An Enhanced Technique for Analyzing Sentiments of Public Reviews - II

Chintan Panjwani, Rashmi Thakur

Abstract: Enhnaced Dual Sentiment Analysis (EDSA) is an improved system which enhances the performance of the existing Dual Sentiment Analysis (DSA) which is implemented in literature. It mainly focuses on improving the efficiency of the existing system by making some modifications to the existing DSA approach. EDSA improves the classification accuracy of the public reviews. Apart from the classification accuracy other parameters considered in EDSA are precision, recall and f-measure. In the first phase, a data pre-processing is performed to clean the data where subjectivity analysis is performed to obtain the subjective reviews and sentiment analysis is performed on subjective reviews only. Second phase deals with negation detection and sentiment word sreversal to obtain the reversed reviews. Third phase performs polarity calculation on the original and reversed reviews to obtain positive and negative reviews based on sentiment score of the reviews. Fourth phase performs the enhanced dual training and prediction where the positive and negative reviews are provided to various classifiers which provides the final results as the output. Final phase is the graphical representation of the various parameter values obtained from the previous phase which helps in comparing the results of the various classifiers.

Keywords: Bag of Words, Enhanced Dual Sentiment Analysis, Polarity Shift problem, Sentiment Analysis, Support Vector Machine

I. INTRODUCTION

Online sales have increased which leads to increase in the reviews available for different products or services bought. A large number of reviews are available on different websites which needs to be analyzed to obtain some meaningful information. Analysis of such reviews is important to provide the thoughts expressed in the reviews. Sentiment analysis is the process of analyzing text to extract useful information and classifying the reviews in different polarity classes. Various approaches have been implemented to perform sentiment analysis. One such approach is the DSA model. In DSA, the reviews are first reversed to obtain polarity reversed reviews which addresses the polarity shift problem [1]. Polarity shift problem can reverse the polarity of a review as a Bag of Words (BoW) approach is used to represent the reviews for classification. BoW breaks a text into individual words without considering the semantic information of the text. This leads to misclassification of reviews which is called the Polarity Shift Problem.

Chintan Panjwani*, M.E Computer Engineering Student, Thakur College of Engineering & Technology, Mumbai, chintan.panjwani@gmail.com

Rashmi Thakur, Assistant Professor, Thakur College of Engineering & Technology, Mumbai, thakurrashmik@gmail.com

DSA approach addresses this issue by reversing the reviews and then performing sentiment analysis on the original and reversed reviews in pair. The Enhanced Dual Sentiment Analysis (EDSA) model which is implemented here which focuses on modifying the existing DSA approach to improve the classification accuracy. EDSA model first breaks the review into individual sentences in case of multiple sentence reviews. Subjectivity analysis is performed on the individual sentences which separates the subjective reviews from the objective ones. Sentiment analysis is performed only on the subjective reviews which reduces the number of reviews and the time required for sentiment analysis. The sentences of a reviews are then split based on the connecting words such as and, or, etc. Negation detection is then performed on the individual sentences of the review. The negation words are removed which have the sentiment words following the negation word. The sentiment words which are out of the scope of negation are reversed to obtain the reversed review. Once the reversed reviews are obtained, polarity calculation is performed on both the original as well as reversed reviews to obtain the polarity class for the review which can be either positive or negative. The reviews are then labelled with the polarity classes obtained for the reviews. Feature extraction is performed on the reviews to extract the features which will be used in the dual training and prediction phase to obtain the polarity classes. Various classifiers are used in the enhanced dual training and prediction phase to train the classifier based on training data and then test the classifiers in the dual prediction phase to obtain the classification accuracy of the various classifiers. Apart from accuracy other parameters used for measuring the efficiency of EDSA model are precision, recall and f-measure. Classifiers used in the dual training and prediction phase are Naïve Bayes, Support Vector Machine, Maximum Entropy and Vote classifier. The results obtained are then graphically represented to compare the performance of the classifiers used. The rest of the paper is organized as follows. Section II provides the details about the classifiers used in EDSA model. Section III presents the methodology used in the implementation of the EDSA model in detail. Section IV gives the experimental results obtained from the implemented system. Section V gives the future scope of the implementation. Section VI is the conclusion which summarizes the learnings of the entire paper and provides certain applications of the proposed system.

II. CLASSIFIERS

The dual training and prediction phase in the EDSA model employs different classifiers for classification of reviews into polarity classes. In the dual training phase, classifiers are trained using labelled reviews from the data set.

