Technical Journal, University of Engineering and Technology (UET) Taxila, Pakistan Vol. 24 No. 4-2019
ISSN:1813-1786 (Print) 2313-7770 (Online)

Sentiment Analysis of Social Media Contents Using Machine Learning Algorithms

R.I.Ahmed[1], M.J.Iqbal[2], M.Bakhsh[3]

[1] *Department of Computer Science, Virtual University of Pakistan*
[2] *Department of Computer Science, University of Engineering and Technology Taxila, Pakistan*
[3] *Department of Computing Abasyn University Peshawar and Pakistan Academy for Rural Development Peshawar*

[2] javed.iqbal@uettaxila.edu.pk

Abstract- In natural language processing, sentiment analysis of social media contents is proven to be very effective to analyse huge and complex amount of unstructured data for better decision making. Social media provides an online environment for the users to show their behaviours and emotions through tweets and post etc. Sentiment analysis of any written text especially social media content is applicable to extract the opinions, emotions and meaningful insights. There are many challenges in the accurate sentiment analysis of available social media contents. The challenges can be both technical and theoretical. Several techniques have been suggested in the past but those failed to overcome the mentioned issues in an optimal way because within their work limited datasets were evaluated. The proposed methodology is helpful to overcome above mentioned issues in data acquisition, feature encoding, data pre-processing, feature selection, and classification. In feature encoding phase, a hybrid approach of bi-gram and tri-gram is used for embedding of words. In the experiments, several benchmark datasets have been utilized to measure the effectiveness of the proposed framework. The proposed methodology gives better or at least comparable results with maximum confidence and with less computational complexity. The average accuracy results were in the range from 89-91 with the multilayer perceptron neural network. The mechanism of this work will be helpful to enhance the sentiment analysis process of multifaceted types of social media and blogs contents..

Keywords- Sentiment Analysis, Machine Learning, Social Media, Tweets, Multilayer Perceptron Neural Network.

I. INTRODUCTION

Due to the progress and innovations in information technology, the world has become a global village. The modern age can be broadly described as digital age [1]. Our dependence on IT based systems is increasing and it results in generation of massive amount data. The data in which we are interested is comprised of text or alpha numeric. Text mining is accountable for explicit aggregation in the conferred text. Social media has become a basic information gathering tool. We can make shopping, exchange opinions and get different kinds of services without going outside. A large percentage of people around the globe are using social media like Facebook, Twitter and WhatsApp. This percentage of people is increasing on daily basis. Popularity of social media is also increasing day by day due to innovations and developments in various cutting-edge technologies. Different companies are using social media for marketing and advertisement purpose. These companies also deal with their clients online. They offer their products to clients for purchase [2]. After choosing the products, the clients place their orders and make payments via credit card. In order to be more successful, it is necessary for the companies to understand the needs, behaviours and emotions of the clients. Measurement of the success and usefulness of the product is essential too. Huge amount of reviews, articles and feedbacks have been created daily in the form of text. It is very difficult to manage big amount of data due to its polarity and complex nature. Text mining plays its vital part in classification of text data in order to extract useful and meaningful information from such corpuses [3]. Text mining is used in all those fields where we need to get important data from huge amount of data and social media is one of these fields.

Sentiment analysis uses natural language processing (NLP) techniques to identify the exact meaning of the opinion, behaviour and attitude [4]. Sentiment analysis shapes the feelings of a spokesman or user with respect to some subject matter. The opinion or emotion maybe one's decision or the expressive state while writing. Sentiment analysis is also referred as opinion mining and proven to be very helpful process in polarity detection. Recently, many machine learning-based techniques have been proposed for pre-processing, classification and analysis of social media generated text data. However, these has some challenges like spam and fake, domain dependency, negation,

NLP overhead, bi-polar words and huge lexicon [5]. In order to make data mining process more effective and efficient, it is very important to overcome above mentioned issues. Previously, researches have conducted their research on sentiment analysis and its challenges. We have chosen sentiment analysis of social media contents using proposed framework because of its significance and impact on any society.

After knowing the importance of sentiment analysis, the mechanism of this work will be helpful to improve the sentiment analysis process in social media. The proposed methodology gives better or at least comparable results with maximum confidence and with less computational complexity.

The rest of paper is structured as follows. Section II contains literature review of relevant work on sentiment analysis. Section III presents methodology adopted in this work. Results are given in Section IV and finally the conclusion of the proposed work is presented in Section V.

