

Neuro Fuzzy Classification based Traffic Sentiment Research

A. Vijay Karthik, P. Sengottuvelan

Abstract--- Sentiment analysis has become the hot area of research in recent years. Less work is done in the field of transportation. In this paper, Term Frequency - Inverse Document Frequency (TF-IDF) is used for feature selection and Neuro Fuzzy Classification algorithm is designed for Traffic Sentiment Analysis (TSA). It is seen from the results that, the proposed Neuro Fuzzy Rule Mining (NFRM) algorithm yields better performance when compared to the Apriori Algorithm (ARMA) and Fuzzy Rule Mining Algorithm (FRMA) in terms of accuracy, precision, recall and time.

Keywords--- Traffic Sentiment Analysis (TSA), Neuro Fuzzy Rule Mining (NFRM), Apriori Algorithm (ARMA), Fuzzy Rule Mining Algorithm (FRMA)

I. INTRODUCTION

Opinion mining also known as sentiment analysis or emotion AI involves the usage of Natural Language Processing (NLP), text analysis, computational linguistics, and biometrics. It deals with the identification, extraction, quantification along with relevant states and information. Sentiment analysis finds its application in raising the voice of the customers through reviews, survey responses, online and social media, and healthcare materials for applications.

To be precise, Sentiment analysis is based on machine learning wherein machines analyse and classify the sentiments, emotions and opinions of humans that are expressed in the form of text or speech.

Modern Intelligent Transportation Systems (ITSs) do not focus on public opinions. Hence, it is essential to collect and analyse the public's opinion.

The steps in sentiment analysis include:

- **Data collection:** The data is collected from user generated blogs, forums and social networks. The data may not be organized in the right way. The data may be expressed in different ways involving diverse vocabularies, slangs and context of writing, thus making manual analysis impossible. Text analytics and NLP simplifies the process of extraction and classification.
- **Text preparation:** It deals with pre-processing of data and makes it ready for analysis. It involves removal of non-textual and irrelevant details from the extracted data.
- **Sentiment detection:** It deals with examining the extracted sentences of the reviews and opinions. Sentences with subjective expressions that include opinions, beliefs and views are retained. The sentences with objective communication that contain facts, factual information are cast-off.

- **Sentiment classification:** In classification, subjective sentences are classified as positive or negative, good or bad, like or dislike using multiple points.
- **Presentation of output:** It is a known fact that sentiment analysis deals with the conversion of unstructured text into meaningful information. The text results are graphically displayed and a sentiment time line with the chosen value over time is constructed for analysis.

1.1 Sentiment Classification Approaches

In sentiment classification, data in a document are categorized into favourable (positive) or unfavourable (negative) class.

Medhat et al (2014) have classified into 3 main levels:

- **Document level:** It considers the whole document a basic information unit and classifies an opinion document based on positive/ negative opinion or sentiment.
- **Sentence-level:** It deals with classifying the sentiment expressed in each sentence. If the sentence is a subjective one, it classifies it based on positive or negative opinions.
- **Aspect-level:** As users have different opinions for diverse aspects of the same entity, aspect-level classification deals with sentiments based on the specific aspects of entities.

1.2 Traffic Sentiment Analysis

Transportation systems serve the public, but the modern Intelligent Transportation Systems (ITSs) do not consider their opinions. To deal with this issue, it is essential to collect and analyse the data related to public knowledge and opinion. Traffic Sentiment Analysis (TSA) deals with processing of traffic information taken from websites. As human opinions are considered, TSA improves the performance of the current ITS space. To be more precise, TSA is a subfield of sentiment analysis, which deals with the issues of traffic. As stated by Pang & Lee (2008) have stated that due to the field sensitivity of sentiment analysis, it is essential to construct specific TSA systems. Hence, TSA looks into the traffic problems at a new angle, and adds to the abilities of the current ITS systems.

The functions of the TSA system include:

- **Investigation:** Collecting the public opinion through the TSA system is more economical and efficient than the public poll.

Manuscript received September 16, 2019.

A. Vijay Karthik, Muthurangam Government Arts College, Vellore, T.N, India.(e-mail: vijayv1r365@gmail.com)

P. Sengottuvelan, Department of Computer Science, PG Extension Center, Government Arts College Campus, Dharmapuri. T.N, India.

