International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:10, No:1, 2016

Documents Emotions Classification Model Based on TF-IDF Weighting Measure

Amr Mansour Mohsen, Hesham Ahmed Hassan, Amira M. Idrees

Abstract**—**Emotions classification of text documents is applied to reveal if the document expresses a determined emotion from its writer. As different supervised methods are previously used for emotion documents' classification, in this research we present a novel model that supports the classification algorithms for more accurate results by the support of TF-IDF measure. Different experiments have been applied to reveal the applicability of the proposed model, the model succeeds in raising the accuracy percentage according to the determined metrics (precision, recall, and f-measure) based on applying the refinement of the lexicon, integration of lexicons using different perspectives, and applying the TF-IDF weighting measure over the classifying features. The proposed model has also been compared with other research to prove its competence in raising the results' accuracy.

Keywords**—**Emotion detection, TF-IDF, WEKA tool, classification algorithms.

I. Introduction

At present, Internet plays an important role in our life in many fields such as communications, online gaming, Entertainment, learning, research, and teaching. However, there is a huge amount of computerized text data and information spread among different countries and different cultures. Text mining field considers this large amount of data in some way needed to be analyzed for many uses, text mining simply targets retrieving useful information from large text data. There are many related fields of text mining including Information retrieval, text summarization, question answering, opinion mining, and emotion analysis.

Blogs and reviews become most important in the Internet nowadays. Besides, the social networks like Facebook and twitter play an important role for showing opinions and feelings about many topics or products. An example of using blogs when a company such as "HTC" offers a new mobile edition in the market and needs to analyze the impact of its mobile on the users. As the analysis of hundreds of blogs written about specific products or topics manually is too difficult, therefore, following a more productive approach is required. Sentimental analysis [1] or emotions analysis is a field considers how to analyze unstructured text written by people and analyze the opinions and the emotions in this text.

The remaining of this research paper discusses the related work in Section II, the proposed approach in Section III, the case study is explained in Section IV, then the experiments are presented and discussed in Section V, and finally the conclusion is presented in Section VI with discussing the points for the future work,

Amr Mansour Mohsen, Management Information Systems Department, Faculty of Commerce and Business Administration, Future University in Egypt, Egypt (e-mail: Amr.Mansour@ fue.edu.eg).

Hesham Ahmed Hassan, Computer Science Department, Faculty of Computers and Information, Cairo University, Egypt (e-mail: h.hassan@fci-cu.edu.eg).

Amira M. Idrees is associate professor in the Faculty of Computers and Information - Fayoum University, Egypt (corresponding author phone: 00201113900394; e-mail: ami04@fayoum.edu.eg).

II. Related Work

Different techniques are proposed for extracting emotions from text such as information retrieval, and machine learning. Information retrieval is considered with getting the relevant information from some unstructured text. According to [2] information retrieval techniques follows either Boolean models or vector space models. Vector space model has been applied in different research such as in [3], and [4]. In [3], A. Gohary, T. Sultan, M. Hana and M. Dosoky in [3] used vector space model to detect emotions in Arabic text by an Arabic lexicon with reaching an accuracy of 65%. Another research in [5] which proposed a system for emotion detection using similarity measures [6] with reaching an accuracy of 70%. In [7], (TF_IDF) measure is applied with support vector machines technique for the training set to detect specific emotions in text with accuracy of 71.64%. Also, in [8] support vector machines are applied with testing their approach on SEMEVAL dataset and got accuracy of 65.54%. Also, in [9] they use Tf-Idf with support vector machines and get accuracy of 71.64%. Another approach by [10] which proposed an improved latent semantic analysis (LSA) to classify emotions and they made an enhancement of 4 % than normal LSA. Finally, lexical affinity [11] has been applied in [12] with applying semantic labeling roles in addition to the support of some search engines like Google for classifying the emotions.

III. Proposed Model

The proposed model consists of a set of determined steps which will be discussed in the next subsections. As any of text mining work, preparing the data resources must to be performed [13]. Preparing data includes data preprocessing tasks which is applied on the input dataset; it includes tokenization, stop words removal, stemming, and lemmatization.

