SENTIMENT ANALYSIS USING NOVEL NORMALISATION, NEGATION HANDLING USING MACHINE LEARNING ALGORITHM

Neha Kumari* , Dr. Mukesh Kumar**
*Department of Computer Science & Engineering
Rabindranath Tagore University, India
Email- @gmail.com
**Department of Computer Science & Engineering
Rabindranath Tagore University, Raisen, Madhya Pradesh, India

Abstract- Due to the progression of technology, there is abrupt usage of microblogging sites such as Twitter for sharing of feelings and emotions towards any current hot topic, any product, services, or any event. Such opinionated data needs to be leveraged effectively to get valuable insight from that data. This research work focused on designing a comprehensive feature-based Twitter Sentiment Analysis (TSA) framework using the supervised machine learning approach with integrated sophisticated negation handling approach and knowledge-based Tweet Normalization System (TNS). We leveraged varieties of features such as lexicon-based features, pos-based, morphological, ngrams, negation, and cluster-based features to ascertain which classifier works well with which feature group. We employed three state-of-the-art classifiers including Support Vector Machine (SVM), for our twitter sentiment analysis framework. We found SVM to be the best performing classifier across all the twitter datasets except #9pm9minutes (DTC turned out to be the best for this dataset). Moreover, our SVM model trained on the SemEval-2013 training dataset outperformed the winning team NRC Canada of SemEval-2013 task 2 in terms of macro-averaged F1 score, averaged on positive and negative classes only.

Keywords: Sentiment Analysis, Twitter Sentiment Analysis, Negation Modelling, Tweet Normalization System, Supervised Machine Learning, Real-Time Twitter Dataset, Benchmark Twitter Dataset.

1. INTRODUCTION

Opinions are the key indicators of one’s behaviours. Subjective feelings such as opinion, emotion, attitude, etc. greatly influenced human behaviour. Our choices and beliefs are somehow conditioned on how others evaluate and feel for the world. Whenever we want to take a decision, we frequently look out for the others opinions. This is not only related to an individual but also for the large organizations, who are always eager to know public opinions on their services or products. Even a person before buying any product, often look out the feedback or reviews provided by the existing customers of that product. For instance, shopping websites such as Amazon.com provides the star ratings and consumer’s reviews for every product they are selling. Such opinionated reviews act as product recommendation for the buyer. There is remarkable progress in the usage of social media like Facebook, Twitter, blogs, etc. by users of different nationalities for expressing their opinions, desires, allegations, and emotions about any services, product or about any topic[1].

In ancient times, before the technological growth, people used to ask their friends, family, and neighbours, whenever an opinion regarding any entity is needed (Liu, 2010). In the same way, organizations in the past arranged for the opinion polls and surveys (using questionnaire) for getting the public opinions on their products or services. That means, in the past opinionated data is limited only. But the explosion of Web 2.0 triggers the people to use the social media (such as blogs, reviews, forums, twitter, etc.) for conveying their opinionated views regarding anything that could be a product, an event, any famous personality[2], politicians, and many more. It is stated by John Scalzi that “Everyone is entitled to their opinion about the things they read (or watch, or listen to, or taste, or whatever). They’re so entitled to express them online”. Thus, opinionated data is no longer limited to friends, neighbours and relatives. It becomes possible to get the opinions from the massive pool of people. .Huge amount of opinionated data is available in digital form from which valuable information can be extracted that in turn would help in taking varieties of decisions regarding any product, service, organization, or topic. Such data is commonly known as User-Generated-Data (UGC). Extracted information from opinionated data have different potential usages i.e. organizations collect the opinionated data which is available in form of comments, reviews, blogs, etc. about their services and products and then analyze those data for extracting valuable information so that decision can be made regarding improvement of their services and products. Even different e-commerce websites use the opioniated data for analyzing the buying pattern of their customers, sentiment of customers regarding their brand, and, then, based on the analysis strategies are made to for sales improvement[3].

2. LEVELS OF SENTIMENT ANALYSIS

1.1.1 Document-level Sentiment Analysis

Document-level analysis is based on the assumption that a document contains direct opinion on a single entity. Thus, it is not suitable in cases (such as blogs and forums) where comparative opinions are present in a document as it provides overall single opinion on a document. Several researchers performed document-level sentiment analysis (Pang & Lee, 2008). It's not a deeper level of analysis and doesn't provide opinion of people on individual feature or an aspect of any product, which is a major disadvantage (Liu, 2012). For example, consider a document regarding Nokia phone:

(1) I bought Nokia phone 2 months ago. (2) Camera quality is good. (3) My sister thought it's very expensive.