II. Literature Review

Several authors have proposed techniques to mitigate issues related to analysis and mining of social media sentiment analysis. Following is the existing work done by different researchers on sentiment analysis.

In [6], authors proposed a fast method for sentiment detection from the piece of text which presents the user's emotions in various languages. The work focused on ConvLstm architecture along with word embedding and lexicon based methods. In this architecture convolution neural network and Long Short-Term Memory (LSTM) are employed on the top of vectors representing the word [7]. Their evaluations showed that convolution neural network oppressed long short-term memory as alternate of merging layer of CNN so that it can minimize the risk of redundant information, creates easiness and deal with long term dependency in corpus. In [8], authors explored machine learning and lexicon-based approaches on Twitter and Facebook data which is collected from tweets and Facebook datasets. The result of their experiments showed that the performance of lexicon-based classifiers was satisfactory. Because of pre-processing and removal of the unnecessary texts, the accuracy was increased. Authors concluded that lexicon-based method could be very effective and efficient, as compared to machine learning based approaches. This method of analysing social media data provides very effective sentiment evaluation techniques [9].

In vector space textual data is represented with the help of vector identifiers [10]. It is mostly utilized in data retrieval and indexing. It was firstly used in Statistic based Retrieval System. Words in every document are represented in the vector form. If the document contains the requested term, then its value is considered as non-zero otherwise the value of term is zero. Both query and document are symbolized in vector forms and weighs are assigned to these vectors. After performing necessary calculations, the similarity is determined between the vectors. Various search engines use VSM to search specific information on the internet. Unlike Boolean model, VSM provides result which are based on ranking. Several methods are used for this calculation. Similarity between query documents are determined by using TD-IDF method.

On social media the violent content has been creating lots of problems by hurting religious and political matters. Researcher are working to avoid this kind of contents from social media. In this, the authors performed their analysis hierarchically and tried to derive a mechanism to detect offensive contents on social media [11]. They used Offensive Language Identification Dataset (OLID). They discussed the core correspondences and variations between OLID and pre-existing datasets for odium dialog identification, aggression detection, and similar responsibilities. In the end they perform testing and training activity to compare the outcomes of various machine learning techniques on OLID.

Machine learning based techniques work with the training of an algorithm for a dataset earlier and then apply it to unknown data. In 2016, a comparison between lexicon approach based and machine learning approaches were conducted [12]. In this comparison, naïve Bayes, Support Vector Machine (SVM) and Maximum Entropy algorithms were applied on Twitter dataset. All mentioned algorithms are considered as machine learning based algorithms. After their investigations they showed that the accuracy rate of NB is74.44% and the accuracy rate of the SVM is 77.73%. The SVM showed better results in their comparison.

The authors in [13] implemented deep learning and linear machine learning algorithm to sentiment classification. Secondly, two techniques were proposed that aggregated their standard classifier with other classifiers commonly used in Sentiment Analysis [14]. Third, they also suggested two approaches in order to combine both proposed and deep learning approaches to fuse information from numerous resources. Fourth, they classified different models with respect to their categories which were proposed in the literature. Fifth, they had performed different performance evaluation tests to measure the effectiveness of those models.

According to this study, deep learning techniques for sentiment analysis have emerged as popular methods [15]. Authors gave automatic function extraction and richer illustrations. Result showed improved performance than traditional characteristic based strategies. Traditional methods are based on manually extracted complicated features. Deep learning has appeared as a good alternative to machine learning technique that works well and gave improved results [16]. With this improved achievement, deep learning-

based techniques are playing its effective role in many tasks related to numeric, text or image data [17].
The authors had compared different techniques of conducting sentiment analysis based on different frameworks [18]. Accordingly, the sentiment analysis of the users were evaluated for better decision making.

