disgust, fear, guilt, joy, sadness and shame. The Affective text dataset was designed for Semeval 2007 task on affective text and consists of news headlines sentences annotated based on the six emotions defined by Ekman (Ekman, 1999) which are anger, disgust, fear, happiness, sadness, surprise. For each sentence is specified its emotional load on a range from 0 to 100. For our experiment, we use emotions having the highest load as the sentence label and are considered only the emotions having a score greater than 50 specified by the experts.

Since the ensemble classifier detects emotional presence in tweets and characterizes them as emotional positive, neutral or negative, the sentences of the datasets where meta-annotated based on their emotional content. The meta-annotation specifies the emotional polarity of the sentences of the datasets as positive, neutral or negative and was based on the emotional theory of Russel that defines a two-dimensional model of affect (Russel, 1980). In this model, emotions can be presented in a dimensional space of two dimensions (Figure 2), where the one dimension represent the emotion's polarity and the other dimension the emotion's activation. The activation characterizes an emotion as activated or deactivated whereas polarity dimension is used to characterize emotions as positive or negative. For the meta-annotation of the sentences of ISEAR that express shame and guilt, the Parrot's analysis of emotions (Parrott 2001) was utilized, which specifies the shame and the guilt emotions to be associated with sadness. In this line, both emotions are meta-annotated to have negative emotional polarity. So, the sentences of the two datasets and also the additional annotated tweets from Sanders corpus were annotated, based on the aforementioned emotion schema, to convey positive, neutral or negative emotional content and a new corpus were formulated for the training of the classifiers for the needs of this study.

The mapping assists in specifying the polarity of a sentence based on its underlying emotional content. That is, in case a sentence is annotated to convey emotions, its emotional polarity is determined and meta-annotated according to the mapping of Russell's space (Russell, 1980). The joy emotion is associated with positive emotional polarity, while the emotions of anger, disgust, fear, sadness, shame and guilt characterize a sentence as emotionally negative. In this line, the surprise emotion can characterize a sentence as emotionally positive, in cases it is accompanied with joy emotion (happy surpise), as negative in cases it is associated with emotions of negative polarity or neutral in other cases.

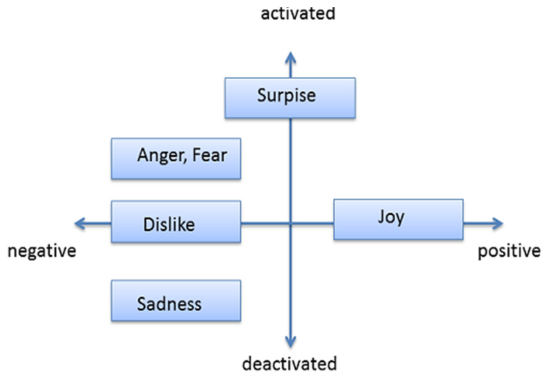

Figure 2: Polarity of basic Ekman emotions on Russel's scale.

The base classifiers are trained on the extended meta-annotated corpus to learn to detect emotional content and recognize its emotional polarity. In the training phase, additional tweets mainly form Sanders corpus that were also emotionally meta-annotated were utilized.

4.3 Ensemble Classifier Methods

The main aim of ensemble classifier is to leverage and benefit from the advantages of the base learners. For the combination of the base learners, various methods have been proposed in the literature and used in ensemble learners. The way that an ensemble classifier is formulated and the base classifiers are combined consist a crucial aspect that can greatly affect its performance. In the study, we examine the ensemble's performance based on the performance of the base learners under different combination methods. In the context of our work, we utilize instance partitioning methods and examine in the ensemble classifier the bagging and the boosting combination methods. Bellow, the nature and the functionality of the two methods are described.

4.3.1 Bagging

Bagging is one of the first combination methods for ensemble classifier. It relays on the principle to train each base classifier using a randomly drawn subset of the whole training dataset aiming to aggregate the multiple hypotheses generated by the same classifier on different distributions of training data.