In the above document there are 3 sentences expressing different opinions on different aspects of Nokia phone. Sentence (1) is neutral, (2) expresses positive sentiment on camera quality of Nokia phone and (3) expresses negative opinion on price feature. If document-level analysis is done on above document then it's very difficult to get individual opinion on each aspect of Nokia phone as document-level provides overall single sentiment on entire document [9]. Moreover, in reality people are interested in knowing the public reviews on each feature of the product (Nokia phone in the above example) rather than on entire product itself. Hence, there is need of sentence-level or aspect-level sentiment analysis for deeper analysis.

1.1.2 Sentence-level Sentiment Analysis

The main aim of sentence-level analysis is to determine the polarity of sentence i.e. whether it is expressing positive, negative or neutral opinion. It is based on the assumption that, a sentence articulates single opinion. In this level, firstly analysis is done to determine whether the sentence is objective or subjective that is separating feelings, beliefs, or views from the facts. That is subjective versus objective classification is done. Subjective sentences are those which contain expressed opinions, feelings, etc. that is non-factual information is represented by them. For instance, Camera quality of Nokia is good. Most of the subjective sentences contain sentiment bearing words such as love, good, etc. but few many not express opinions. Such type of subjective text needs to be handled carefully. For instance, "He came yesterday".

Objective sentences are those which contain facts but such sentences do not convey any sentiments, e.g., "I purchased a Nokia phone 2 months before". It's an objective sentence having no opinion. However, there are few objective sentences with implied opinion in them. For instance, "This machine has stopped working suddenly" expressing negative polarity implicitly, though there are no sentiment bearing words in it. Next polarity classification is done to determine whether a sentence entails negative or positive connotations (Liu, 2012).

1.1.3 Aspect-level Sentiment Analysis

It's a finer level of analysis on individual aspect (feature or attribute of an entity) of each entity so, it has more possible usages. It is also known as entity or feature level in some earlier works (Hu & Liu, 2004; Pang & Lee, 2008). The concept of aspect extraction from opinionated data was first introduced by Hu and Liu (2004). Identifying sentiments expressed in a piece of text has varieties of applications such as tracking people's mood towards movies, products, politics, etc. This helps in improvement of customer relation model. In many of the above mentioned applications, it is important to determine the sentiment associated with an entity or features (aspect) of that entity because generally, customers express their opinions on different aspects of a product or services they have consumed. As an illustration, “Camera quality of my Nokia phone is awesome”. If aspect-level analysis is done on that piece of text then, it determines the positive sentiment on camera quality feature of Nokia phone[10]. It might be possible that overall sentiment towards a product or service is positive, but opposite sentiment towards the aspect or feature of that entity.

.