TABLE I: COMPARISON OF DIFFERENT SENTIMENT ANALYSIS TECHNIQUES FROM LITERATURE

Author's Name	*Dataset(s)*	*Encoding Technique*	*Classifiers*	*Results*	*Limitations*	*Complexity*
(Giatsoglou, Vozalis et al. 2017)	Movies IMDB	Word2vec	Lex1, Lex2 Lex3, Lex4	Hybrid vector shown better result	Difficult to handle OOV words.	Medium
(Yousefpour, Ibrahim et al. 2017)	Movies Books Music	Word2vec	SBN NB ME	POS-based features shown effectiveness	To select an optimal Feature subset is to an exhaustive.	High
Araque, O., et al. (2017).	Twitter Movies Google	Word2vec	Fixed rule model, Meta model.	Google Net perform higher then baseline classifiers	I. Domain dependence. II. Negation. III. Topic nature.	Low
Vateekul, P. and T. Koomsubha (2016)	Twitter	Word2vec	LSTM DCNN	DCNN shown better results	It is critical to mine a large and relevant sample of data	Medium
(Hassan and Mahmood 2017)	SST Bank, IMDB	Word2vec	SBN NB	Accuracy of 85.86% on STS dataset	The dataset size affects the deep learning.	High
(Baroni, Dinu et al. 2014)	English Wikipedia ,	Word2vec	ME LDF	ConvLstm model performance was satisfactory	I. Domain dependence. II. Huge lexicon.	Low
(Ain, Ali et al. 2017)	SentiBank Twitter, T&C News	CNN, Word2vec.	NB MLP	Satisfactory	I. Requires large data sets. II. Costly to train.	Medium
(Zampieri et al., 2019)	Twitter	Word2vec	NB SVM	A lexicon-based approach was better	I. Spam& fake. II. Bi-Polar words.	High
(Kharde and Sonawane 2016)	Twitter	Lexicon-based	NB SVM.	Accuracy NB= 74.44% SVM= 77.73%	I. Negation. II. Domain dependence. III. Huge lexicon.	Low

In their studies, the authors proposed frequency-based integration of two methods for feature vector representation and feature subset selection [19]. An ordinal based integration of various feature vectors was proposed to attain simple feature vectors. The attained features are dependent on the sequence of the features used in the previous vectors. The final feature vector was gained from a hybrid method of wrapper and filter in the feature selection step. The authors conducted a review on the challenges and issues in sentiment analysis. The clarity of sentiments was beneficial for people related to any field of life. The given text was said meaningful only when the mentioned text is refined with the help of organized technique of text mining and sentiment analysis [20]. But there were many issues and challenges in the accurate and reliable sentiment analysis of social media content. The challenges can be technical and theoretical. The authors explored latest advancements in Recurrent Neural Networks (RNN) for large scale Language Modeling [21]. They also worked to overcome the challenges and issues in recurrent neural network. Two major issues in RNN task were vocabulary sizes and long term structure of language.

In this paper, authors performed a comprehensive evaluation on several lexical semantics tasks with several parameter settings [22]. By using sentiment analysis, the authors analysed the data from social media tweets and posts. In their analysis, they evaluated the emotions, attitudes and opinions of the users [23]. The others tried to develop a dictionary for the word used in social media.

Finally, the analysis of literature review shows that there are still some challenges in sentiment analysis and evaluation of social media contents. The proposed work aims to mitigate these and enhances overall system performance.

The strengths and weakness of the literature is given in Table I. It also represents the summary of all the paper which are reviewed.

III. Methodology

The following framework is implemented to carry out work on sentiment analysis. The proposed framework consists of various phases. In each phase, specific tasks were performed. The next subsections provide detailed illustration of each phase of the framework. Moreover, the implemented methodology is graphically represented in Figure 1.

A. Extraction of Data

In sentiment analysis, the first step is extraction of data from the benchmark data sources. The data is extracted from tweets, posts, comments and product reviews. Search criteria are defined before the extraction of data, search for topics, and extract social media data from required media. Datasets are extracted with the use of