An Enhanced Technique for Analyzing Sentiments of Public Reviews - II

In the dual prediction phase a different set of unlabeled reviews are provided to the classifiers which predicts the polarity class based on which the classification accuracy, precision, recall and f-measure are calculated to measure the performance of the classifiers. The DSA model uses Naïve Bayes, Support Vector Machine and Logistic regression classifiers in this phase. In EDSA, we have considered Naïve Bayes, Support Vector Machine, Maximum Entropy and Vote classifier for classification of reviews into polarity classes and calculating the classification accuracy of the EDSA model based on the accuracy of classifiers. The following sections provides the details about the classifiers used.

A. Naïve Bayes

Naïve Bayes is a simple yet effective classifier which is used in sentiment analysis [2]. It classifies a text or review in this context into appropriate classes (positive/negative) based on the vector of feature values [3]. Vector is a simple array of individual words in the reviews. A review is assigned a particular class based on the word count in the review. The word frequencies are considered as the features for classification of reviews. The reviews are classified into their respective classes using the following equation [3].

$$\hat{y} = \underset{k \in \{1,...,k\}}{argmax} p(C_k) \prod_{i=1}^{n} p(x_i | C_k)$$

where,

$\hat{y}$ represents the value which will give the class of the review,

$\underset{k \in \{1,...,k\}}{argmax}$ – k belongs to the classes considered for classification,

Here pos and neg class is considered, the review is classified based on the higher value obtained for the class considered,

$p(C_k)$ is the probability of the k_{th} class considered for classification,

$\prod_{i=1}^{n} p(x_i | C_k)$ is the product of probabilities of the individual features of the review considered for sentiment analysis given that class is k.

Given a review, the positive probability and negative probability is calculated using the above equation. The value of k will be 2 as two classes (positive, negative) are considered here. The value of n will be equal to the number of unique words(features) in the review. The review will be classified as positive or negative based on the higher value of probability obtained for a class.

B. Support Vector Machine

Support Vector Machine (SVM) is a supervised(feed-me) machine learning algorithm which can be used for both classification or regression challenges [2]. It performs classification by finding the best hyper-plane that differentiate the classes we plotted in n-dimensional space. The hyper plane is plotted in such a manner that it maximizes the distance between the support vectors and hyper plane. Fig. 1 illustrates how SVM classification is performed using a hyper plane to classify the reviews into positive and negative classes.
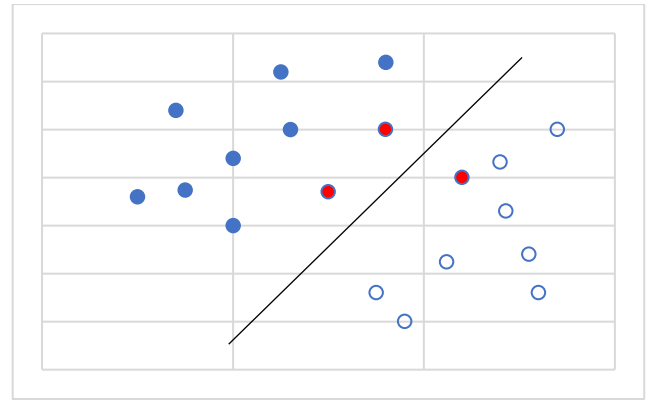

Figure 1. SVM graphical representation

In Fig. 1, the blue coloured points represent the positive labelled reviews, the blue bordered points represent the negative labelled reviews, the red coloured points are the support vectors, the line separating the positive and negative reviews is the hyper plane. The hyper plane is plotted such that its distance from the support vectors is maximum. These points are the training data. Given an unseen review, the SVM classifier classifies it as either positive or negative based on the training data available.

Text classification problems such as sentiment analysis of reviews is generally linearly separable [5]. The linear kernel is advantageous when there are a lot of features available [4]. Mapping the data to higher dimensions does not help to improve the classifier performance much [6]. Linear kernel is faster for text classification problems. So SVM used for classification of reviews uses a linear kernel function as only two data labels (positive, negative) are used in the classification. For classification having multiple classes, a non-linear kernel is used and the graph plotted in a 3D space with x, y and z axes.