- **Evaluation:** The performance of traffic services and policies can be evaluated using the computational production of the TSA system.
- **Prediction:** The trends of some social events can be predicted using the extended TSA system.

1.3 Issues in TSA

The appropriate selection of sentiment analysis approaches is a challenge in TSA as they deal with Web based data. The performance of both the rule and learning based approaches are determined by the data to some extent, the features of web data should be identified first. Following are the properties of data.

- Texts with varying lengths.
- Diverse stylistic features of the texts
- Frequent updations of internet expressions

II. RELATED WORK

In this section, various methodologies proposed for sentiment analysis by various authors in recent years are discussed.

Sentiment analysis deals with mining affecting information from data and identifying the sentiment polarity in the information.

Hatzivassiloglou &McKeown (1997) have dealt with the analysis of the semantic orientation of adjectives. The sentiment polarities are extracted using expressions such as or ‘fast but inaccurate’ or ‘beautiful and smart’.

Various techniques of sentiment analysis are widely used in text filtering, tracking of public opinion, and customer relationship management (Liu et al (2005), Nasukawa& Yi (2003), Popescu&Etzioni (2007), Riloff et al (2006)).

Pang & Lee (2008) and Zhang et al (2009) have classified the former studies based on different standards.

Based on the study by Zhang et al (2009), the previous studies are discussed based on their level of granularity, type of analytical technique and based on language.

2.1 Level of Granularity

The issues in sentiment analysis at different levels of granularity, from the document level to the sentence level are discussed by various authors.

Pang et al. (2002) have adopted a standard bag-of-features framework to classify the sentiments of articles focussing on unigrams and bigrams of words. Document sentiment classification approach classifies movie reviews by using supervised machine learning method.

Turney et al (2002)have propounded Point-Wise Mutual Information and Information Retrieval (PMI-IR), an unsupervised learning algorithm that deals with the prediction of the semantic orientations of an article. It computes the similarity of phrases and classifies them as either ‘excellent’ or ‘poor’. The relationship between a polarity-unknown word and a set of manually selected seeds are considered for classifying into positive or negative class.

Tsou et al (2005) have considered the spread, density and intensity of polar lexical terms to increase the performance of sentiment classification.

Mullen & Collier (2004) have used semantic orientation of words defined by Turney et al (2002) and information from the Web and thesaurus. The accuracy is improved by

the semantic orientation of words and the lemmatized word unigram.

Kudo & Matsumoto (2004) have used sentence level classification approach by considering word dependency trees as features for sentence-wise sentiment polarity classification.

2.2 Type of Analytical Technique

Current methods of sentiment analysis can also be categorized into rule and learning based approaches.

Rule-based approaches

Rule-based approaches use an expert-definite dictionary containing subjective words. Bloom et al (2007) have predicted the polarity of a sentence or document by analysing the patterns of such words that occur in the text.

Wiebe et al (2004)have provided a dictionary of subjectivity clues including verbs, adjectives and nouns along with their polarity(positive, negative, or neutral) and strength (strong or weak). The drawback in this approach is that the original polarity of a word can be found. But the actual polarity of a word may be modified based on its context in a sentence.

Approaches that consider the context of words to decide the sentiment orientation of words are also proposed in the literature.

Yuen et al (2004) have designed a methodology wherein the semantic polarity of words is derived based on morphemes.

Knowledge sources like WordNet are also used to measure the semantic polarity of adjectives (Kamps et al 2004).

Learning-based approaches

Hu & Liu (2006) have designed an approach to extract features from product reviews based on class sequential rules which are linguistic patterns. They are mined from a set of labelled training sequences of words and part-of-speech tags.

Pang et al (2002) have applied three machine-learning methods to predict their sentiment on the reviews taken as a bag of unigram/bigram features. It is found that machine learning algorithms do not yield better perform when compared to traditional topic categorization tasks.

As stated by Turney et al (2002), these sentiment classification techniques demands manually labelled, suitably large training datasets containing positive and negative samples, which are often not cost and time effective.

2.3 Language

Most of the studies on sentiment analysis have focused on English and have shown a notable success in several applications.

In contrast, as stated by Che et al (2010), Chinese sentiment analysis is not greatly investigated as the linguistic characteristics of the Chinese language involve several technical challenges.