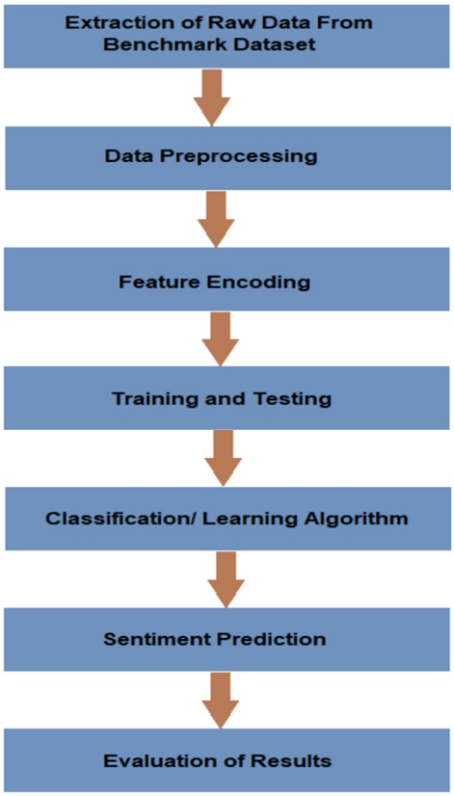

Fig 1: Block Diagram of Sentiment Analysis System

different application specific APIs, crawlers, benchmarks and it can also be obtained manually form the related websites. In the proposed work, some data is extracted manually from the related sites and benchmark datasets. The remaining data is extracted from the datasets used by other researchers in the past. That data is available on the different websites. Commonly used datasets are about Twitter, Movies, News, Reviews and Facebook. The extracted data is used for data analysis and insightful knowledge extraction. The details of the extracted datasets are shown in Table II.

B. Data Pre-processing

Data pre-processing is a vital and critical phase in data analysis and mining [24]. A huge amount of complexity occurs due to the duplications and redundant information in tweets, posts and reviews etc. Data pre-processing is used as a filtering tool to normalize the data. Data pre-processing includes normalization, Erase Punctuation, convert the text data to lowercase, tokenization of the text [25], removal of stop words, normalize the words using the Porter stemmer, removal of the hyperlink, removal of hash tag, conversion of the abbreviations and flying words into original word, translating other languages into English, removal of unnecessary spaces, POS tagging and conversion of emotion into meaningful text. Conversion of Emotion into meaningful Text will be a new activity in

pre-processing phase. The use of emoji's is increasing rapidly; that is why it is important to convert them into text form.

C. Feature Encoding

The extracted datasets may not be in a format, suitable for performing any statistical or mathematical operations. There is a need for a suitable feature encoding method that extracts numeric features from the available text data [26]. We have to propose mathematical model that properly represents each tweets of the sample in a way that captures the true or actual semantic of words or phrases in it. The proposed numeric features are then used in the next phase of the methodology for further processing and analysis.

i. Word2vec with Bi-gram and Tri-gram

The used of word2vec technique with n-gram is found to be more effective and shows better performance. Word2vec is a collection of correlated models which might be utilized to supply word embedding [27]. These models are shallow, layered neural networks which might be skilled to rebuild semantic contexts of words. A huge corpus data is taken by the word2vecas input and convert it into vector form. This graphical representation become easy to understand and evaluate. In this work, the combination of both bi-gram and tri-gram is used for word embedding. The hybrid approach of bi-ram and tri-gram will increase the accuracy because the effectiveness of both techniques will give the high rate of accuracy.

Figure 2 and Figure 3 shows the working of bi and tri gram feature encoding model.

TABLE II: DETAILS OF DATASETS

Sr. No.	Datasets	Positive Instances	Negative Instances	Neutral Instances	Total Instances
1	Stanford Twitter Sentiment (STS)	2532	1758	750	5000
2	Movies	850	835	315	2000
3	SemEval2013	932	861	522	2315
4	FBData	196	138	250	584

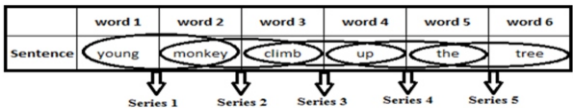

Fig 2: Example of Bi-gram

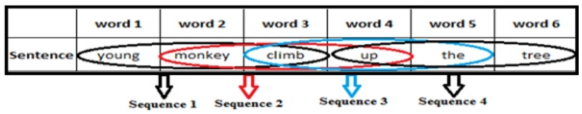

Fig 3: Example of Tri-gram

D. Classification/ Learning Algorithm

As the volume of personal information appears on social networks is increasing constantly, the classification is becoming critical in NLP particularly sentiment analysis [28]. In the classification phase, we have used best performing algorithms such as, multilayer perceptron neural network (MLP), decision tree and support vector machine. A multilayer perceptron is a class of neural network. The MLP consists of minimum three layers of nodes. Except for the input nodes, every node represents a neuron that uses a nonlinear activation function. SVM is a machine learning classifier which is used for supervised text classification [29]. It may be used for regression analysis, but it is mostly used for classification task. The decision tree is a graphical representation of the results of a decision. It is used in data mining to make the complex challenges and decisions simpler. It describes the text data and also classifies the data according to its nature.