C. Maximum Entropy

In Maximum Entropy (MaxEnt) classification, the probability that a review belongs to a particular class given a context must maximize the entropy of the classification system [7]. By maximizing entropy, it is ensured that no biases are introduced into the system. It is a classifier which prefers the uniformity or maximum entropy if no data is observed. But as it sees the data, it moves away from the maximum entropy by explaining data. After it has explained the data, it again tries to maximize the entropy on whatever remaining is not seen [8]. MaxEnt classifier does not assume that the features are conditionally independent of each other. It can be used to solve multiple text classification problems such as topic classification, sentiment analysis and many more which is the reason we have considered the classifier for classification of reviews in the EDSA model. For each word w and class c ∈ C present in the review, a joint feature f(w, c) = N is defined where N is the number times that word w occurs in the review in class c [9].

Using an iterative optimization, a weight is assigned to each joint feature so as to maximize the log-likelihood of training data. The probability that a given review d belongs to class c using weights λ is given by the following equation.

$$P(c|d, \lambda) = \frac{\exp \sum_i \lambda_i f_i(c,d)}{\sum_{c' \in C} \exp \sum_i \lambda_i f_i(c',d)}$$

As a search procedure is involved in this to find the optimized maximum entropy, it is somewhat difficult to implement compared with other classifiers such as Naïve Bayes used earlier. MaxEnt classifier provides good results compared to other classifiers but sometimes it produces worst results which is based on the features considered for classification.

D. Vote Classifier

As the name suggests, Vote classifier considers all the previously mentioned classifiers i.e. Naïve Bayes, Support Vector Machine and Maximum Entropy to classify the review as either positive or negative. It is based on voting for the considered classifiers. The class to which the review belongs is based on the maximum votes received by the review for a class by the classifiers under consideration. Vote classifier works by combining the classifiers and considers the classification of each classifier as a vote. Based on the votes, a class is assigned to the review having maximum votes for a particular class [10].Vote classifier iterates over the list of classifiers considered for classification. Then each classifier classifies the review based on the features. Once each classifier classifies the review, mode of votes (most popular vote) is calculated to predict the class for the review. Apart from the mode of votes, another parameter considered here is the confidence. It is calculated by finding the votes for and against the winning vote. For example, 3/5 votes for positive is weaker than 5/5 votes. Using this we can return the ratio of votes which is the confidence indicator. The review will be assigned to a class with maximum confidence. Vote classifier accuracy depends on the accuracy of individual classifiers. The more number of classifiers with good accuracy, the more accurate will be the Vote Classifier. The problem of misclassification is minimized by using the Vote classifier as multiple classifiers are considered in a combined manner to predict the class of the review.

III. METHODOLOGY

The Enhanced Dual Sentiment Analysis (EDSA) model is developed to improve the classification accuracy of the existing Dual Sentiment Analysis (DSA) system. EDSA model makes some modifications to the DSA model for providing improved performance compared to the existing model. EDSA model is divided into multiple phases to obtain the classification of reviews and map the classification accuracy to verify the improvement over the existing model. It also considers other parameters to measure the performance of the system which are precision, recall and f-measure. EDSA model is presented as a phased system consisting of the following phases.

IV. DATA COLLECTION AND PRE-PROCESSING:

Datasets from multiple domains can be considered for the EDSA model. The datasets considered in the existing system are book, DVD, electronics and some other datasets having around 1000 positive and negative reviews. In the EDSA model, we have considered airline services dataset. The dataset is divided into multiple individual datasets based on the number of reviews. A total of 12 different datasets are created from airline services dataset which contains reviews from 500 to 9000. Six datasets contain same number of positive and negative reviews called as balanced review dataset and remaining six datasets have uneven number of positive and negative reviews called the unbalanced dataset. The results obtained from EDSA system for these datasets is represented in Table 1 and 2. The data from these datasets is refined using the following two methods.

a. Subjectivity analysis: The reviews from the dataset can contain both objective as well as subjective statements. Sentiment analysis should be performed only on the subjective reviews as objective statements do not express any opinion or sentiment. In subjectivity analysis, a review is first broken down into individual sentences. The individual sentences are then analyzed to find out the subjectivity score of the sentence. The reviews having a subjectivity score below a predefined threshold are discarded and only the subjective sentences from a review are considered for sentiment analysis. This reduces the length of reviews and in some cases it can eliminate entire review which reduces the number of reviews to be considered so that the overall processing time reduces. The subjective reviews are then further processed in the negation detection phase.

b. Negation detection: To address the polarity shift problem as discussed earlier, the reviews are reversed and sentiment analysis is performed on the original and reversed reviews simultaneously. This reduces the problem of misclassification as the review is analyzed for both positive as well as negative sentiment so that a positive review will have the least possible negative sentiment score and vice versa. This increases the confidence with which the EDSA system can classify the reviews. Negation detection is performed on the reviews to obtain the reviews for the reverse data creation phase. The negative words are detected such as not, don't, neither, etc. The reviews are now provided to the next phase to detect the sentiment words and obtain the reversed dataset.

