Sentiment Mining of Movie Reviews using Random Forest with Tuned Hyperparameters

Hitesh Parmar
Department of Computer Engineering
Marwadi Education Foundations Group of Institution, Rajkot, India
hitesh.parmar915@gmail.com

Sanjay Bhanderi
Department of Computer Engineering
Marwadi Education Foundations Group of Institution, Rajkot, India
sanjay.bhanderi@marwadieducation.edu

Glory Shah
Department of Computer Engineering
Marwadi Education Foundations Group of Institution, Rajkot, India
glory.shah@marwadieducation.edu.in

***Abstract*—Sentiment analysis is an automated mining of user generated opinionated text data such as reviews, comments and feedback. Sentiment analysis system classify those text data into their respective sentiments of positive, negative or neutral. Most of the previous research in this domain focused on using one of the three classifiers like SVM, Naïve Bayes and Maximum Entropy. In machine learning there are number of classifier models available. There is a huge potential that other robust classifier can also provide comparable or better results. Different machine learning classifier have their own set of configuration parameters, which are required to be tuned before a model gets trained. These parameters are known as hyperparameters. If these hyperparameters are tuned properly model can give state of the art result for a problem else it will do nothing but waste of time. Proposed approach focuses on a classifier model known as random forest. Few of the researchers have started using this model for sentiment classification but none of them focused on importance of hyperparameters. Random forest contains few set of hyperparameters which requires manual tuning. This paper focuses on tuning those set of hyperparameters manually. Proposed approach provided very promising results on standard datasets. Out of two standard movie review datasets, for one dataset it outperformed all the previous results with accuracy of 87.85% and for other standard movie review dataset it provided accuracy of 91.00%.**

Keywords—Sentiment Analysis, Opinion Mining, Random Forest classifier, Hyperparameters of Random Forest classifier

I. INTRODUCTION

Sentiment can be defined as “A personal Positive or Negative feelings” [1], “Sentiment is a usually formulated as two class classification problem, positive and negative” [2]. With the rapid growth of the online discussion group, social network sites, and increased usage of the micro blogging there is the increase in the number of people providing their opinion online and labeling those sentiments can provide the great summaries to all those people who are looking forward to some advice or help from the online opinions [3].

Sentiment Analysis is a process of mining on this user generated text content and determining the sentiment of users towards any particular thing like person, product or event and sentiments can be Positive, Negative or may be Neutral. Sentiment Analysis has become a very popular research area since 2000 [2]. After the research work published by [3] and [4] in 2002, it really provided the very good directions to many of the researchers who are working in the domain of sentiment analysis. This domain is also known as the Opinion mining as well.

This would be the first time starting from the internet era we are overwhelmed with a very huge volume of opinionated data over the social media sites and many other blogs, websites and forums, and without this data lot of research would not have been even possible. This led many of the researchers to focus on this area which is also having the huge potential for applications in many different areas. Opinions are always important to everyone whether it is individual, brands and services, governments or any other organization in the world, they play a very vital role in decision making. Business organizations are always in hurry to know that whether people like their products and services, what do people think about them, what kind of things people really like and don't like about their organization, product, service which may really help organizations to make decisions in a better way. Nowadays most of the people do not buy things without making some product analysis over the internet, people check for the product reviews and then make their decisions. Back in the time when organizations needed the public or consumers opinions they used to conduct the surveys and opinion polls which will require human resource and will be expensive as well as time consuming.

A. Levels of Sentiment Analysis

Sentiment analysis can mainly be carried out on any of the following two levels, known as document level sentiment analysis and sentence level sentiment analysis.

- *Document Level:* This focuses on classifying the whole document into its respective sentiments of positive, negative or neutral. Movie reviews or product reviews generally fall into this category. Most of the previous research in the sentiment analysis focused on document level only and many of them worked on the movie reviews only. This paper focuses on sentiment analysis of movie reviews which is nothing but the document classification, more information about the movie review dataset and experiments are provided in section 4: Experiments and Result Analysis.