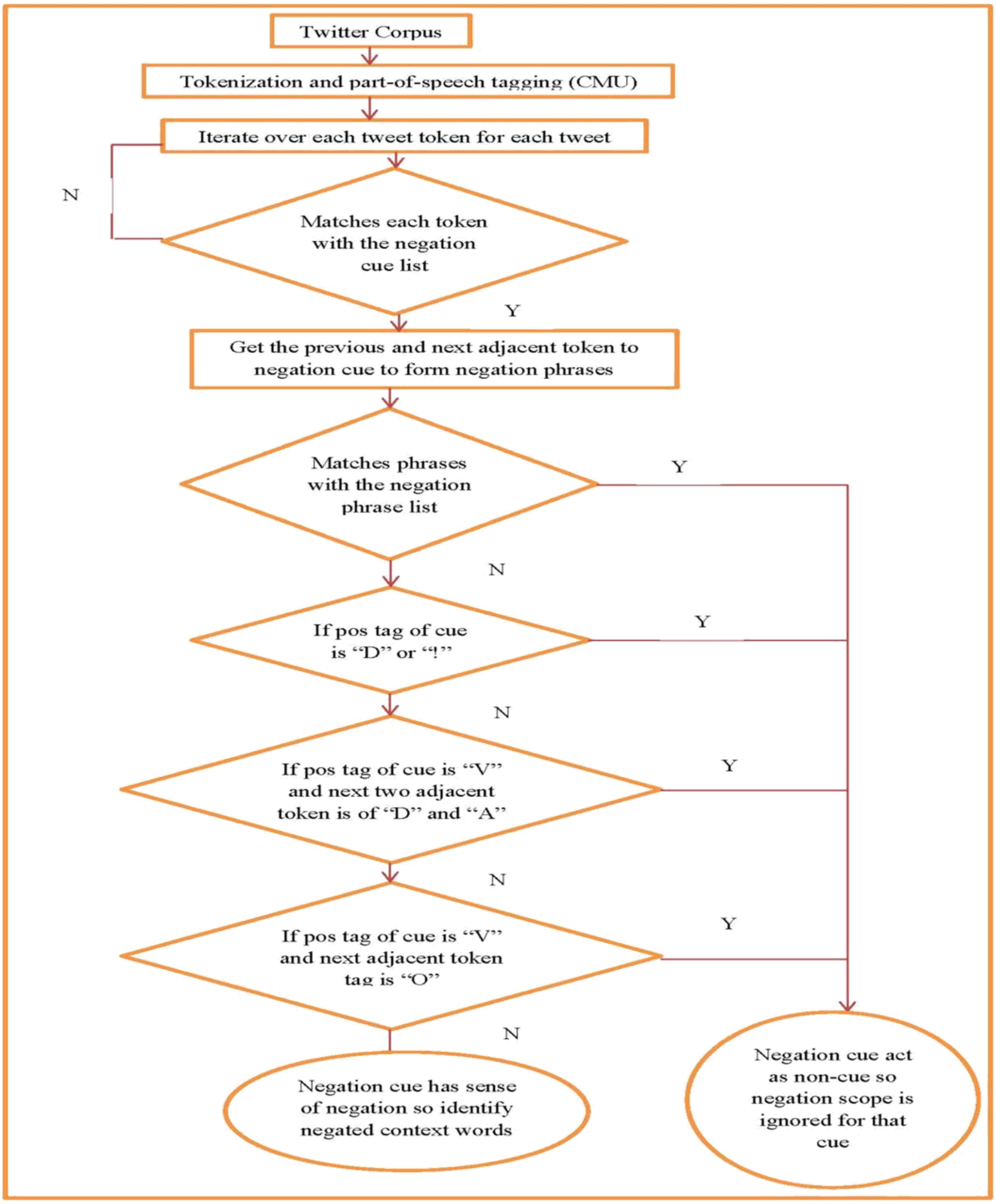


Figure 4.1: Procedure for handling negation exception cases