A. Lexicon Preparation

Preparing the lexicon is also one of the main steps that should be performed. We have used two lexicons in the proposed system, they are NRC emotion lexicon (National Research Council of Canada) [14] and SentiWordNet sentiment lexicon [15]. The first step is refining NRC lexicon by mapping the lexicon with the dataset and selecting the related words, then applying a pattern-based technique approach to extend the lexicon with more words that are related to the dataset.

To extend the lexicon, words in the dataset are used, then extraction of the related patterns is performed, and finally more words are extracted from the dataset which follow the same emotion category of its related seeds. The words are related to a seed if it is extracted by the pattern that is related to this seed. This pattern can be defined as the text that is surrounding the word, which means the words that are prior and posterior of the word. We focused on the classification of all words in the NRC lexicon according to only five emotions, they are (Anger, Disgust, Fear, Joy, and Sadness), and Table I show an example of the word classification in NRC.

TABLE I
CROSS SECTION FROM THE NRC LEXICON

Emotion	Value
Anger	0
anticipation	1
disgust	1
fear	0
joy	1
sadness	0
surprise	1
trust	1

After preparing NRC lexicon and extending its words, the SentiWordNet sentiment lexicon is used to extend the sentimental classification of all the words in NRC lexicon by adding the polarity of being either “Negative or Positive”. SentiWordNet is a sentiment lexicon that gives different terms from the WordNet lexicon positive and negative scores according to each term. These scores represent the polarity between (-1) which means the word is negative, (+1) means that the word is positive and (0) means that the word is neutral. This classification is performed using the TF/IDF measure; the whole procedure is discussed in Section III C.

B. TFIDF Weighting

TFIDF stands for term frequency index document frequency which measures the importance of word according to specific document and the whole documents [7]. The measurement reflects how the word is important in the text. First, we calculate TF (term frequency) with in the same document or sentences as the following (1). Also the IDF (index document frequency) will be calculated or IDF in whole documents as follows in (2). After that we multiply the TF by idf to get the value of all words in the training set.

$$tf(t,d) = \frac{0.5 * f(t,d)}{\max\{f(w,d) : w \in d\}} \tag{1}$$

$$idf(t,D) = \log \frac{N}{|\{d \in D : t \in d\}|} \tag{2}$$

TF-IDF measure is applied for preparing the training dataset to provide weighting for each word in each sentence in the training dataset, the following procedure is applied for preparing the training dataset, and we will also reveal the following steps by presenting simple examples.

Using the NRC lexicon, in each word in the sentence, we assign a value which is equal to its weight in the emotion class multiplied by its occurrences in the training set sentences that are following this class, then we calculate the log of the result.

For example, if the sentence “I misunderstood my friend” is following the “anger” class, and the word misunderstand has the occurrences to all the emotion class which exists in the enriched lexicon as (anger = 4, joy = 0, sad = 2, fear = 0, disgust = 0) in addition to assuming the weight of the word in the lexicon to be equal 0.5. Then the “misunderstood” words will have weight in the sentence as (anger = 4*0.5, sad = 2*0.5, joy = 0, fear = 0, disgust = 0). This procedure is repeated for all words in the sentence. And finally, the sentence will have a weighting for each emotion class by summing all the values for all the words for that class and calculate its log.

The second step is by using the SEMEVAL lexicon, an addition to other polarity is applied, it is positive and negative, using the SEMEVAL lexicon.

After preparing the training data, then we apply the same procedure for the testing data with noticing that the weighting and occurrences of each word is calculated according to the training data. While preparing the testing data, the occurrence of the word increase when the word is found in the testing data which in turn update its TF-IDF of this word. This step has revealed to an increase to the accuracy results.