V. REVERSED DATA CREATION

This phase focuses on creation of reversed dataset from the original dataset. It takes the preprocessed data and creates sentiment reversed reviews which is used for dual sentiment analysis to address the polarity shift problem and improve the classification accuracy. A reversed reviews file is created as follows.

a. Sentiment words reversal: This phase identifies the sentiment words from the review obtained from negation detection phase. It reverses the sentiment words to its antonym to obtain a reversed review. The antonyms of sentiment words are obtained using the WordNet antonym dictionary. The sentiment words which follow the negation words such as not, don't, etc are not reversed, only the negative words in such sentences are discarded to create a reversed review.

b. Reversed dataset creation: A reversed dataset is created from the reviews obtained from previous phase. The individual sentiment reversed sentences of a review are merged to form a single review text and then appended to a new reversed review file. This phase leads to creation of two different files, one corresponds to the original dataset and other is the one which contains the reversed reviews. Both the original and reversed dataset are further input to the next phase where polarity of each review is calculated.

VI. POLARITY CALCULATION

The original and reversed review polarity is calculated in this phase to create labelled data for the data processing phase. Polarity calculation labels the reviews as either positive or negative based on the sentiment score of the review. The labelled data is used in the dual training phase for training the classifiers. Polarity of the reviews is calculated as follows.

a. Sentiment score analysis: Sentiment of a text is its positivity or negativity [11]. The positivity or negativity is decided based on the sentiment score. The review is divided into individual words and each word is checked in a database containing of sentiment scores to obtain its sentiment score. The more positive sentiment score of words makes the review as positively polar. Similarly, negative sentiment score is considered to label reviews as negative. Based on the sentiment score, this phase labels the reviews from original and reversed dataset as either positive or negative.

b. Dataset segregation: The labelled reviews from previous phase are separated into different files. Positive and negative review files are created which serves as the input for data processing phase. The labelled dataset is fed to the classifiers for further processing. This helps in training the classifiers and the training data is used for prediction of unseen reviews.

VII. DATA PROCESSING

The refined dataset obtain from previous phase is applied to the data processing phase. The main aim of this phase is to train the classifiers using the training data and then test/predict the class for unseen reviews based on the trained classifiers. The dataset is divided into training and testing set. ¾ dataset is considered as the training dataset and the remaining ¼ is considered for testing phase.

a. Enhanced dual training: This phase employs multiple classifiers viz. Naïve Bayes, Support Vector Machine, Maximum Entropy and Vote classifier for training using the training dataset which contains the original and reversed reviews. The training is performed using original and reversed training data simultaneously. The labelled training dataset is provided to the classifiers which trains the classifiers for the prediction phase. Each review is considered to extract features (individual words/unigram features) which are fed to the classifiers for training purpose. Training involves making the classifiers familiar with positive/negative features in the reviews. This phase is called enhanced dual training as some modifications are performed to the existing system in terms of data pre-processing and classifiers such as the Vote classifier is used in addition to the classifiers used in existing system which improves the system performance. The classifiers are now applied to the testing/prediction phase to classify the reviews and calculate performance of the EDSA model.

b. Enhanced dual prediction: Enhanced dual prediction phase classifies the reviews into appropriate class (positive/negative) from the testing dataset using the classifiers trained in the previous phase. It is called dual prediction as both the original and reversed reviews are considered simultaneously for prediction. The testing dataset contains the previously unseen data. The performance of the EDSA model is measured by considering various parameters whose values are calculated based on the predictions from this phase. The parameters considered are accuracy, precision, recall and f-measure. These parameters provide measure for the overall performance of the system. EDSA model improves the performance of existing system by improving the classification accuracy of the system. Support Vector Machine provides the highest classification accuracy followed by the Vote classifier.