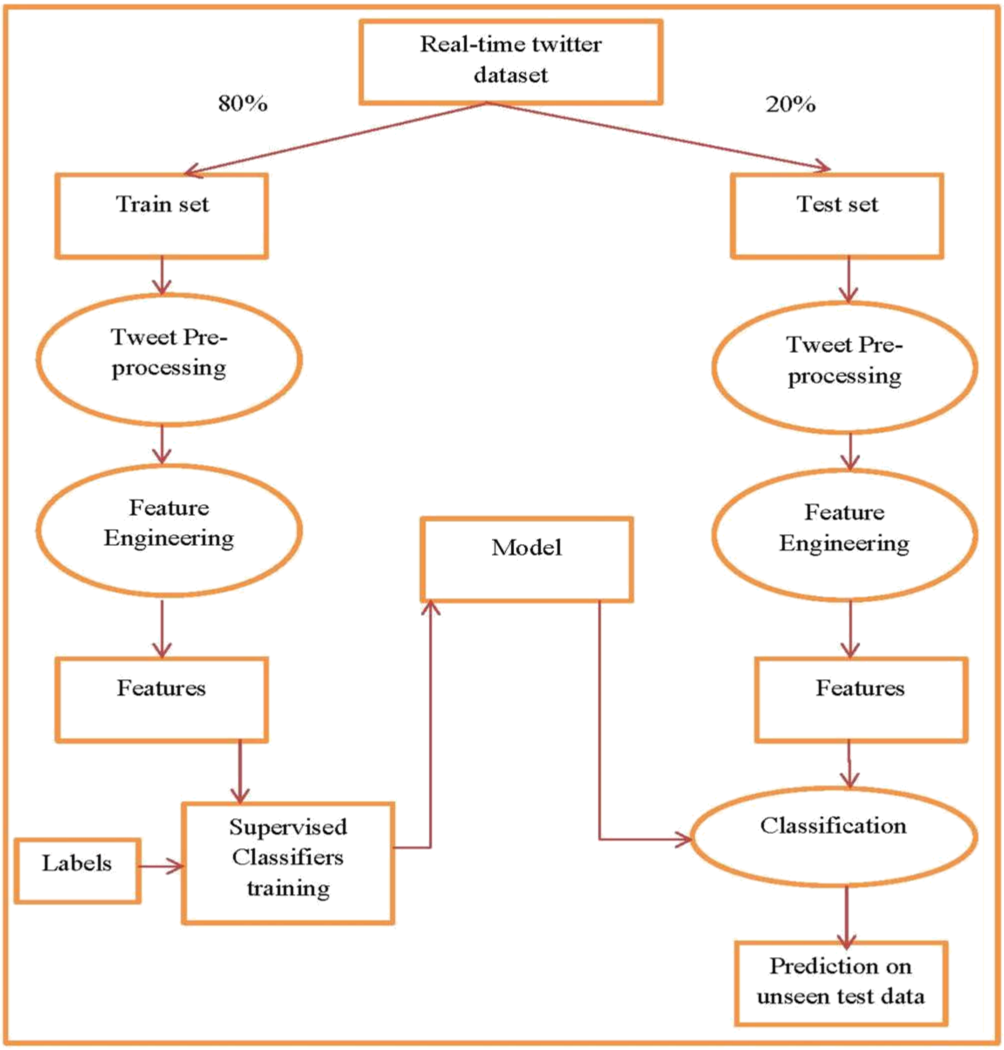
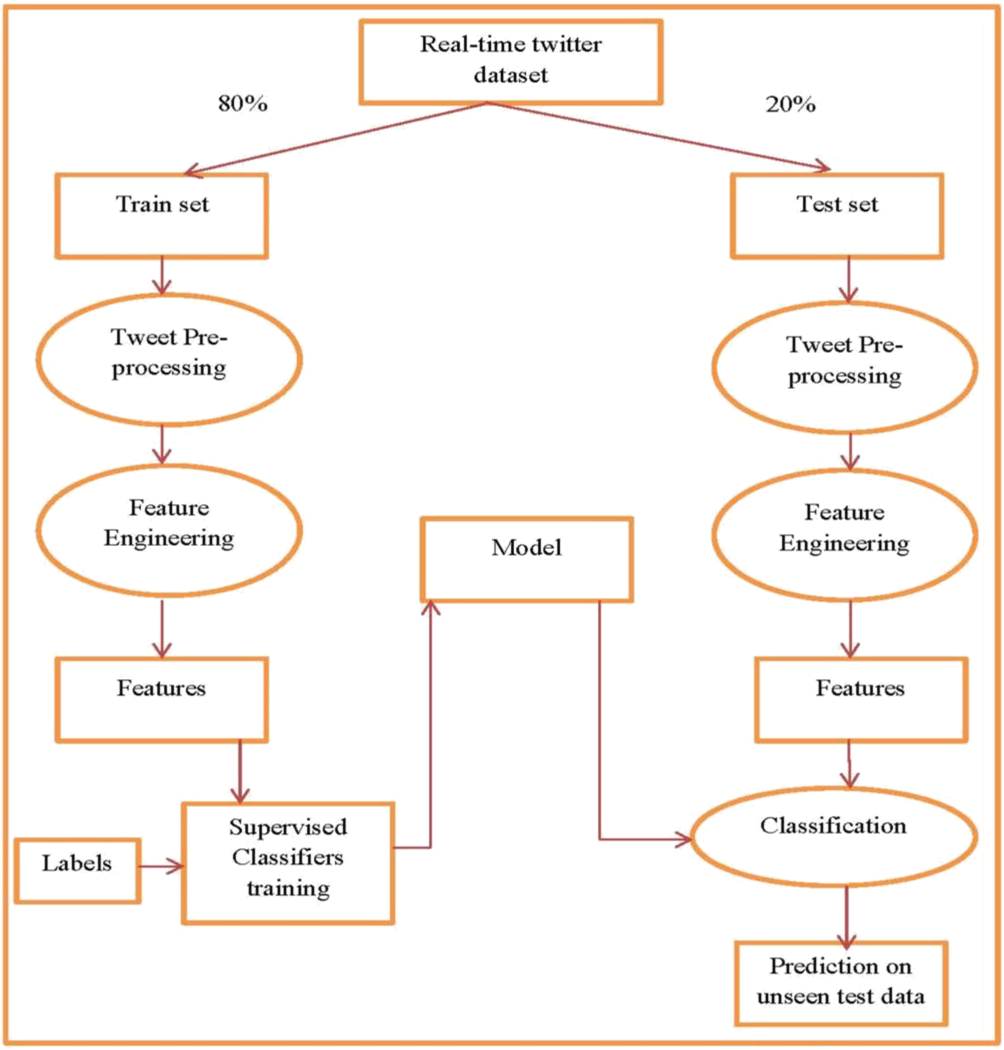