C. Feature Selection Calculations

In the proposed model, seven features are used in the learning procedure; they are five emotions (anger, disgust, fear, joy and sadness) and the two SentiWordNet polarities (positive, negative). In this step, a measurement for weighting each feature for the word is performed according to the following:

- TF-IDF weighting is calculated for each word in the lexicon.
- In the training dataset, the emotion vector is retrieved for each word representing the emotion classification of the word using the NRC lexicon.
- For each class that the word follow, the TF-IDF measure is added as weight of this word in the sentence it is part of.
- After repeating the previous steps on all words in the training dataset, each sentence will have a weighting in all emotion classes according to the TF-IDF measure of the words that constitute the word.
- The same procedure is repeated to calculate the weight of the sentence polarity (positive and negative) according to the SentiWordNet lexicon.
- Finally, each sentence will be classified to seven classes with a determined confidence base on the TF-IDF measure for its constituting words; these classes are

(anger, disgust, fear, joy and sadness, positive, negative)

IV. Case Study Setup

In our case study, we applied the proposed model on the ISEAR dataset [16]. The dataset consists of 7,666 situations divided into the following; 1096 classified as 'anger', 1096 as 'disgust', 1095 'fear', 1095 'joy', 1093 'guilt' and 1096 'shame'. We had 1040 sentences for each of the five emotions that we focus on as mentioned earlier with total sentences equal 5200 sentences. Each sentence from the dataset is classified as (anger, disgust, fear, joy and sadness).

In our experiment, we applied different perspectives, we applied two validation techniques in our experiments, a 10-fold cross validation technique, and 70% training with 30% testing. The second perspective, we applied the experiment twice, the first time with no change in the original lexicons, and the second time after refining and extending the lexicon as previously discussed. The third perspective is considered with the applied classification algorithm, we applied six classification algorithm using WEKA tool [17], they are SMO (Sequential Minimal Optimization) [18], Nearest Neighbor (IBK) [19], KStar [20], Bagging [21], Logistic model trees (LMT) [22], and J48 [23].

V. Experiments and Evaluation

In this section, we demonstrate all the experimental results when applying the proposed model with providing a discussion for these results, the following subsections discuss each group of experiments.

TABLE II
Experiment before Enrichment When Put 70 % Training 30 % Testing

Machine Learning algorithm	Emotions	Precision	Recall	f-measure
Naive Bayes	*Anger*	39.19%	26.36%	31.52%
	Disgust	55.77%	17.58%	26.73%
	Fear	53.61%	26.97%	35.89%
	Joy	31.36%	46.67%	37.52%
	Sadness	22.04%	44.55%	29.49%
	Average	40.40%	32.42%	32.23%
IBK	*Anger*	33.42%	36.97%	35.11%
	Disgust	38.96%	38.48%	38.72%
	Fear	46.35%	44.24%	45.27%
	Joy	44.15%	45.76%	44.94%
	Sadness	31.79%	29.09%	30.38%
	Average	38.93%	38.91%	38.88%
SMO	*Anger*	43.58%	33.94%	38.16%
	Disgust	49.03%	45.76%	47.34%
	Fear	57.81%	41.52%	48.32%
	Joy	46.20%	46.06%	46.13%
	Sadness	34.49%	54.24%	42.17%
	Average	46.22%	44.30%	44.42%
KSTAR	*Anger*	38.29%	36.67%	37.46%
	Disgust	46.15%	41.82%	43.88%
	Fear	51.26%	43.03%	46.79%
	Joy	42.75%	51.82%	46.85%
	Sadness	35.75%	38.79%	37.21%
	Average	42.84%	42.42%	42.44%
Bagging	*Anger*	38.33%	34.85%	36.51%
	Disgust	43.10%	46.36%	44.67%
	Fear	47.62%	45.45%	46.51%
	Joy	45.03%	52.12%	48.31%
	Sadness	34.90%	31.52%	33.12%
	Average	41.80%	42.06%	41.83%
LMT	*Anger*	37.46%	33.03%	35.10%
	Disgust	47.85%	47.27%	47.56%
	Fear	53.11%	49.09%	51.02%
	Joy	42.73%	56.06%	48.49%
	Sadness	39.32%	35.15%	37.12%
	Average	44.09%	44.12%	43.86%
J48	*Anger*	33.84%	33.94%	33.89%
	Disgust	41.44%	41.82%	41.63%
	Fear	40.22%	43.64%	41.86%
	Joy	40.80%	40.30%	40.55%
	Sadness	32.12%	29.39%	30.70%
	Average	37.68%	37.82%	37.72%