Further details about these classifiers can be found from existing literature.

E. Sentiment Prediction

Sentiment prediction phase is useful in the process of sentiment analysis. The prediction results are provided on each dataset with different classification algorithms [30]. After, sufficient training of the model, the sentiment prediction of the query tweets is observed. For generalization of the algorithms, several iterations of technique may be required.

F. Sentiment Evaluation

After performing all above mentioned phases as analyst, we are in a position to describe the polarity of the text. In sentiment evaluation step, the results of the analysis are comprehensively described. It is the stage of shaping the polarity of the text. As for as meaning of the text concern, it can have positive, negative and neutral meanings. Deriving the exact meanings from the text is also term as opinion mining. The results of proposed method are compared with the existing best approaches from the literature. The overall performance of the proposed system is evaluated using performance metrics such as accuracy, specificity, recall, precision and F-measure [31]. The description of each performance measure is as follows.

(*a) Accuracy*

Accuracy is the most prominent performance measure used for properly motive. It is extraordinarily beneficial, easy to compute and recognize. Accuracy measures the capacity of a predictor in successfully identifying all samples, irrespective of its far effective or negative [31].

$$Accuracy = \frac{TP + TN}{P + N} \quad \text{(i)}$$

(b) Specificity

Specificity is termed as the true negative rate. It finds the percentage of real negatives which are efficiently identified as such [32]. In a scientific test specificity is the quantity to which real negatives are classified.

$$Specificity = \frac{TN}{N} \quad \text{(ii)}$$

(c) Sensitivity/Recall

Sensitivity can be termed as the true positive rate. In a few finds the percentage of actual positives which can be efficiently recognized [33]. Higher sensitivity reflects less false negatives and lower sensitivity means more false negatives. Sometimes while we improve the sensitivity, as a result the precision decreases.

$$Sencitivity = \frac{TP}{P} \quad \text{(iii)}$$

(d) Precision

The precision shows the correctness of the classifier. High precision means that less False Positive and low precision means less positive. It is inversely proportion to the sensitivity, the improvement in precision be the cause of lower sensitivity.

$$\Pr ecision = \frac{TP}{TP + FN} \quad \text{(iv)}$$

(e) F-measure

The combination of precision and sensitivity is called frequency measure. It is the weighted harmonic mean of precision and sensitivity. Frequency measure is proved itself as beneficial as accuracy.

$$F - Score = \frac{2 \times precision \times sensitivity}{precision + sensitivity} \quad \text{(v)}$$

IV. Results and Discussion

This section shows experimental results and discussion about the proposed work. The chosen datasets are evaluated by applying MLP neural network, Decision Tree and SVM classifiers.

Figure 4 illustrates experimental results when the Stanford Tweeter Sentiment (STS) Dataset is evaluated through different classifiers (i.e. Decision Tree, MLP, and SVM).

The evaluation of the Movies Review Dataset through different classifiers (Decision Tree, MLP, and SVM) is shown in Figure 5. Finally, Figure 6 and 7 shows the comparison of performance measures when different classifiers are applied on SemEval2013 and FBData dataset.

1. Comparison of Experimental Results

Figure 8 provides a comprehensive comparison of three classifiers (Multilayer Perceptron, Decision Tree and Support Vector Machine) used in the experiments. Table III shows the overall experimental results for all three classifiers. From experimental results, it is clear that the performance of the chosen classifiers is better in all aspects. The tweets in every dataset belongs to three classes such as, Positive, Negative and Neutral. After evaluating all these classes, the overall average performance is shown subsequently. Figure 6 shows that the performance of all the datasets is higher with the MLP classifier. Thus we concluded that Multilayer Perceptron neural network has high prediction about the polarity than other technique. Figure 6 presents that sentiment prediction of Decision Tree and SVM is also notable and these techniques are reliable, but these have slight low performance than multilayer perceptron. On chosen data, MLP has given 83%, 82%, 84% and 82% performance, whereas the performance of SVM and Decision tree remained somewhere between 74% and 78% which is lower than MLP. If the ranges of accuracies for each classifier are observed, then we can say that the multilayer perceptron neural network has comparatively better results.