VIII. DATA VISUALIZATION

Data visualization provides a graphical representation of results for values of various parameters obtained from the dual prediction phase. The graphs plot the accuracy, precision, recall and f-measure values for different classifiers to provide a visual comparison of the efficiency of the classifiers.

a. Graphical result representation: The parameter values for balanced and unbalanced datasets are plotted for visualizing the classification results obtained from the classifiers in dual prediction phase. A total of 16 different graphs are plotted, 8 for balanced dataset results (4 – single fold, 4 – n-fold) and remaining 8 for unbalanced dataset results (4 – single fold, 4 – n-fold). Fig. 2 represents the graphs for balanced dataset results. Fig. 3 represents the graphs for unbalanced dataset results. Similarly, n-fold results are plotted for all the classifiers representing the parameter values for each classifier.

b. Result comparison: The parameter values plotted in the graphs show that SVM provides the best classification accuracy of 0.836 for unbalanced dataset and 0.822 for balanced dataset among the four classifiers considered. Vote classifier provides the second highest classification accuracy followed by Naïve Bayes and Max Ent. The classification accuracy of the system has improved compared to the existing system. Similarly, other parameters precision, recall and f-measure have higher values for SVM classifier and provide better results compared to other classifiers.

IX. RESULTS

Table I shows the experimental results for all the classifiers using unbalanced dataset. Table II shows the experimental results obtained using balanced dataset. The results are obtained by considering single fold (entire dataset) and n fold cross validation where the dataset is shuffled and divided into n smaller datasets, here we have considered n as 5. Cross validation is performed to improve the performance of the system by testing the model using more unseen data so that the classification accuracy can be improved.

TABLE I. Unbalanced dataset results

No. of reviews / Classifiers			Unbalanced reviews						
			500	1000	1500	3000	7500	9000	Average
Naïve Bayes	Single fold	Accuracy	0.7418	0.7429	0.722	0.7	0.7572	0.76	**0.7374**
		Positive precision	0.7886	0.76	0.756	0.804	0.77	0.766	**0.7741**
		Negative precision	0.6441	0.7	0.649	0.606	0.7282	0.746	**0.6788**
		Precision	0.7163	0.73	0.703	0.705	0.7491	0.756	**0.7265**
		Recall	0.7079	0.701	0.688	0.711	0.724	0.72	**0.7088**
		F-measure	0.7114	0.7086	0.693	0.699	0.7311	0.729	**0.712**
	N-fold	Accuracy	0.76	0.7736	0.772	0.793	0.7912	0.784	**0.7791**
		Precision	0.7472	0.7821	0.79	0.8	0.8272	0.828	**0.7957**
		Recall	0.7102	0.7216	0.717	0.77	0.7395	0.725	**0.7306**
		F-measure	0.7199	0.732	0.729	0.777	0.7524	0.738	**0.7413**
SVM	Single fold	Accuracy	0.7637	0.7714	0.777	0.873	0.9126	0.921	**0.8366**
		Positive precision	0.9213	0.9375	0.927	0.921	0.956	0.966	**0.9383**
		Negative precision	0.6129	0.6316	0.643	0.815	0.8523	0.859	**0.7355**
		Precision	0.7671	0.7845	0.785	0.868	0.9042	0.912	**0.8369**
		Recall	0.7928	0.8024	0.804	0.877	0.9165	0.926	**0.8531**
		F-measure	0.7592	0.7697	0.775	0.871	0.909	0.918	**0.8336**
	N-fold	Accuracy	0.8552	0.9007	0.905	0.915	0.9366	0.937	**0.9083**
		Precision	0.8459	0.8976	0.899	0.912	0.9338	0.934	**0.9038**
		Recall	0.8338	0.8897	0.899	0.912	0.9319	0.932	**0.8996**
		F-measure	0.8385	0.8925	0.899	0.912	0.9327	0.933	**0.9012**
MaxEnt	Single fold	Accuracy	0.6758	0.6457	0.653	0.623	0.6418	0.645	**0.6475**
		Positive precision	0.6667	0.6404	0.644	0.611	0.6339	0.638	**0.639**
		Negative precision	1	0.875	0.938	0.925	0.881	0.869	**0.9145**
		Precision	0.8333	0.7577	0.791	0.768	0.7574	0.753	**0.7768**
		Recall	0.5391	0.5247	0.537	0.541	0.5338	0.533	**0.5347**
		F-measure	0.4725	0.4404	0.462	0.458	0.4552	0.455	**0.4573**
	N-fold	Accuracy	0.6579	0.6443	0.647	0.621	0.631	0.634	**0.6391**
		Precision	0.594	0.8194	0.811	0.794	0.8018	0.799	**0.7698**
		Recall	0.5164	0.5218	0.527	0.538	0.518	0.516	**0.5227**
		F-measure	0.4278	0.4305	0.441	0.449	0.4199	0.417	**0.4309**
Vote classifier	Single fold	Accuracy	0.7747	0.7629	0.752	0.832	0.8106	0.803	**0.7893**
		Positive precision	0.7895	0.7645	0.76	0.824	0.7824	0.774	**0.7824**
		Negative precision	0.7347	0.7582	0.732	0.847	0.8991	0.902	**0.8122**
		Precision	0.7621	0.7614	0.746	0.835	0.8407	0.838	**0.7973**
		Recall	0.7262	0.7154	0.708	0.815	0.7646	0.751	**0.7466**
		F-measure	0.7369	0.7256	0.716	0.821	0.7796	0.767	**0.7577**
	N-fold	Accuracy	0.9876	0.9764	0.957	0.963	0.9177	0.904	**0.9509**
		Precision	0.9907	0.9819	0.968	0.97	0.941	0.933	**0.9641**
		Recall	0.9819	0.9686	0.942	0.955	0.8923	0.874	**0.9356**
		F-measure	0.986	0.9745	0.952	0.961	0.9083	0.892	**0.9457**