A. Applying the Proposed Model without Applying Lexicon Refinement or TF-IDF Weighting

In this experiment, a direct weighting for the emotion in the sentence is performed by summing the weights of the words' emotions. The lexicons are used in their original form without applying the refinement step; our target of this experiment is to prove the impact of the refinement step, in addition to the impact of applying the TF-IDF weighting step. Table II presents the results of the applied experiment with dividing the dataset 70% training data and 30% testing distributed equally on the five emotions based on the six different machine learning algorithms to show the best results machine learning algorithm used. This Experiment shows that SMO algorithm is the best with f-measure 44.42 %. The experiment is also applied using a 10-fold cross validation technique which shows that LMT algorithm produces the best average results with f-measure 64.78 % Table III.

TABLE III
10-FOLD CROSS VALIDATION BEFORE ENRICHMENT

Machine Learning algorithm	Emotions	Precision	Recall	f-measure
Naive Bayes	*Anger*	*59.31%*	*42.88%*	*49.78%*
	Disgust	*87.79%*	*32.50%*	*47.44%*
	Fear	*66.18%*	*38.94%*	*49.03%*
	Joy	*40.40%*	*58.46%*	*47.78%*
	Sadness	*31.96%*	*59.81%*	*41.66%*
	Average	*57.13%*	*46.52%*	*47.14%*
IBK	*Anger*	*50.29%*	*50.87%*	*50.57%*
	Disgust	*56.80%*	*54.62%*	*55.69%*
	Fear	*58.83%*	*58.27%*	*58.55%*
	Joy	*57.69%*	*60.58%*	*59.10%*
	Sadness	*50.39%*	*49.71%*	*50.05%*
	Average	*54.80%*	*54.81%*	*54.79%*
SMO	*Anger*	*69.35%*	*55.48%*	*61.65%*
	Disgust	*71.94%*	*61.63%*	*66.39%*
	Fear	*76.98%*	*60.77%*	*67.92%*
	Joy	*64.87%*	*61.25%*	*63.01%*
	Sadness	*47.61%*	*76.63%*	*58.73%*
	Average	*66.15%*	*63.15%*	*63.54%*
KSTAR	*Anger*	*58.49%*	*54.33%*	*56.33%*
	Disgust	*64.66%*	*59.81%*	*62.14%*
	Fear	*65.33%*	*58.17%*	*61.55%*
	Joy	*56.81%*	*65.38%*	*60.80%*
	Sadness	*53.61%*	*59.23%*	*56.28%*
	Average	*59.78%*	*59.38%*	*59.42%*
Bagging	*Anger*	*62.92%*	*59.23%*	*61.02%*
	Disgust	*68.41%*	*64.13%*	*66.20%*
	Fear	*67.32%*	*66.35%*	*66.83%*
	Joy	*60.82%*	*69.71%*	*64.96%*
	Sadness	*57.92%*	*57.31%*	*57.61%*
	Average	*63.48%*	*63.35%*	*63.33%*
LMT	*Anger*	*65.11%*	*57.60%*	*61.12%*
	Disgust	*69.57%*	*67.50%*	*68.52%*
	Fear	*72.07%*	*64.52%*	*68.09%*
	Joy	*62.25%*	*69.62%*	*65.73%*
	Sadness	*56.92%*	*64.42%*	*60.44%*
	Average	*65.19%*	*64.73%*	*64.78%*
J48	*Anger*	*53.20%*	*51.15%*	*52.16%*
	Disgust	*59.18%*	*59.81%*	*59.49%*
	Fear	*60.50%*	*62.60%*	*61.53%*
	Joy	*59.10%*	*60.29%*	*59.69%*
	Sadness	*52.67%*	*51.25%*	*51.95%*
	Average	*56.93%*	*57.02%*	*56.96%*

B. Applying the Proposed Model with Applying Lexicon Refinement and TF-IDF Weighting

In this experiment, the lexicons are used after applying the refinement step, and the sentences are classified in the training data by applying the proposed TF-IDF measures as previously discussed in the proposed model. Table IV and V presents the results of using the enriched lexicon and the proposed weighting technique with different machine learning algorithms which shows that LMT is the best results with f-measure 61.36 % with about 18 % increase in the accuracy.