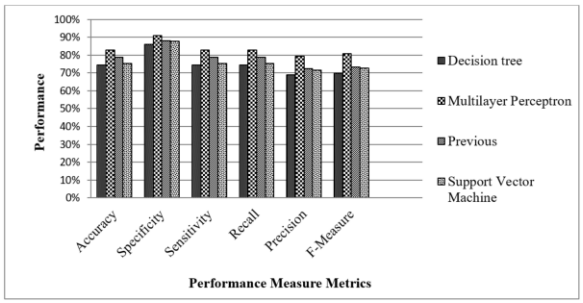

Fig 4: Comparison of Performance Measure Matrices for STS Dataset.

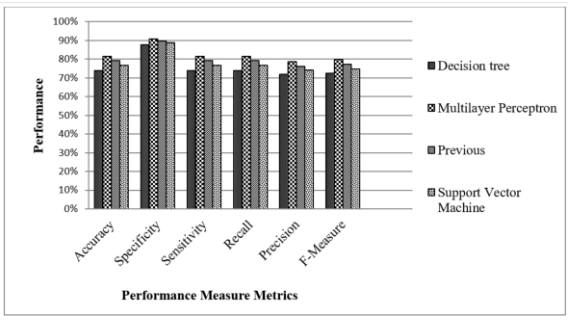

Fig 5: Comparison of Performance Measure Matrices for Movies Dataset.

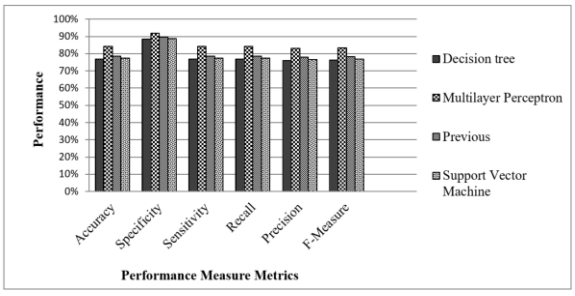

Fig 6: Comparison of Performance Measure Matrices for SemEval2013 Dataset.

Table IV shows comparison of presented work with previous approaches.

After evaluating the chosen data on these techniques, the summary of derived results is shown in Figure 8.

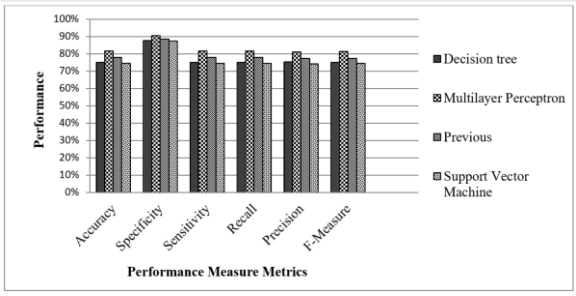

Fig 7: Comparison of Performance Measure Matrices for FBData Dataset.

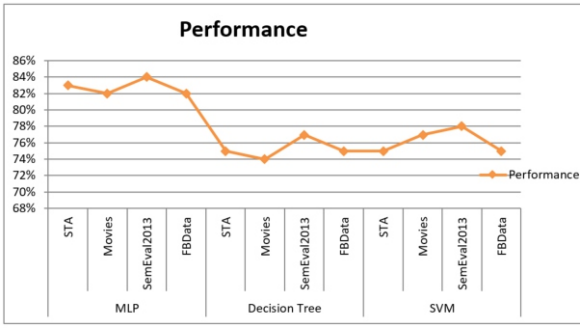

Fig 8: Accuracy Comparison of Different Techniques

TABLE III: COMPARISON OF SENTIMENT ANALYSIS RESULTS USING DIFFERENT CLASSIFIERS.

Classifier	Dataset	Positive	Negative	Neutral	Average
MLP	STS	0.83	0.81	0.85	***0.83***
	Movies	0.82	0.83	0.80	***0.82***
	SemEval2013	0.85	0.81	0.86	***0.84***
	FBData	0.81	0.83	0.82	***0.82***
Decision Tree	STS	0.71	0.67	0.86	***0.75***
	Movies	0.75	0.80	0.67	***0.74***
	SemEval2013	0.80	0.76	0.75	***0.77***
	8FBData	0.77	0.69	0.80	***0.75***
SVM	STS	0.73	0.81	0.72	***0.75***
	Movies	0.77	0.80	0.73	***0.77***
	SemEval2013	0.82	0.74	0.77	***0.78***
	FBData	0.71	0.73	0.79	***0.75***