TABLE II. Balanced dataset results

Classifiers / No. of reviews			Balanced reviews						
			500	1000	1500	3000	7500	9000	Average
Naïve Bayes	Single fold	Accuracy	0.6512	0.6716	0.671	0.683	0.7514	0.765	**0.6988**
		Positive precision	0.7191	0.7432	0.692	0.798	0.7889	0.786	**0.7546**
		Negative precision	0.5783	0.5871	0.617	0.589	0.6934	0.729	**0.6322**
		Precision	0.6487	0.6651	0.655	0.693	0.7411	0.758	**0.6934**
		Recall	0.652	0.6697	0.631	0.697	0.7395	0.749	**0.6898**
		F-measure	0.6481	0.6657	0.633	0.682	0.7403	0.753	**0.687**
	N-fold	Accuracy	0.7088	0.7785	0.765	0.786	0.7816	0.804	**0.7706**
		Precision	0.7114	0.78	0.757	0.79	0.8151	0.828	**0.7803**
		Recall	0.694	0.7542	0.742	0.765	0.7393	0.771	**0.7442**
		F-measure	0.6928	0.7605	0.747	0.771	0.749	0.781	**0.7502**
SVM	Single fold	Accuracy	0.657	0.7485	0.873	0.87	0.8901	0.897	**0.8227**
		Positive precision	0.8333	0.8662	0.904	0.92	0.9455	0.932	**0.9001**
		Negative precision	0.5625	0.6464	0.829	0.812	0.8224	0.85	**0.7537**
		Precision	0.6979	0.7563	0.866	0.866	0.884	0.891	**0.8269**
		Recall	0.684	0.7639	0.87	0.874	0.8959	0.898	**0.831**
		F-measure	0.655	0.7477	0.868	0.868	0.8876	0.894	**0.8201**
	N-fold	Accuracy	0.8206	0.8704	0.893	0.901	0.9248	0.932	**0.8904**
		Precision	0.8188	0.8662	0.891	0.898	0.9227	0.93	**0.8878**
		Recall	0.8144	0.865	0.885	0.898	0.9206	0.929	**0.8854**
		F-measure	0.8146	0.8653	0.887	0.898	0.9215	0.93	**0.8861**
MaxEnt	Single fold	Accuracy	0.6512	0.6272	0.632	0.634	0.6204	0.615	**0.6301**
		Positive precision	0.6226	0.6135	0.623	0.616	0.6127	0.608	**0.616**
		Negative precision	1	1	0.895	0.947	0.9355	0.895	**0.9454**
		Precision	0.8113	0.8067	0.759	0.782	0.7741	0.751	**0.7807**
		Recall	0.589	0.5435	0.536	0.56	0.5268	0.525	**0.5469**
		F-measure	0.5349	0.4602	0.456	0.492	0.4326	0.429	**0.4673**
	N-fold	Accuracy	0.6206	0.6259	0.627	0.621	0.6201	0.616	**0.6217**
		Precision	0.7313	0.7953	0.801	0.792	0.7982	0.795	**0.7853**
		Recall	0.5578	0.5434	0.527	0.543	0.526	0.526	**0.5372**
		F-measure	0.4875	0.4598	0.434	0.458	0.4294	0.428	**0.4494**
Vote classifier	Single fold	Accuracy	0.6744	0.7278	0.743	0.834	0.8263	0.826	**0.7718**
		Positive precision	0.7216	0.75	0.714	0.828	0.8018	0.794	**0.7682**
		Negative precision	0.6133	0.6885	0.864	0.845	0.8852	0.907	**0.8005**
		Precision	0.6675	0.7193	0.789	0.836	0.8435	0.851	**0.7843**
		Recall	0.6686	0.7093	0.686	0.819	0.7972	0.795	**0.7459**
		F-measure	0.668	0.7125	0.689	0.825	0.8085	0.807	**0.7516**
	N-fold	Accuracy	0.9838	0.9741	0.977	0.965	0.9162	0.93	**0.9579**
		Precision	0.9864	0.9789	0.982	0.972	0.9382	0.947	**0.9674**
		Recall	0.9811	0.9685	0.972	0.958	0.8956	0.914	**0.9482**
		F-measure	0.9834	0.9729	0.976	0.964	0.909	0.925	**0.9551**