C. Experiments after Enrichment and Using SentiWordNet Classification Weighting Measures

In this experiment, SentiWordNet classification is used for all sentences as previously discussed in the proposed model with applying TF-IDF measure for weighting the polarity of the sentences. This step has revealed to an increase of the results' accuracy to be 62.93% on the accuracy in the Table VI for SMO algorithm and 76.12 % for LMT algorithm in Table VII.

TABLE IV
10-FOLD CROSS VALIDATION BEFORE ENRICHMENT

Machine Learning Algorithm	Emotions	Precision	Recall	f-measure
Naive Bayes	*Anger*	49.78%	33.94%	40.36%
	Disgust	34.33%	62.73%	44.37%
	Fear	73.01%	36.06%	48.28%
	Joy	40.97%	66.67%	50.75%
	Sadness	56.56%	20.91%	30.53%
	Average	50.93%	44.06%	42.86%
IBK	*Anger*	50.79%	48.79%	49.77%
	Disgust	55.16%	51.82%	53.44%
	Fear	57.69%	68.18%	62.50%
	Joy	58.47%	62.73%	60.53%
	Sadness	49.46%	41.82%	45.32%
	Average	54.32%	54.67%	54.31%
SMO	*Anger*	55.33%	56.67%	55.99%
	Disgust	65.20%	58.48%	61.66%
	Fear	74.32%	66.67%	70.29%
	Joy	56.21%	72.73%	63.41%
	Sadness	55.97%	49.70%	52.65%
	Average	61.41%	60.85%	60.80%
KSTAR	*Anger*	52.14%	55.45%	53.74%
	Disgust	62.89%	55.45%	58.94%
	Fear	67.07%	66.67%	66.87%
	Joy	60.37%	68.79%	64.31%
	Sadness	54.93%	50.61%	52.68%
	Average	59.48%	59.39%	59.31%
Bagging	*Anger*	52.96%	56.97%	54.89%
	Disgust	61.07%	55.15%	57.96%
	Fear	63.66%	70.61%	66.95%
	Joy	61.74%	64.55%	63.11%
	Sadness	57.69%	50.00%	53.57%
	Average	59.42%	59.45%	59.30%
LMT	*Anger*	53.20%	57.88%	55.44%
	Disgust	62.78%	60.30%	61.51%
	Fear	69.25%	70.30%	69.77%
	Joy	61.71%	67.88%	64.65%
	Sadness	60.87%	50.91%	55.45%
	Average	61.56%	61.45%	61.36%
J48	*Anger*	48.04%	52.12%	50.00%
	Disgust	55.59%	49.70%	52.48%
	Fear	60.99%	67.27%	63.98%
	Joy	57.67%	61.52%	59.53%
	Sadness	50.18%	42.73%	46.15%
	Average	54.50%	54.67%	54.43%

D. Comparison with Other Systems

We compared our system with [24] which used support vector machines and stemming for building his system. In Table VIII, we compare our evaluation to this paper.

E. Discussing the Experiments' Results

We have applied a set of result to clarify the impact of each step in the proposed model. We had the best result when applying a direct classification for the dataset with a direct frequency weighting by to be 64.78%, while after applying the lexicon refinement and the TF-IDF measure for weighting the features, we had an accuracy of 75.72%, and finally when applying the step of SentiWordNet classification, the measure proves that it revealed to the best accuracy of all the experiments with a value of 76.12%. Moreover, the comparison is applied between the proposed model and other research which used the same dataset which also revealed to the advancement of our proposed model in emotion classification.