TABLE IV: COMPARISON WITH PREVIOUS RESULTS

Reference	Dataset	Technique	Specificity	Sensitivity	Accuracy
(Giatsoglou, Vozalis et al. 2017)	MOVIES	SVM	-	-	0.73
Vateekul, P. and T. Koomsubha (2016)	BOLDFACE IS THE WINNER	DCNN	-	-	0.75
		SVM	-	-	0.74
		ME	-	-	0.75
(Kharde and Sonawane 2016)	SSTb	SVM	-	-	0.76
Proposed Work	STS	MLP	0.90	0.83	0.83
		D-Tree	0.86	0.75	0.75
		SVM	0.88	0.75	0.75
	Movies	MLP	0.89	0.82	0.82
		D-Tree	0.88	0.74	0.74
		SVM	0.89	0.77	0.77
	SemEval2013	MLP	0.92	0.84	0.84
		D-Tree	0.89	0.77	0.77
		SVM	0.89	0.78	0.78
	FBData	MLP	0.90	0.82	0.82
		D-Tree	0.88	0.75	0.75
		SVM	0.87	0.75	0.75

It is observed from experimental results that average accuracy of each classifier is varied on different datasets. The main reason of this variation is the complexity of data. Moreover, various other factors also affect the performance of classifiers used, which include,

I. Pre-processing plays a significant role in text mining and analysis. If the process of pre-processing is not performed properly then it is very difficult for any classifier to produce accurate results on the given data. This ambiguity is the cause of poor results in some cases. To avoid this situation, the methods of decimal scaling and zero mean normalization are used.

II. In many cases, noise greatly affects the performance of the classifier. Noise is defined as the typical words gave inconsistent meanings. It is typically a minority of the datasets. Noise produces lots of ambiguity for the classifiers and they produce poor results.

III. Sometimes, data has some attribute which are not understandable for the classifiers. An attribute or set of attributes may have the duplications or they mix up the meanings with each other. It also produces problems for classifiers to produce good results.

IV. As a classifier uses the training dataset to conduct the classification, normally it is expected that the dataset has the instances which are not similar to each other. It must be different from each other. If the data lack the diversity, then it will produce problems for the classifiers. Meanwhile, the size of dataset is also significant, and it affects the performance of the classifiers.

V. The single experiment cannot measure the classification accuracy. To assure the performance of a classifier cross validation is a good exercise. In cross validation, more than one experiments are conducted and the average accuracy of all experiments is treated the final and authentic accuracy.

V. Conclusion

Social media is generating huge types and amount of structured or unstructured complex data on daily basis through Facebook, Twitter, WhatsApp and Viber etc. Sentiment analysis is becoming very effective in analysing this data to extract useful information and finding text's polarity. This work focuses on the proposal of an improved methodology that overcome the issues in existing systems. The methodology adopted in this work is implemented in different phases for highly accurate sentiment analysis and better system performance. The pre-processing and classification phase are very much critical in such type of research work. We have investigated the effect of three classification techniques such as MLP, Decision Tree and SVM for finding polarity of source datasets. The performance is evaluated on four benchmark datasets which includes Stanford Twitter Sentiment (STS), Movies Review, SemEval2013 and FBData. MLP has shown better accuracy results than SVM and decision tree. The major finding of this work is that the accuracy of a classifier is varied with respect to different datasets. Thus, the experiment results are very much dependent on the type and complexity of given data.

Reference

[1] Qin, S.J., *Process data analytics in the era of big data.* AIChE Journal, 2014. 60(9): p. 3092-3100.

[2] Tarhini, A., et al., *An analysis of the factors influencing the adoption of online shopping.* International Journal of Technology Diffusion (IJTD), 2018. 9(3): p. 68-87.

[3] Salloum, S.A., et al., *Using text mining techniques for extracting information from research articles*, in *Intelligent natural language processing: Trends and Applications*. 2018, Springer. p. 373-397.

[4] Collobert, R., et al., *Natural language processing (almost) from scratch.* Journal of machine learning research, 2011. 12(Aug): p. 2493-2537.

[5] Bhati, R.G., *A SURVEY ON SENTIMENT ANALYSIS ALGORITHMS AND DATASETS.* 2019.

[6] Giatsoglou, M., et al., *Sentiment analysis leveraging emotions and word embeddings.* Expert Systems with Applications, 2017. 69: p. 214-224.