Enhanced Dual Semtiment Analysis

Balanced Naive Bayes Single Fold; Balanced SVM Single Fold; Balanced MAXENT Single Fold; Balanced Vote Classifier Single Fold
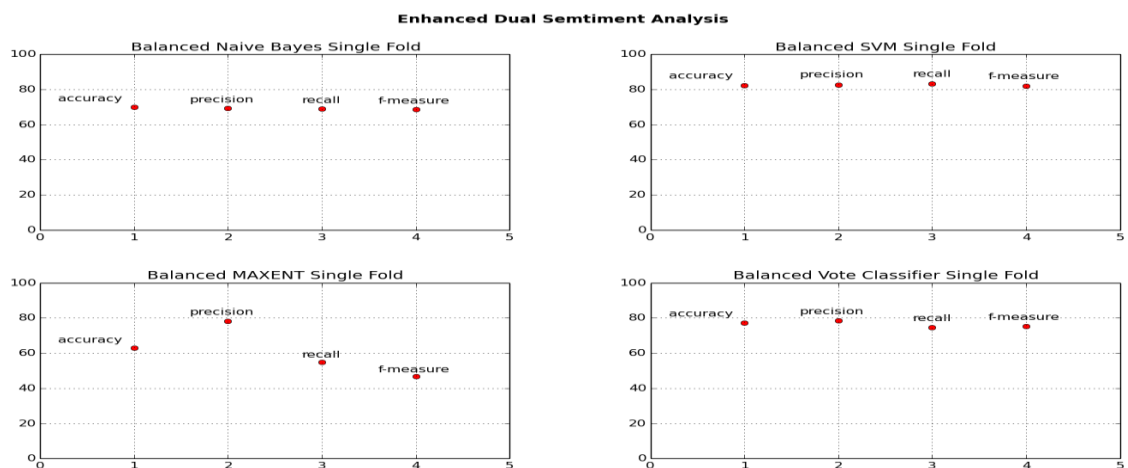
Figure 2. Balanced dataset single fold results