- *Sentence Level*: Other approach is known as sentence level sentiment analysis in which only sentences are going to be analyzed and then will be classified as positive and negative polarity of a sentence. Sentence level sentiment analysis is being very popular nowadays because of the popularity of micro-blogging sites such as twitter and many other, which deals with short sentences which are limited to only 140 characters, and also influence many researchers to work on this sort of platforms.

B. Techniques for Sentiment Analysis

There are mainly two methods to carry out the sentiment analysis, first is known as Supervised approach or Machine Learning based approach which make use of machine learning classification techniques and other is known as Unsupervised or Lexicon based approach, which is also known as dictionary based approach.

- *Supervised Learning:* In supervised learning test data or unclassified data is going to be classified based on the data available in the training dataset.

  Training dataset is the one which is already labeled and uses the classifier algorithm to classify new data based on the labeled data or training data. Number of classifier algorithms like Support Vector Machine (SVM), Naïve Bayes and Maximum Entropy are mostly used classifier algorithm to carryout sentiment analysis. [3] were the first one to use the concept of supervised learning classifier in the area of sentiment analysis, they worked with above mentioned three classifiers used the concept of Unigrams for the feature selection and they found that Support vector Machine performs better compared to other classifiers. Reason to use above mentioned three classifier is because they work greatly in the area of text classification [Bo pang]. More details on these approaches are discussed in section 2-Related Work.

- *Unsupervised Learning:* Lexicon based approach is also known as the dictionary based approach or semantic based approach. This approach do not require separate training and testing dataset but instead of that list of words or dictionary of words will be used to classify the text data in form of sentence or document. Much of the research based on lexicon approach make use of available lexical resources such as dictionary of positive and negative words which are going to be used to classify the sentence or document. As if there are some positive words in the sentence then it means that sentence represents positive polarity, and if there are negative words then it represents negative polarity of a sentence or document. This concept was first started by [4].

  [4] carried out the sentiment analysis on the movie reviews data set and he worked with the approach of classifying reviews by average semantic orientation of phrases in the reviews and results returned using this approach was not that promising compared to that of the results returned by machine learning approach.

C. Feature Model

Feature selection model also plays a very important role in classification purpose, this refers to an approach which defines in which way those features are going to be used classify new data in to the specific type of class.

Following are different feature selection model while dealing with text classification or sentiment classification.

- *Unigram Model*: This model whole sentence will be divided into number of words, and those words are going to be used as feature.

  Unigram. Example: "it is a good movie".

  Output feature set {it, is, a, good, movie}

- *Bigram Model:* This model is also similar to Unigram model, but instead of working with single words this model will use combination of words to create the feature vector.

  Bigrams. Example: it is not bad movie.

  Output feature set {it is, is not, not bad, bad movie}

- *N-Gram Model:* When we use combination of more than two words for the feature vector that model will be referred as N-grams model. Which refers to combination of more words together to generate the feature vector and use that feature vector for classifying new or testing data.

For the purpose of sentiment analysis is unigram model is considered to be best as far as the results are considered [1, 2, 6]. All the experiments and evaluation provided in this report make use of Unigram as a feature selection model and which also provides some good results compared to other model like bigram [11].

D. Feature Reduction Approach

Sentiment analysis is nothing but the problem of text classification which classify text sentence or text document in to positive or negative class based on the words of text data and their polarity. One of the biggest problem in dealing with text data is that they are usually available in very high dimensions which may really affect the classifier performance. So there is a need for such technique which can help in reducing those features by removing or eliminating those features which are not relevant and keeping those features which are much important and which help to discriminate the sentence into available class labels like positive and negative. Among number of feature reduction techniques, Information Gain and Gain Ratio are the two methods which are very much popular and in terms of results as well these methods are consistent and provide better results compared to other feature reduction technique [7].

- *Information Gain:* Information Gain works on the basis of information required for a document to be classified in a respective class, depending on presence and absence of word in that document. Information gain is mainly responsible for finding importance of a feature in decreasing overall entropy.

- *Gain Ratio:* This approach is a modification of Information Gain. According to newly published paper of [7] Gain Ratio works better compared to information gain. In gain ratio contribution of all features will be normalized before classifying the document.