[7] Hassan, A. and A. Mahmood. *Deep learning approach for sentiment analysis of short texts*. in *2017 3rd international conference on control, automation and robotics (ICCAR)*. 2017. IEEE.

[8] Vateekul, P. and T. Koomsubha. *A study of sentiment analysis using deep learning techniques on Thai Twitter data*. in *2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. 2016. IEEE.

[9] Bottou, L., F.E. Curtis, and J. Nocedal, *Optimization methods for large-scale machine learning.* Siam Review, 2018. 60(2): p. 223-311.

[10] Abu-Salih, B., *Applying Vector Space Model (VSM) Techniques in Information Retrieval for Arabic Language*. arXiv preprint arXiv:1801.03627, 2018.

[11] Zampieri, M., et al., *Predicting the Type and Target of Offensive Posts in Social Media.* arXiv preprint arXiv:1902.09666, 2019.

[12] Kharde, V. and P. Sonawane, *Sentiment analysis of twitter data: a survey of techniques.* arXiv preprint arXiv:1601.06971, 2016.

[13] Araque, O., et al., *Enhancing deep learning sentiment analysis with ensemble techniques in social applications*. Expert Systems with Applications, 2017. 77: p. 236-246.

[14] Ain, Q.T., et al., *Sentiment analysis using deep learning techniques: a review.* Int J Adv Comput Sci Appl, 2017. 8(6): p. 424.

[15] Goodfellow, I., et al., *Deep learning. vol. 1*. 2016, MIT press Cambridge.

[16] LeCun, Y., Y. Bengio, and G. Hinton, *Deep learning.* nature, 2015. 521(7553): p. 436.

[17] Vyrva, N., *Sentiment analysis in social media*. 2016.

[18] Devika, M., C. Sunitha, and A. Ganesh, *Sentiment analysis: a comparative study on different approaches*. procedia computer Science, 2016. 87: p. 44-49.

[19] Yousefpour, A., R. Ibrahim, and H.N.A. Hamed, *Ordinal-based and frequency-based integration of feature selection methods for sentiment analysis.* Expert Systems with Applications, 2017. 75: p. 80-93.

[20] Cambria, E., et al. *SenticNet 4: A semantic resource for sentiment analysis based on conceptual primitives*. in *Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers*. 2016.

[21] Jozefowicz, R., et al., *Exploring the limits of language modeling*. arXiv preprint arXiv:1602.02410, 2016.

[22] Baroni, M., G. Dinu, and G. Kruszewski. *Don't count, predict! a systematic comparison of context-counting vs. context-predicting semantic vectors*. in *Proceedings of the*

52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2014.

[23] Singh, R. and R. Kaur, *Sentiment analysis on social media and online review.* International Journal of Computer Applications, 2015. 121(20).

[24] Shoukry, A. and A. Rafea. *Preprocessing Egyptian dialect tweets for sentiment mining*. in *The Fourth Workshop on Computational Approaches to Arabic Script-based Languages*. 2012.

[25] Carus, A.B., *Method and apparatus for improved tokenization of natural language text*. 1999, Google Patents.

[26] Kantorov, V. and I. Laptev. *Efficient feature extraction, encoding and classification for action recognition*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.

[27] Joulin, A., et al., *Fasttext. zip: Compressing text classification models*. arXiv preprint arXiv:1612.03651, 2016.

[28] Mohamed, E. *Morphological segmentation and part of speech tagging for religious Arabic*. in *The Fourth Workshop on Computational Approaches to Arabic Script-based Languages*. 2012.

[29] Bottou, L., F.E. Curtis, and J. Gautam, G. and D. Yadav. *Sentiment analysis of twitter data using machine learning approaches and semantic analysis*. in *2014 Seventh International Conference on Contemporary Computing (IC3)*. 2014. IEEE.

[30] Kim, J., et al. *Sentiment prediction using collaborative filtering*. in *Seventh international AAAI conference on weblogs and social media*. 2013.

[31] Caruana, R. and A. Niculescu-Mizil. *Data mining in metric space: an empirical analysis of supervised learning performance criteria*. in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. 2004. ACM.

[32] Cios, K.J. and G.W. Moore, *Uniqueness of medical data mining.* Artificial intelligence in medicine, 2002. 26(1-2): p. 1-24.

[33] Schell, M.J., et al., *Evidence-based target recall rates for screening mammography.* Radiology, 2007. 243(3): p. 681-689.