Figure 6.1: Proposed Supervised Learning Process

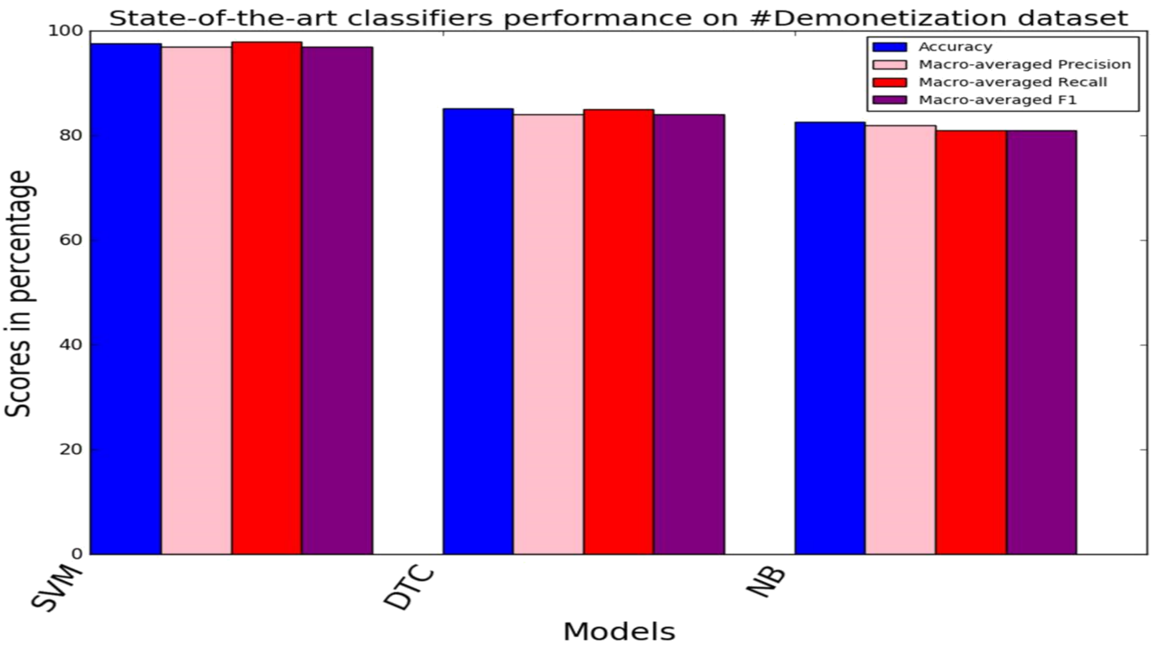
First of all, each real-time dataset was divided into training and test set by the usage of train_test_split ( ) function such that 80% of each dataset was used for training and rest 20% of dataset was used for testing. The reason for this splitting is to learn a model that will generalize on unseen data; otherwise model will mimic and gives 100% accuracy. It is very important to test the performance of learned model or trained model on unseen data (that was not seen during the training). Also, for benchmark dataset there is no need for splitting because we already have separate training and test tweets for it. Table 6.1 presents the statistics of training and testing tweets for each of the real-time dataset as well as benchmark dataset and figure 6.1



Classifier Evaluation Results on Real-time Twitter Datasets

As we have mentioned earlier, trained classifiers need to be evaluated on the test tweet for the performance evaluation. We recorded the trained classifiers performances on each of the real-time test tweets with several feature groups. We have presented several evaluation experiments for determining the significant performance of our feature-based twitter sentiment analysis system with incorporated knowledge-based tweet normalization system and negation handling procedure. Furthermore, we evaluated the performance in terms of accuracy, macro-averaged precision, recall, and F1 respectively. However, macro-averaged recall and macro-averaged F1 was considered as primary metrics for evaluation.