Main intention of this paper is to focus on random forest classifier and its hyperparameters, detail explanation of above mentioned approaches are out of the scope of this paper but researchers can go through [5][7][12] for more details on these approaches. Random Forest is discussed in section 3- Proposed Approach.

II. RELATED WORK

There are number of way through which sentiment analysis can be carried out, but this paper only focuses on sentiment analysis using machine learning approaches. Number of researchers have used different classification models to classify sentiments but most of them focused on Support Vector Machine, Naive Bayes, and Maximum Entropy as these classifiers works well for the general text classification problem. There are hardly few papers [8], [9] which focuses on Random forest for sentiment classification and none of them focused on influence of hyperparameters on output prediction. This paper focuses on fine tuning of those hyperparameters of Random Forest which can lead to good accuracy results compared to those of the previous results on standard datasets. Following are details of some of the influential work in the area of sentiment analysis.

In the first paper of sentiment analysis using machine learning approach [2] researchers worked with three supervised learning approach which are Naïve Bayes, Support Vector Machine and Maximum Entropy, and for feature selection model they used Unigram model, Bigram Model and POS Tagging and according to their experimental results Unigram Model works best for the purpose of Sentiment analysis whereas Bigrams did not provide good results. Support vector machine outperformed other two classifiers with the highest accuracy of 82.90\%. They worked on movie review dataset which contains 700 positive reviews and 700 negative reviews also known as Movie Reviews dataset V1.0. provided by [2]. This paper shows that using random forest classifier on the same dataset V1.0 with tuned hyperparameters can provide better results with increase of 4.95% compared to baseline results. [5] worked on the movie reviews dataset V2.0. provided by [2], which contains 1000 positive and 1000 negative movie reviews and achieved highest accuracy of 95.55% using Support vector machine and using hybrid feature selection approach known as entropy weighted genetic algorithm, in which features are selected with the combination of Information Gain and Genetic algorithm. Authors also mentioned that number of features which ranges near about 5000 can possibly work well for the purpose of sentiment analysis. Another research towards the hybrid feature selection for sentiment analysis is carried out by [6], Authors introduces the new concept for the feature reduction based on the Rough Set Theory (RST) and they are using Support Vector Machine for classification and as per authors' findings, if features are reduced in a well-defined manner then it can provide very good classification accuracy. Authors focuses on combining Information Gain with the Rough Set Theory for feature reduction and those reduced features are going to be used by the Support Vector Machine classifier and result provided by that classifier can be much better compared to the accuracy results of some other feature reduction approaches. Using this approach authors managed to get accuracy of 87.70% for a movie reviews dataset V2.0.

Feature selection also plays very important part in sentiment analysis, research work of [7] focuses on appropriate method for feature selection and classifier model selection in the area of sentiment analysis. Authors carried out experiments on five different feature selection techniques like Gain Ratio, Information Gain, Document Frequency, CHI statistics and Relief-F algorithm. For Classification they worked with Support vector machine, Naive Bayes, Decision tree, Maximum entropy, K-nearest neighbor and Window classifier, as per their experiments they concluded that Gain Ratio worked well for feature selection and support vector machine worked well on movie review dataset V2.0. with the highest accuracy of 90.90%.