TABLE V
10-FOLD CROSS VALIDATION BEFORE ENRICHMENT

Machine Learning Algorithm	Emotions	Precision	Recall	f-measure
Naive Bayes	*Anger*	58.57%	46.35%	51.74%
	Disgust	48.02%	58.37%	52.69%
	Fear	75.51%	42.69%	54.55%
	Joy	41.73%	78.85%	54.58%
	Sadness	63.21%	34.04%	44.25%
	Average	57.41%	52.06%	51.56%
IBK	*Anger*	63.36%	62.02%	62.68%
	Disgust	64.69%	65.00%	64.84%
	Fear	72.16%	71.54%	71.85%
	Joy	68.94%	70.87%	69.89%
	Sadness	61.91%	61.73%	61.82%
	Average	66.21%	66.23%	66.22%
SMO	*Anger*	73.87%	72.60%	73.23%
	Disgust	80.46%	73.27%	76.70%
	Fear	85.43%	75.58%	80.20%
	Joy	69.37%	82.12%	75.21%
	Sadness	69.91%	72.60%	71.23%
	Average	75.81%	75.23%	75.31%
KSTAR	*Anger*	68.30%	69.62%	68.95%
	Disgust	76.45%	69.90%	73.03%
	Fear	80.12%	75.19%	77.58%
	Joy	70.91%	79.23%	74.84%
	Sadness	67.65%	68.37%	68.01%
	Average	72.69%	72.46%	72.48%
Bagging	*Anger*	68.90%	70.10%	69.49%
	Disgust	74.30%	71.73%	72.99%
	Fear	78.41%	77.88%	78.15%
	Joy	72.68%	77.50%	75.01%
	Sadness	70.38%	67.40%	68.86%
	Average	72.94%	72.92%	72.90%
LMT	*Anger*	72.61%	72.40%	72.51%
	Disgust	77.96%	75.87%	76.90%
	Fear	82.84%	77.98%	80.34%
	Joy	73.01%	79.33%	76.04%
	Sadness	72.74%	72.88%	72.81%
	Average	75.83%	75.69%	75.72%
J48	*Anger*	63.16%	65.77%	64.44%
	Disgust	69.90%	67.21%	68.53%
	Fear	72.30%	74.04%	73.16%
	Joy	70.13%	71.35%	70.73%
	Sadness	65.29%	62.40%	63.82%
	Average	65.29%	62.40%	63.82%

VI. CONCLUSION

In this research, we proposed a model for enhancing the accuracy results for emotion classification machine learning algorithms, the model is based on applying TF-IDF measure for providing weighting scheme on the classifying features. The proposed model also depended on integrating two lexicon having two polarities in classification with adapting the lexicons by the refinement step. The proposed model has been verified by applying different experiments in addition to comparing the system with other research. The results of the experiments have revealed to the best accuracy measure which reached 776.12 % when applying the proposed model with LMT algorithm.

TABLE VI
EXPERIMENT AFTER ENRICHMENT WHEN PUT 70 % TRAINING 30 % TESTING AND ADD SENTIMENT FEATURES

Machine Learning Algorithm	Emotions	Precision	Recall	f-measure
Naive Bayes	*Anger*	48.50%	29.39%	36.60%
	Disgust	40.29%	33.94%	36.84%
	Fear	63.85%	41.21%	50.09%
	Joy	33.57%	85.76%	48.25%
	Sadness	46.55%	16.36%	24.22%
	Average	46.55%	41.33%	39.20%
IBK	*Anger*	49.53%	47.88%	48.69%
	Disgust	54.69%	51.21%	52.90%
	Fear	57.67%	66.06%	61.58%
	Joy	61.22%	63.64%	62.41%
	Sadness	53.16%	48.48%	50.71%
	Average	55.25%	55.45%	55.26%
SMO	*Anger*	56.93%	57.27%	57.10%
	Disgust	68.31%	58.79%	63.19%
	Fear	74.83%	65.76%	70.00%
	Joy	58.22%	79.39%	67.18%
	Sadness	57.82%	51.52%	54.49%
	Average	63.22%	62.55%	62.39%
KSTAR	*Anger*	47.97%	50.00%	48.96%
	Disgust	54.67%	47.88%	51.05%
	Fear	61.19%	62.12%	61.65%
	Joy	61.41%	66.06%	63.65%
	Sadness	51.07%	50.61%	50.84%
	Average	55.26%	55.33%	55.23%
Bagging	*Anger*	53.31%	58.48%	55.78%
	Disgust	61.82%	55.45%	58.47%
	Fear	63.99%	70.00%	66.86%
	Joy	63.04%	66.67%	64.80%
	Sadness	56.03%	47.88%	51.63%
	Average	59.64%	59.70%	59.51%
LMT	*Anger*	52.25%	56.36%	54.23%
	Disgust	62.54%	59.70%	61.09%
	Fear	69.21%	71.52%	70.34%
	Joy	63.16%	69.09%	65.99%
	Sadness	59.93%	50.30%	54.70%
	Average	61.42%	61.39%	61.27%
J48	*Anger*	48.25%	50.00%	49.11%
	Disgust	57.19%	51.82%	54.37%
	Fear	60.11%	64.85%	62.39%
	Joy	59.13%	61.82%	60.44%
	Sadness	50.97%	47.58%	49.22%
	Average	55.13%	55.21%	55.11%