Initially, the dataset is transformed into multiple data sets using sampling and iteration methods and

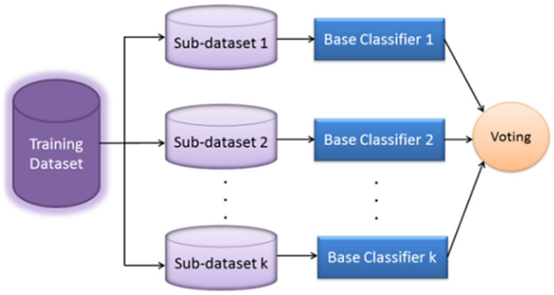

Figure 3: The Bagging combination method.

each set is assigned to a classifier. The diversity is secured by using bootstrapped replicas of the training dataset. The combination strategy of the base classifiers in bagging is the majority voting. Bagging assumes a dataset D and a learning system which trains a base classifier for each training set (i.e. bags) b = 1, 2, .., B sampled with replacement from D. The learning system is able to infer the label for each sentence of the testing set by aggregating over all the bags according to a majority voting decision rule.

4.3.2 Boosting

Boosting incrementally builds an ensemble by training each new model to emphasize those instances that previous models misclassified. The basic idea of boosting consists of three main stages.

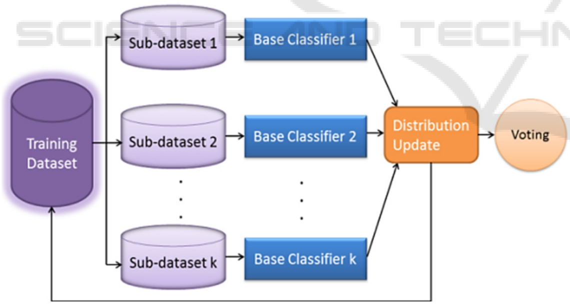

Figure 4: The AdaBoost method.

In the first stage, an iterative search to locate the examples that are more difficult to predict is performed, in the second stage the accurate predictions on those examples in each iteration are rewarded and in the third stage the rules from each iteration are combined (Schapire 1999). The workflow of the ensemble combination method is presented in Figure 3. In our work, the AdaBoost (Adaptive Boosting) algorithm was utilized. Both combination methods are examined on how they can enhance the performance of the base learners. The development of the combination methods and the base machine learning classifiers was implemented in Python language.

5 EVALUATION

An experimental evaluation study was designed and conducted to provide an insight of the performance of the ensemble approach examined under bagging and boosting combination methods. Initially, for the study we retrieve a wide range of posts published by different people on various topics on Twitter platform. To collect data, the Sanders Twitter sentiment corpus and the Twitter API ware utilized. The Sanders corpus consists of tweets collected from 4 search terms (@apple, #google, #microsoft, #twitter) which are characterized by an expert as neutral, irrelevant, positive and negative. The Twitter API was also used to access core Twitter data and to collect additional tweets. After that and for the needs of our study, we formulated a corpus consisting of 300 tweets and then a human expert was used to emotionally annotate each Tweet. The expert annotation would be used as a golden standard for the experimental evaluation. For each tweet, the expert specified the existence of emotional content and also, in case it exists, its emotional polarity. Based on the expert's annotations, the emotional polarity is specified, characterizing a Tweet as emotionally positive, negative or emotionally neutral.

5.1 Performance Evaluation

The evaluation study consists of two main stages. Initially, the ensemble classifier is evaluated in detecting emotional presence in tweets and after that in specifying the emotional polarity. For the evaluation we use the accuracy, precision, sensitivity and specificity metrics to assess the performance of both the sole classifiers and the ensemble classifier.

Table 1: The performance results of the classifiers.

Metric	N.B.	MaxEnt	K.B. Tool	E.C. Bagging	E.C. Boostin g
Accuracy	0.82	0.80	0.76	0.83	0.84
Precision	0.87	0.87	0.82	0.87	0.88
Sensitivity	0.78	0.78	0.78	0.79	0.80
Specificity	0.87	0.86	0.75	0.86	0.87

Initially, for the first part of the study that examines the classifiers performance in characterizing a tweet as emotional or neutral, the results obtained are illustrated in Table 1.