Sentiment analysis on sentence level is becoming very famous research trend with the growth of social media. [1] was the first paper which focused on classifying the sentiments of tweets messages, they also made use of the machine learning approach to classifying those tweets, authors worked with three classifier which are Support Vector Machine, Naïve Bayes and Maximum Entropy and for the feature selection model they worked with Unigram Model and they created their own dataset of 1.6 millions of tweets using different emoticons like :) , :(etc. As per their result conclusion they found that Support Vector Machine worked well providing the result accuracy of 82.20% whereas Naïve Bayes provided accuracy of 81.30% and Maximum Entropy provided 80.50% of accuracy. [8] was the first paper in the area of sentiment analysis which focused on using Random Forest for classification. They also worked with the other two classification approaches as well which are Naïve Bayes and KNN classifier. They focused on three class classification problem that is Positive, Negative and Neutral class classification. They managed to get 300 reviews form the android market and tried to classify those reviews. For those 300 reviews Random forest was able to achieve accuracy of 60\% where as Naïve Bayes managed to get 55% of accuracy and at last KNN managed to achieve 52% of accuracy. So from these results it is possible that Random Forest can achieve better results with proper techniques. [9] focuses on classifying the sentiments of tweet messages in to three classes which are positive class and negative class and neutral class. Authors have worked with many different classifiers like Random Forest, Naïve Bayes Multinomial, Naïve Bayes, Support Vector Machine, J48(Implementation of C4.5 classifier). For the experiments they worked on the data instances which ranged from 5000 – 17000. And for their experiments they concluded that performance of Support Vector Machine and Random Forest are very much acceptable compared to accuracy results of other classifiers. For their Experiments Random Forest provided the accuracy of 79.79% whereas for SVM it gave accuracy of 77.80%.

After achieving good results in supervised approach [1, 3, 5, 6, 7, 8, 9] many of researchers were inspired to use supervised approach for sentiment analysis. Many of researchers have also tried to come up with some different classification approach like Decision Trees, Random Forest, K-nearest neighbors [7, 8, 9], and few researchers came up with concept of ensemble approach [10].

III. Proposed Approach

This section focuses on random forest classifier, hyperparamters of random forest, their impact on accuracy and some of the features of random forest classifier.

Random Forests [14] was the first paper which brought the concept of ensemble of decision trees which is known Random Forest, which is composed by combining multiple decision trees. While dealing with the single tree classifier there may be the problem of noise or outliers which may possibly affect the result of the overall classification method, whereas Random Forest is a type of classifier which is very much robust to noise and outliers because of randomness it provides [28].

Random Forest classifier provides two types of randomness, first is with respect to data and second is with respect to features. Random Forest classifier uses the concept of Bagging and Bootstrapping [14]. Random Forest works as shown below.

Algorithm 1. Random Forest

Input: B = Number of Trees, N = Training Data, F = Total-Features, f = Subset of Features
Output: Bagged class label for the input data.

1. For each tree in Forest B:
   a) Select a bootstrap sample S of size N from training data.
   b) Create the tree Tb by recursively repeating the following steps for each internal node of the tree.
      i. Choose f at random from the F.
      ii. Select the best among f.
      iii. Split the node.

2. Once B Trees are created, Test instance will be passed to each tree and class label will be assigned based on majority of votes.

Bootstrapping is considered as technique for improving the quality of estimators, in which predefined amount of portion of total dataset will be selected and that will be used for training, so the classifier will not actually get to see the overall data but a small portion of it. Whereas Bagging refers Bootstrap Aggregating which is mainly used to improve the stability and accuracy of classification algorithms, it is mainly used to get rid of variance because single tree is considered to be of high variance but to get rid of that variance number of trees can be combined and the average result of those combined trees will be free from variance.

A. *Hyperparameters of Random Forest*

As Random Forest is the combination of decision Trees, it deals with multiple number of hyperparameters which are:

- Number of Trees to construct for the Decision Forest
- Number of features to select at random
- Depth of each trees.

All these hyperparameters are required to be set manually which will be time consuming and does not guarantee that it will give good results for the parameter that we have set manually. Each of the hyperparameters have their own importance and influence towards the output prediction. First hyperparameter is Number of Trees in the forest, increase in number of trees linearly increase accuracy of the model. Larger the size forest better the accuracy, but the accuracy will not be changed at certain level when even there is an increase in number of trees. Number of features also plays a very important role in classification. Random forest does not work on all the features but instead of that there are two values of features which are very famous in the literature [13] [14]and they probably may provide good accuracy results compared to other values of features, but it is worth trying random forest with other values for selecting features at random. Depth of tree is also a very critical hyperparameter in random forest, if smaller value is been selected for Depth then model will suffer from under fitting. More about influence of these hyperparameters are discussed in section 4: Experiments and Results.

B. *Features of Random Forest*

Random Forest is considered to be an accurate and robust classifier because of following two reasons.