Word segmentation is required as an additional step in Chinese language processing as words in sentences are not segmented by space (Zeng et al 2008). Further, as the language contains various adverbs, the use of them can lead to subtlety and ambiguity in sentences.

Other approaches

Ku et al (2006) have proposed algorithms for opinion extraction, summarization and tracking. The opinion tracking system not only provides text and graph based opinion summaries, but also presents the trend of opinions from many information sources.

Popescu&Etzioni (2007) have introduced an unsupervised information extraction system named OPINE which mines reviews and builds a model which is a combination of significant product features, evaluation by reviewers and their relative quality across products. It uses relaxation labelling to find the opinion phrases and their polarity by finding the semantic orientation of words in context.

Kaur & Gupta (2013) have done a survey on approaches used for sentiment analysis which include:

- Subjective lexicon approach: It includes a list of words with a score that indicates its nature - positive, negative or objective.
- n-gram modelling approach: It uses unigram, bigram, trigram or combination of these for the sentiments classification.
- Machine learning approach: It performs semi and/or supervised learning by extracting features from the text and learning the model.

Maynard & Funk (2011) have suggested three types of techniques for Sentiment Classification, machine learning approach, lexicon-based approach and hybrid approach.

Machine learning approach predicts the polarity of sentiments based on trained as well as test datasets by applying ML algorithms and using linguistic features. The frequently used machine learning approaches include Bayesian Networks, Naive Bayes Classification, Maximum Entropy, Neural Networks and Support Vector Machine.

The commonly used lexicon-based approaches of sentiment classification include Dictionary based approach, Novel Machine Learning Approach, Corpus based approach and Ensemble Approaches.

Syed et al (2014) dealt with the lexicon based approach applied on a morphologically rich language: Urdu. In addition to the morphological structure of the words, it focuses on the sentence grammatical structures. Extraction and linking are done using shallow and dependency parsing.

Abdulla et al (2014) have designed a lexicon-based approach for Arabic sentiment analysis by building the lexicon and the sentiment analysis tool. It aids the researchers in their on-going efforts in improving the lexicon-based sentiment analysis.

III. EXISTING SYSTEM

Cao et al (2014) have proposed methods and models in TSA, and have highlighted the advantages and disadvantages of rule and learning based approaches for web data. They have applied the rule based approach to deal with real problems, and have presented an architectural design,

constructed related bases, demonstrated the process, and discussed the online data collection.

The authors have taken 2 cases to demonstrate the efficiency of their proposed approach - the 'yellow light rule' and 'fuel price' in China.

A rule-based approach is adopted to address the distinct challenges posed by the Chinese dataset. The architecture includes the following components: web data collection, pre-processing, extraction of subjects and objects, extraction of sentiment properties, sentiment calculation and classification, evaluation or applications and feedback, and improve the construction of the sentiment, rule and TSA object bases.

As Chinese documents do not segment words by spaces in sentences, they are to be processed. IT includes the following steps - Segmentation of text, Labelling of words, and Replacement of synonymous expressions.

The first two steps are done by a Chinese segmentation tool namely the Chinese Lexical Analysis System 3 launched by the Chinese Academy of Sciences, Beijing, China in 2011 (ICTCLAS).

Pre-processing is done based on the material and the demand of the algorithms to circumvent useless disturbances and improve precision (Zeng et al 2008). Hence, Cao et al (2014) have constructed 'sentiment base' in the application of TSA.

Subjects and objects are extracted using context mining and document analysis (Riloff et al 2005, Dave et al 2003). In context mining, models are to be designed based on the datasets and resources. In the second approach of extracting subjects and objects is text analysis, the opinion-oriented information are extracted through the pure text.

The properties are extracted based on sentiment, modifier, and rule bases. Three-step strategy proposed by Zhang et al (2009) is applied. The rule and object bases are to be updated with time since the topics and fashion terms discussed online change frequently.

As the regulation of a language is relatively static, the rule base is relatively consistent and hence updated semi-automatically. As the topics in the object base change quickly, the related topics and objects, as well as their attributions and components are to be summarized.

The approach is evaluated on a scientifically constructed standard dataset (Whitelaw et al 2005, Kobayashi et al 2005) before application. The efficiency and precision of the algorithms are tested. If the test performance is to be improved, proper words are to be identified and updated to improve the related bases.