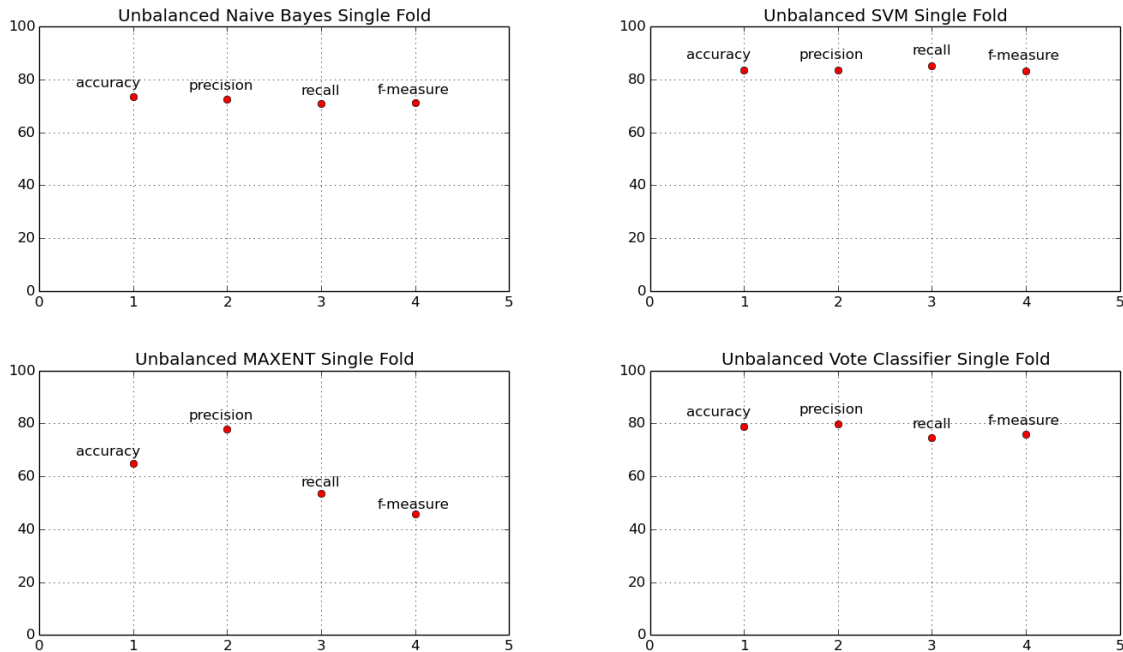**Enhanced Dual Semtiment Analysis**

Figure 3. Unbalanced dataset single fold results

X. FUTURE SCOPE

This paper presents an improved model for sentiment analysis. It mainly focuses on improving the classification accuracy by making changes to the existing system in the data pre-processing phase as well as by using additional classifiers. Future scope includes addition of natural language detection such as identifying the emoticons in the reviews which can improve the classification of complex text in the reviews. Another addition can be sarcasm detection as same sentence can have different meanings based on the context in which it is used. Finally, classifiers with higher classification accuracy can be considered for vote classifier so that the overall accuracy of the model can be improved further.

XI. CONCLUSION

This paper presents an enhanced model for sentiment analysis of public reviews. It presents a technique whose aim is to improve the classification accuracy of the existing system. To improve the classification accuracy, certain modifications are made to the existing system in the data pre-processing phase where subjectivity analysis is considered to filter the reviews which discards the objective text from the reviews. This reduces the computation time as fewer reviews need to be considered for the process. This model also includes a phase for polarity calculation which labels the reviews as positive/negative based on sentiment score. In the dual training and prediction phase a vote classifier is used which works on the voting of individual classifiers for classification of reviews. The experimental results presented in section IV justifies the improvement in the classification accuracy of the existing system where the SVM classifier has the highest classification accuracy along with other parameters such as precision, recall and f-

measure. Finally, in section V we have presented the future scope which can be considered to better the accuracy further.

REFERENCES

1. Rui Xia, Feng Xu, Chengqing Zong, Qianmu Li, Yong Qi, and Tao Li, Dual Sentiment Analysis: Considering Two Sides of One Review, *IEEE Transactions on Knowledge and Data Engineering*, 2015
2. https://medium.com/@vasista/sentiment-analysis-using-svm-338d418e3ff1 accessed on 06-08-19,12:32 pm
3. https://en.wikipedia.org/wiki/Naive_Bayes_classifier accessed on 16-08-19,13:15 pm
4. https://www.svm-tutorial.com/2014/10/svm-linear-kernel-good-text-classification/ accessed on 16-08-19,13:37 pm
5. Thorsten Joachims, Text Categorization with Support Vector Machines: Learning with Many Relevant Features, ECML 1998: Machine Learning: ECML-98 pp 137-142, 2005
6. D. O. Computer, C. wei Hsu, C. chung Chang, and C. jen Lin. A practical guide to support vector classification chih wei hsu, chih-chung chang, and chih-jen lin. Technical report, 2003
7. Mehra, N., Khandelwal, S., & Patel, P., Sentiment Identification Using Maximum Entropy Analysis of Movie Reviews, Working paper, 2002
8. https://www.quora.com/What-is-maximum-entropy-in-the-simplest-terms accessed on 18-08-19,11.52 am
9. https://blog.datumbox.com/machine-learning-tutorial-the-max-entropy-text-classifier/ accessed on 18-08-19,12.03 pm
10. https://pythonprogramming.net/combine-classifier-algorithms-nltk-tutorial/ accessed on 22-08-19,16.34 pm
11. https://readable.help/en/articles/2533315-what-is-a-sentiment-score-and-how-is-it-measured accessed on 23-08-19,17.50 pm

13