- *Robust:* As Random Forest uses the concept of Bootstrapping, so each tree works on the subset of the total training data, and due to that each tree is trained on the different value of training data. So that it is very much robust in terms of noise [14].
- Accurate: Random Forest make use of concept of Bagging so that output of all decision classifier will be averaged, as there is a logical proof that when infinite number of data is provided to a single classifier then the result will not be consistent, whereas if those data is divided into number of classifier, then averaging of the result of those classifier will be consistent [15].

IV. Experiments and Result analysis

All the experiments carried out in this section are computed using open source tool Weka 3.7.10 under Windows machine running on core-i3 processor with 3GB of main memory. Following subsection discusses more on content of dataset, preprocessing of dataset and tuning of hyperparameters of random forest classifier.

A. *Data set Description:*

This paper focuses on experimental evaluation on two standard Movie Review datasets. Dataset is available at [18], which is usually conceded as the gold standard data set for the researchers working in the domain of the Sentiment Analysis. First Dataset is known as Movie Review Dataset V1.0 which consist of 1400 movie review out of which 700 reviews are positive and 700 reviews are negative. Second dataset consist of total 2000 Movie reviews and 1000 of which are positive and 1000 of which are negative. The main reason for using this data set is that, they are already classified in to the two classes which are Positive and Negative, so all the reviews which are positive by their contextual sentiments they are kept in to the positive directory and the one that are negative by their contextual sentiment are kept in the negative directory. All those reviews are in the text file format.

TABLE I. DESCRIPTION OF DATASET

Dataset	Content of the Dataset
Movie Review Dataset V1.0	1400 Movie Reviews (700+ & 700-)
Movie Review Dataset V2.0	2000 Movie Reviews (1000+ & 1000-)

B. Data Cleaning:

Data preprocessing plays a very important role in classification, as both the datasets are in text format which are required to be preprocessed in an effective way to get good accuracy results. For experiment evaluation both datasets are required to be converted in to the arff format with which Weka can work for further classification process. Once file is ready with the arff format [27], process of converting that file in to word vector is required which will convert those text files in to feature vector or word vector. For this process Weka [17] filter function stringtowordvector() is used. Once these word vector is ready, data reduction technique Gain ratio is applied. For gain ratio, ranker search is used with threshold value is set to '0' and then all those reduced features are given to the random forest classifier for the classification purpose.

C. Experiments and Result Analysis:

Random Forest is evaluated on the available sentiment analysis standard datasets. These are two gold standard movie review datasets V1.0 and V2.0 of Cornell University [18] are selected, which are also been used by many of the researchers in the field of sentiment analysis. After performing experiments on available datasets using Random Forests, results achieved are very much promising and competitive. Out of two datasets for V1.0 Random Forest outperformed all previous research results by providing accuracy result of 87.85% whereas previous highest result was 87.52%, comparison is shown in Table II. Experiment results on dataset V1.0

TABLE II. EXPERIMENT RESULTS ON DATASET V1.0

Sr. No.	Author	Approach	Feature Selection	Accuracy %
1	Pang and Lee [3]	Naïve Bayes, SVM, Maximum Entropy	Unigram, Bigram	82.90%
2	Mullen and Collier [10]	Support Vector Machine	Unigrams, syntactic relations	86.00%
3	E. Riloff et al. [19]	Lexicon Based Approach	Unigram, Bigram	82.70%
4	Xue Bai [20]	Two Stage Markov Blanket Classifier	All words, and their subsets	87.52%
5	*Proposed Approach*	*Random Forest*	*Unigrams*	*87.85%*

And for dataset V2.0 accuracy results provided by random forest was 91.00%, comparison is shown in Table III. Results are very promising and which can still be improved. All the results achieved using random forest are by tuning those hyperparameters of a classifier manually.