The results show a very good performance of the three classifiers and the ensemble classifier schema. The ensemble formulated performs robustly better in all experiment than the sole classifiers better in both the bagging and the boosting combination methods perform. A main reason for this concerns the good accuracy of the classifiers and the fact that the classification is performed with very good performance by each one of three classifiers of the ensemble schema. So, in cases that one of the classifiers fails to make a correct prediction, the final prediction is corrected by the remaining two. The results show that Naïve Bayes has the better performance among the base learners. Also, the ensemble classifier combined under boosting is performing slightly better than under bagging.

After that, in the second stage of the evaluation, the performance of the classifiers is evaluated in specifying the emotional polarity of tweets. The results are presented in Table 2.

Table 2: The performance results of the classifiers.

Metric	N.B.	Max Ent	K.B. Tool	E.C. Bagging	E.C. Boosting
Accuracy	0.80	0.78	0.73	0.81	0.82
Precision	0.88	0.86	0.84	0.85	0.88
Sensitivity	0.77	0.87	0.71	0.79	0.80
Specificity	0.85	0.77	0.70	0.85	0.86

The three base classifiers demonstrate very good performance in the recognition of the emotional polarity of emotional tweets. The ensemble classifier formulated in both combination methods is performing better than the base learners. Also, results show the boosting method to slightly outperform bagging once again in this part of the study. In the context of this study, the results show that the machine learning approaches achieve a satisfactory performance. In addition, the ensemble classifier approaches can enhance the performance and sole classification approaches in sentiment analysis of Tweets. Both combinations are suitable and can enhance the performance of sole classifiers so that the ensemble schema to perform robust better in detecting emotional presence in Tweets. Regarding the combination methods of the base classifiers in the ensemble, the results indicate the boosting method to perform slightly better than bagging in both stages of the evaluation study. Finally, the machine learning approaches have achieved a quite satisfactory performance. Given that their training was based also on sentences from ISEAR and the Affective Text datasets, it seems that both datasets are valuable and can assist in the training of machine learning algorithms.

6 CONCLUSIONS

In this paper, we present a classifier ensemble approach to detect emotional content in social media and specify their emotional polarity and examine its performance under bagging and boosting methods. The ensemble combines three classifiers, that are two machine learning and a knowledge based tool. The knowledge based tool performs deep analysis of the sentence structure, utilizes lexical resources to detect emotional worlds and specifies emotional content of a sentence based on the word dependencies. The two statistical machine learning classifiers are a Naïve Bayes and a Maximum Entropy trained using ISEAR, Affective Text datasets and annotated tweets. The evaluation indicated that the ensemble formed by diversified learners is a valuable approach on sentiment analysis of social media. Regarding the combination methods, results indicated boosting method to slightly outperform bagging and that both can perform robust better than the base classifiers.

As a future work a larger scale evaluation will be conducted to provide a deeper insight of the performance of the ensemble approach. Also, a next step regarding the feature representation would be to examine feature construction based on linguistic aspects and in addition examine and utilize SVM classifiers which are suitable for sparse representations. Moreover, the ensemble classifier utilizes the bagging and boosting combination methods which are instance partitioning methods and as a future work we plan to examine additional methods such as random subspace that is a feature partitioning method.

REFERENCES

Augustyniak, L., Kajdanowicz, T., Szymanski, P., Tuliglowicz, W., Kazienko, P., Alhajj, R., & Szymanski, B., 2014. Simpler is better? Lexicon-based ensemble sentiment classification beats supervised methods. In *Advances in Social Networks Analysis and Mining (ASONAM), 2014* IEEE/ACM International Conference on (pp. 924-929). IEEE.

Bollen, J., Mao, H., & Zeng, X., 2011. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1-8.

Cambria, E., Schuller, B., Xia, Y., & Havasi, C., 2013. New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, 28(2), 15-21.

da Silva, N. F., Hruschka, E. R., & Hruschka, E. R. ,2014. Tweet sentiment analysis with classifier ensembles. *Decision Support Systems*, 66, 170-179.

Devi, K. L., Subathra, P., & Kumar, P. N., 2015. Tweet Sentiment Classification Using an Ensemble of Machine Learning Supervised Classifiers Employing Statistical Feature Selection Methods. In *Proceedings of the Fifth International Conference on Fuzzy and Neuro Computing (FANCCO-2015),* pp. 1-13. Springer International Publishing.