Dataset	System	Accuracy	Macro-avg. precision	Macro-avg. recall	Macro avg. F1
#Demonetization	a) SVM (without negation exception rules)	95.8	95	96	95
	a.1) SVM (without negation exception rules+ without double negation)	94.8	94	95	94
	b) SVM (our system)	97.6	97	98	97
	c) DTC (without negation exception rules)	84.3	83	84	83
	c.1) DTC (without negation exception rules+ without double negation)	83.3	82	82	82
	d) DTC (our model)	85.2	84	85	84
	e) NB(without negation exception rules)	82.0	81	80	81
	e.1) NB (without negation exception rules+ without double negation)	81.5	81	80	80
	f) NB (our model)	82.6	82	81	81



3. CONCLUSION

we have shown the process of training of SVM, classifiers on real-time as well as benchmark twitter dataset. Classifiers were trained on the 4 feature sets that we have extracted from the training corpus and evaluated on the unseen test corpus. We conducted series of ablation experiments to conclude which classifier works well with which feature group. We performed this set of experiment across all the datasets (real-time and benchmark both) and with all the three classifiers. Results showed that lexicon-based features to be the most influential across all the twitter datasets and classifiers. Further we discussed the impact of linguistic phenomenon negation on classifiers performance across all the twitter datasets that we have used in this work. Specifically, we analyzed the contribution of handling negation exception cases on our twitter sentiment analysis framework. We observed consistency in performance improvement with negation exception algorithm across all the datasets and classifiers. At last, we showed the impact of each pre-processing module of our tweet normalization system on classification performance of SVM model across each real-time as well benchmark dataset. We observed basic cleaning tasks and removal of stop words to be the most influential in performance gain of SVM across all the twitter datasets.

.

REFERENCES

[1] Agarwal, A., & Sabharwal, J. (2022, December). End-to-end sentiment analysis of Twitter data. In Proceedings of the Workshop on Information Extraction and Entity Analytics on Social Media Data (pp. 39-44).

[2] Agarwal, A., Xie, B., Vovsha, I., RamBOW, O., & Passonneau, R. J. (2011, June). Sentiment analysis of Twitter data. In Proceedings of the Workshop on Language in Social Media (LSM 2011) (pp. 30-38).

[3] Akhtar, M. S., Ekbal, A., Narayan, S., Singh, V., & Cambria, E. (2018). No, that never happened!! Investigating rumors on Twitter. IEEE Intelligent Systems, 33(5), 8-15.

[4] Alharbi, A. S. M., & de Doncker, E. (2019). Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information. Cognitive Systems Research, 54, 50-61.

[5] Alistair, K., & Diana, I. (2005). Sentiment classification of movie and product reviews using contextual valence shifters. In Proceedings of FINEXIN.

[6] Angiani, G., Ferrari, L., Fontanini, T., Fornacciari, P., Iotti, E., Magliani, F., & Manicardi, S. (2016). A Comparison between Preprocessing Techniques for Sentiment Analysis in Twitter. In KDWeb.

[7] Anjaria, M., & Guddeti, R. M. R. (2014). A novel sentiment analysis of social networks using supervised learning. Social Network Analysis and Mining, 4(1), 181.

[8] Appel, O., Chiclana, F., Carter, J., & Fujita, H. (2016). A hybrid approach to the sentiment analysis problem at the sentence-level. Knowledge-Based Systems, 108, 110-124.

[9] Araque, O., Corcuera-Platas, I., Sánchez-Rada, J. F., & Iglesias, C. A. (2017). Enhancing deep learning sentiment analysis with ensemble techniques in social applications. Expert Systems with Applications, 77, 236-246.

[10] Asghar, M. Z., Khan, A., Ahmad, S., Qasim, M., & Khan, I. A. (2017). Lexicon-enhanced sentiment analysis framework using rule-based classification scheme. PloS one, 12(2).

[11] Asghar, M. Z., Kundi, F. M., Ahmad, S., Khan, A., & Khan, F. (2018). T-SAF: Twitter sentiment analysis framework using a hybrid classification scheme. Expert Systems, 35(1), e12233.

[12] Asiaee T, A., Tepper, M., Banerjee, A., & Sapiro, G. (2012, October). If you are happy and you

know it... tweet. In Proceedings of the 21st ACM international conference on Information and knowledge management (pp. 1602-1606).

[13] Baccianella, S., Esuli, A., & Sebastiani, F. (2010, May). Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In Lrec (Vol. 10, No. 2010, pp. 2200-2204).

[14] Bakliwal, A., Foster, J., van der Puil, J., O'Brien, R., Tounsi, L., & Hughes, M. (2013, June). Sentiment analysis of political tweets: Towards an accurate classifier. Association for Computational Linguistics.