Providing an additional step by providing a link between the document content to find the dependency and the causality of the classified sentence, however, can provide more success to our model. Moreover, as the proposed model used only five emotions, more research for adding other emotions can be a further successful extension to our work.

TABLE VII
10-FOLD CROSS VALIDATION AFTER ENRICHMENT AND ADDING SENTIMENT FEATURES

Machine Learning Algorithm	Emotions	Precision	Recall	f-measure
Naive Bayes	*Anger*	58.18%	40.00%	47.41%
	Disgust	47.76%	44.04%	45.82%
	Fear	67.35%	47.40%	55.64%
	Joy	37.27%	85.00%	51.82%
	Sadness	65.88%	26.73%	38.03%
	Average	55.29%	48.63%	47.74%
IBK	*Anger*	59.51%	58.65%	59.08%
	Disgust	62.12%	61.35%	61.73%
	Fear	69.53%	68.46%	68.99%
	Joy	67.17%	68.65%	67.90%
	Sadness	59.57%	60.77%	60.16%
	Average	63.58%	63.58%	63.57%
SMO	*Anger*	75.05%	71.44%	73.20%
	Disgust	81.65%	73.17%	77.18%
	Fear	85.86%	75.29%	80.23%
	Joy	68.20%	86.83%	76.40%
	Sadness	71.40%	71.54%	71.47%
	Average	76.43%	75.65%	75.69%
KSTAR	*Anger*	59.17%	57.69%	58.42%
	Disgust	65.47%	59.62%	62.41%
	Fear	71.59%	67.12%	69.28%
	Joy	64.74%	74.13%	69.12%
	Sadness	60.58%	62.50%	61.52%
	Average	64.31%	64.21%	64.15%
Bagging	*Anger*	69.81%	69.13%	69.47%
	Disgust	74.11%	72.12%	73.10%
	Fear	78.17%	78.17%	78.17%
	Joy	73.31%	79.23%	76.16%
	Sadness	70.93%	67.79%	69.32%
	Average	73.26%	73.29%	73.24%
LMT	*Anger*	73.39%	72.12%	72.74%
	Disgust	78.32%	75.38%	76.83%
	Fear	83.66%	78.27%	80.87%
	Joy	72.25%	82.60%	77.07%
	Sadness	73.99%	72.21%	73.09%
	Average	76.32%	76.12%	76.12%
J48	*Anger*	64.04%	64.04%	64.04%
	Disgust	66.16%	66.35%	66.25%
	Fear	72.07%	72.21%	72.14%
	Joy	70.68%	71.63%	71.16%
	Sadness	63.08%	61.92%	62.49%
	Average	65.29%	62.40%	63.82%

TABLE VIII
EVALUATION COMPARISON

System	F1	Accuracy
[24]	67.5	67.4
Our proposed	*76.12%*	*76.12%*

REFERENCES

[1] Bing Liu, Sentiment Analysis and Opinion Mining. Chicago, USA: Morgan & Claypool Publishers, 2012.