Dietterich, T. G., 2000. Ensemble methods in machine learning, *Proceedings of the First International Workshop on Multiple Classifier Systems, MCS'00*, Springer-Verlag, London, UK, 2000, 1–15.

Ekman, P., 1999. *Basic emotions. Handbook of cognition and emotion*, 45-60.

Fersini, E., Messina, E., & Pozzi, F. A., 2014. Sentiment analysis: Bayesian Ensemble Learning. *Decision Support Systems*, 68, 26-38.

Firmino Alves, A. L., Baptista, C. D. S., Firmino, A. A., Oliveira, M. G. D., & Paiva, A. C. D., 2014. A Comparison of SVM Versus Naive-Bayes Techniques for Sentiment Analysis in Tweets: A Case Study with the 2013 FIFA Confederations Cup. In *Proceedings of the 20th Brazilian Symposium on Multimedia and the Web* (pp. 123-130). ACM.

Go, A., R. Bhayani, and L. Huang. 2009. *Twitter sentiment classification using distant supervision. Technical report*, Stanford Digital Library Technologies Project.

Li, S., Zong, C., & Wang, X., 2007. Sentiment classification through combining classifiers with multiple feature sets. In *Natural Language Processing and Knowledge Engineering*, 2007. NLP-KE 2007. International Conference on 135-140, IEEE.

Liu, B., 2015. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press.

Medhat, W., Hassan, A., & Korashy, H., 2014. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093-1113.

Nigam, K., Lafferty, J., & McCallum, A., 1999. Using maximum entropy for text classification. In IJCAI-99 workshop on machine learning for information filtering, Vol. 1, 61-67.

Pang, B., & Lee, L., 2008. Opinion mining and sentiment analysis. *Foundations and trends in information retrieval,* 2(1-2), 1-135.

Parrott, W. G., 2001. *Emotions in social psychology: Essential readings*. Psychology Press.

Perikos, I., & Hatzilygeroudis, I., 2013, Recognizing emotion presence in natural language sentences. *Engineering Applications of Neural Networks*. Springer Berlin Heidelberg, 2013. 30-39.

Perikos, I., & Hatzilygeroudis, I., 2016, Recognizing emotions in text using ensemble of classifiers, Engineering Applications of Artificial Intelligence.

Qiu, L., Lin, H., Ramsay, J. and Yang, F., 2012. You Are What You Tweet: Personality Expression and Perception on Twitter. *Journal of Research in Personality,* Vol. 46, Issue 6, 710-718.

Ravi, K., & Ravi, V., 2015. A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems*, 89, 14-46.

Russell, J. A., 1980. A circumplex model of affect. *Journal of personality and social psychology*, 39(6), 1161.

Schapire, R. E., 1999. A brief introduction to boosting. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI),* Vol. 99, pp. 1401-1406.

Scherer, K. R., & Wallbott, H. G., 1994. Evidence for universality and cultural variation of differential emotion response patterning. *Journal of personality and social psychology*, 66(2), 310.

Shaheen, S., El-Hajj, W., Hajj, H., & Elbassuoni, S., 2014. Emotion Recognition from Text Based on Automatically Generated Rules. In Data Mining Workshop (ICDMW), 2014 IEEE International Conference on, pp. 383-392. IEEE.

Strapparava, C., & Mihalcea, R., 2007. Semeval-2007 task 14: Affective text. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, 70-74, Association for Computational Linguistics.

Wang, G., Sun, J., Ma, J., Xu, K., & Gu, J., 2014. Sentiment classification: The contribution of ensemble learning. *Decision support systems,* 57, 77-93.

Wang, Y., & Pal, A., 2015. Detecting Emotions in Social Media: A Constrained Optimization Approach. *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015),* 996-1002.

Whitehead, M., & Yaeger, L. 2010. Sentiment mining using ensemble classification models. In *Innovations and advances in computer sciences and engineering* (pp. 509-514). Springer Netherlands.

Xia, R., Zong, C., & Li, S., 2011. Ensemble of feature sets and classification algorithms for sentiment classification. *Information Sciences*, 181(6), 1138-1152.