TABLE III. EXPERIMENT RESULTS ON DATASET V2.0

Sr. No.	Author	Approach	Feature Selection	Accuracy %
1	Pang and Lee [21]	Naïve Bayes, SVM	Graph based Approach	87.20
2	Kennedy & Inkpen [22]	SVM	Unigrams, Bigrams, based on Terms Counting	85.90
3	Zhu Jian [23]	Back propagation	Unigrams	86.00%
4	Rui Xia [24]	Naïve Bayes, SVM, Maximum Entropy	Unigrams, bigrams, dependency grammar	86.40%
5	Agarwal and Mittal [8]	SVM	Unigrams + Rough set theory	87.60%
6	Prabowo et. al. [25]	ID3, SVM	Document Frequency	89.00%
7	Sharma and Dey [6]	NB, SVM, ME, DT	Unigrams	90.90%
8	Konig and Brill [26]	Hybrid approach	n-grams	91.00%
9	Xue Bai [20]	Two stage Markov blanket classifier	All words, subset of words	92.00%
10	*Proposed Approach*	*Random Forest*	*Unigrams*	*91.00%*
11	A. Abbasi et al. [5]	Support Vector Machine	Hybrid feature selection of Information Gain + Genetic Algorithm	95.55%

Above mentioned accuracy results for dataset V1.0 and V2.0 are achieved using random forest by manually changing values of all three different hyperparameters. Both the dataset are evaluated on train and test split by keeping the ratio of 80% for training and 20% for Testing. Most of the researchers uses cross validation for evaluating their results but in the case of random forest it works on the bootstrapped data only so there is no need for cross validation or it can be avoided [14].

Dataset V1.0 was first converted into word vector of 10000 words with removing stop words, after this Information Gain was applied for feature reduction with threshold value of 0.002. This resulted in total of 2275 features. Random forest classifier with hyperparameters values for number of trees 900, number of features at random 12, depth value was set to unlimited and that provided classification accuracy of 87.85%. Dataset V2.0 was first converted into word vector of 10000 words with removing stop words, after this Gain Ratio was applied for feature reduction with threshold value of 0.00. This resulted in total of 1942 features.

Random forest classifier with hyperparameters values for number of trees 400, number of features at random 11, depth value was set to unlimited and that provided classification accuracy of 91.00%. There number of different values of each hyperparameter are tried for both the datasets and above mentioned values of hyperparameters are the ones that provided good results. Though different values of hyperparameters are tried manually for each iteration and based on the accuracy returned in that iteration hyperparameters values will be updated for the next iteration. However we have mainly focused on two hyperparameter that is number of trees and number of features. In which increase in number of trees linearly increases accuracy up to certain values and after that there will not be any drastic change in accuracy results.

V. CONCLUSION

This paper focused on using the Random Forest to carry out the sentiments analysis of movie review dataset. Tuning of hyperparameters in random forest needs a special attention because they are required to be tuned manually so it is very time consuming and complex. On the basis of experimental results, random forest performed well on the standard movie review datasets. For dataset V1.0 it resulted in highest accuracy of 87.85% and for dataset V2.0 it provided very promising results with accuracy of 91.00%. Most of the previous research has focused on mainly SVM, Naive Bayes and Maximum Entropy for the sentiment classification, but according to experiments carried out in this paper random forest can also provide good and competitive results and can provide better results if hyperparameters are fine tuned.

VI. FUTURE WORK

For future work, we focuses on performing automatic tuning of those hyperparameters of Random Forest Classifier. Hyperparameter optimization techniques such as random search or gird search, can provide some even better results compared to those results which are achieved by manually tuning of hyperparameters. Random forest with optimized hyperparameters can provide better results for the purpose of sentiment classification.