[2] Manish Sharma and Rahul Patel, "A Survey on Information Retrieval Models, Techniques and Applications," International Journal of Emerging Technology and Advanced Engineering, vol. 3, no. 11, pp. 542-545, 2013. (Online). http://www.ijetae.com/files/Volume3Issue11/IJETAE_1113_90.pdf

[3] Amira F. El Gohary, Torky I. Sultan, Maha A. Hana, and Mohamed M. El Dosoky, "A Computational Approach for Analyzing and Detecting Emotions in Arabic Text," International Journal of Engineering Research and Applications (IJERA), vol. 3, no. 3, pp. 100-107, May-Jun 2013. (Online). http://www.ijera.com/papers/Vol3_issue3/S33100107.pd

[4] Xuren Wang and Qiuhui Zheng, "Text Emotion Classification Research Based on Improved Latent Semantic Analysis Algorithm," in Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013), 2013.
[5] Barbara Martinazzo, Mariza Miola Dosciatti, and Emerson Cabrera Paraiso, "Identifying Emotions in Short Texts for Brazilian Brazilian Portuguese," in Brazilian conference on intelligent systems, Redes Neurais, 2012.
[6] Christian S. Perone, "Machine Learning: Cosine Similarity for Vector Space Models", (2013) pyevolve.sourceforge. (Online). http://pyevolve.sourceforge.net/wordpress/?p=2497
[7] Dan Jurafsky and Christopher Manning. Coursera. (Online). https://class.coursera.org/nlp/lecture/187
[8] D. Inkpen, F. keshtkar, and D. Ghazi, "Analysis and Generation of Emotion In Texts," International Conference on Knowledge Engineering Principles and Techniques, 2009.
[9] Chaitali G. Patil and Sandip S. Patil, "Use of Porter Stemming Algorithm and SVM for Emotion Extraction from News Headlines," International Journal of Electronics, Communication & Soft Computing Science and Engineering, vol. 2, no. 7, pp. 9-13, 2013.
[10] Landauer T. K. and S. T. Dumais, "A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge," Psychological Review, vol. 104, no. 1, pp. 211-240, 1997.
[11] Rish, Irina. (2001). "An empirical study of the naive Bayes classifier". IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence. (available online: PDF (http://www.research.ibm.com/people/r/rish/papers/RC22230.pdf), PostScript (http://www.research.ibm.com /people/r/rish/papers/ijcai-ws.ps))
[12] Cheng-Yu Lu, Jen-Shin Hong, and Samuel Cruz-Lara, "Emotion Detection in Textual Information by Semantic Role Labeling and Web Mining Techniques," in the Third Taiwanese-French Conference on Information Technology, Nancy/France, 2006.
[13] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze, Introduction to Information Retrieval. Cambridge: Cambridge University Press, 2008.
[14] Saif Mohammad and Peter Turney, "Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon," in the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, California, 2010.
[15] Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani, "SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining," in Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10), Valletta, 2010, pp. 2200–2204.
[16] Mikula, G., Scherer, K. R., & Athenstaedt, U. (1998). The role of injustice in the elicitation of differential emotional reactions. Personality and Social Psychology Bulletin, 24(7), 769-783
[17] Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten Mark Hall. (2009) the University of Waikato. (Online). http://www.cs.waikato.ac.nz/ml/weka/
[18] J. Platt, "Fast Training of Support Vector Machines using Sequential Minimal Optimization," in Advances in Kernel Methods - Support Vector Learning, B. Schoelkopf and C. Burges and A. Smola, Ed. UN: MIT Press, 1998, pp. 1-.
[19] B. S., Landau, S., Leese, M. and Stahl, D. Everitt, Miscellaneous Clustering Methods, in Cluster Analysis, 5th ed. Chichester, UK: John Wiley & Sons, 2011.
[20] G. Cleary and Leonard, E. Trigg John, "K*: An Instance- based Learner Using an Entropic Distance Measure,", 1995.
[21] Amit, Anshuman Sahu, Daniel Apley, and George Runger Shinde, "Preimages for Variation Patterns from Kernel PCA and Bagging," IIE Transactions, vol. 46, no. 5, pp. 1-, 2014.
[22] Niels Landwehr and Mark Hall and Eibe Frank, "Logistic Model Trees," Machine Learning, vol. 95, no. 1, pp. 161-205, 2005.
[23] data-mining business-intelligence. (Online). http://data-mining.business-intelligence.uoc.edu/home/j48-decision-tree
[24] T., & Alpkocak, A. Danisman, "Feeler: Emotion classification of text using vector space model," AISB 2008 Convention Communication, Interaction and Social Intelligence, vol. 1, no. 1, p. p. 53, 2008.