REFRENCES

1. A. Go, R.Bhayani, and L.Huang, Twitter sentiment classification using distant supervision, CS224N Project Report, Stanford, 2009.
2. B. Liu, Sentiment Analysis and Opinion Mining, Morgan Claypool Publishers, May 2012, pp. 1-167.
3. B. Pang , L. Lee, and S. Vaithyanathan, Thumbs up?:sentiment classification using machine learning techniques, Proceedings of the ACL-02 conference on Empirical methods in natural language processing, vol.10, 2002, pp. 79-86.
4. P. Turney, Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews Proceedings of Annual Meeting of the Association for Computational Linguistics, 2002.
5. A. Abbasi, H Chen, A Salem, Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums, ACM Trans. Inf. Syst. 26, 3, Article 12 (June 2008).
6. A. Sharma, S. Dey, Comparative Study of Feature Selection and Machine Learning Techniques for Sentiment Analysis ,Proceedings of the 2012 ACM Research in Applied Computation Symposium, ACM, New York, NY, USA, 1-7.
7. N. Jebaseeli, E. KirubaKaran, M- Learning Sentiment Analysis with Data Mining Technique, International Journal of Computer Science and Telecommunications, Volume 3, Issue 8, August 2012 PP. 45-48.
8. B. Agarwal, N. Mittal, Sentiment Classification using Rough Set based Hybrid Feature Selection, WASSA 2013: 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis
9. B. Gokulakrishnan, P. Priyanthan, T. Ragavan, N. Prasath, A. Perera, Opinion mining and sentiment analysis on a Twitter data stream, IEEE International Conference on Advances in ICT for Emerging Regions (ICTer), pp:182-188, 2012
10. T. Mullen, N. Collier, Incorporating topic information into sentiment analysis models, In Proceedings of the ACL 2004 on Interactive poster and demonstration sessions , Article 25
11. B. Pang, L. Lee, Opinion Mining and Sentiment Analysis, In Foundations and Trends in Information Retrieval, vol. 2, pp. 1-135, 2008.
12. A. Sharma, S. Dey, Performance Investigation of Feature Selection Methods and Sentiment Lexicons for Sentiment Analysis, IJCA Special Issue on Advanced Computing and Communication Technologies for HPC Applications (3), pp:15-20, 2012
13. S Bernard, L. Heutte, S. Adam,Influence of Hyperparameters on Random Forest Accuracy, Multiple Classifier Systems, Lecture Notes in Computer Science, Volume 5519, Springer , pp: 171-180, 2009
14. L. Breiman, Random forests, Machine Learning, vol. 45. Issue 1, pp. 5-32, 2001.
15. T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, Springer Series in Statistics, Springer New York Inc., New York, NY, USA, Year 2001.
16. S. Bhattacharyya, Lecture Note on: Using multiple models: Bagging, Boosting, Ensembles, Forests, IDS 572: Data Mining for Business, University of Illinois at Chicago, 2005
17. Weka, http://www.cs.waikato.ac.nz/ml/weka/.
18. Dataset : cs.cornell.edu/people/pabo/moviereview-data/.
19. E. Riloff, S. Patwardhan, J. Wiebe, Feature subsumption for opinion analysis , ACL, Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, pp. 440-448, 2006.
20. X. Bai, Predicting consumer sentiments from online text, Decision Support Systems 50 (4), pp. 732-742, 2011.
21. B. Pang, L. Lee, A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts, Proceedings of ACL, 2004.
22. A. Kennedy, D. Inkpen, Sentiment Classification of Movie Reviews Using Contextual Valence Shifters, Computational Intelligence, Vol. 22, No. 2 . pp. 110-125. 2006.
23. Z. Jian, X. Chen, Han-shi, Sentiment classification using the theory of ANNs, The Journal of China Universities of Posts and Telecommunications, 17(Suppl.): 58 - 62, 2010.
24. R. Xia, C. Zonga, S. Li, Ensemble of feature sets and classification algorithms for sentiment classification, Information Sciences, Elsevier, 181, PP.1138-1152, 2011.
25. R. Prabowo, M. Thelwall, Sentiment analysis: A combined approach, Journal of Informatics, Volume 3, Issue 2, PP. 143-157, 2009.
26. A. Konig, E. Brill, Reducing the human overhead in text categorization, In Proceedings of the 12th ACM SIGKDD conference on knowledge discovery and data mining, pp. 598-603, 2006.
27. Arff Format: www.cs.waikato.ac.nz/ml/weka/arff.html.
28. H. Parmar, G. Shah, Experimental and Comparative Analysis of Machine Learning Classifiers. International Journal of Advanced Research in Computer Science and Software En-gineering, 3(10), pp. 955