Text sentiment calculations are categorized into three levels namely, word, sentence, and document levels. Computation of sentiment polarity of words is the first step in the construction of the sentiment word base.

Text processing involves two main parts, the polarity calculation of the sentence and document level text.

The method includes two major steps, i.e., the sentence sentiment analysis and document sentiment aggregation.

The document is decomposed into constituting sentences and the sentiment polarity of each sentence is determined. The overall polarity of the entire document is found by computing the polarity scores of all the sentences.

IV. PROPOSED SYSTEM

The proposed methodology uses Neuro Fuzzy Classification algorithm for sentence classification. It takes the twitter and facebook datasets.

The data is pre-processed to remove symbols and display the words in the dataset. It involves the following steps.

- Sentence Extraction&Segmentation - The sentence is extracted and the words in the dataset are segmented. It takes the whitespace as separator and segments the sentence.
- Sentiment words extraction-The words in the sentence are extracted. Important words which are required for sentiment analysis are taken. The stop words which are present in the sentence are removed and the words in sentence are filtered.
- Degree words extraction - The degree words are categorized into six intensity levels, and each degree word is assigned with a value according to its intensity level.
- Negative words extraction - The negative word lexicons are scanned.
- Sentence polarity calculation - The polarity of a sentence or document is predicted by analysing the occurring patterns of such words in the text.
- Sentence weight calculation - For weight calculation and classification, the existing KNN (k-Nearest Neighbour) and the proposed weight Neuro-fuzzy classifier algorithm is used. The weights of each sentence are calculated and the sentence is classified as positive or negative. For positive/negative classification, positive/negative words dataset is used. The words in the input sentence are compared with the positive/negative dataset and classification is done.
- Sentence Classification using Neuro-Fuzzy Classification algorithm - Neuro-fuzzy classification approach normally applies the concept of adaptive neural network.

Features

Unigram, bigram and N-gram features are used in feature representation (Pang et al (2002), Liu (2012), Mouthami et al (2013), Bhadane et al (2015)

Let,

$$doc = n_1(d), n_2(d), n_3(d) \dots n_m(d) \quad (1)$$

where

doc - document

f - features

$n_i(d)$ - Number of times the feature f_i occurs in the document

Selecting a single feature is termed as unigram feature. Feature hashing reduces the number of features, so that

original high-dimensional space is reduced to lower-dimensional space by mapping features to hash keys.

Feature Selection

Various feature selection methods like Term Frequency - Inverse Document Frequency (TF-IDF), Information Gain (IG), Mutual Information (MI), Feature Vector, Unigram, Bigram and N-gram methods are available in the literature.

Chi square method yields better results for both positive and negative classes.

Feature set is used to construct the feature vector.

TF-IDF score is used in balancing the most weighted and less weighted words (Tripathy et al 2015).

- Count Vector - It is the number of occurrences of the feature under consideration.
- Term Frequency - Inverse Document Frequency (TF-IDF) - It is the product of frequency of word under consideration (TF) and the frequency of the word in whole corpus (IDF) (Mouthami et al 2013).

$$W_i = f_{i,j} * \log\left(\frac{n}{n_i}\right) \quad (2)$$

W_i - Weight of a term 'i'

$f_{i,j}$ - Frequency of term 'i' in sample 'j'

n - Total number of samples in the corpus

n_i - Number of samples containing term 'i'

- **Information Gain** - Information gain is the used extensively in attribute selection in the area of sentiment analysis. It is used to determine the appropriate features to predict review by dealing with the presence/absence of a feature in the document.

$$IG(f, c) = -\sum_{c \in C} P(c) \log P(c) + \sum_{f \in F} P(f) \cdot \sum_{c \in C} P(c|f) \log P(c|f) \quad (3)$$

where,

$P(c|f)$ - Joint Probability

c - Class

f - Feature

$P(c)$ - Marginal probability

- **Mutual Information** - Mutual Information deals with the selection of features that are not uniformly distributed across the sentiment classes as they yield the necessary information of their classes. Mutual Information gives much priority to only few terms.

$$MI(f, c) = \sum_{c \in C} \sum_f P(f, c) \log \frac{P(f,c)}{P(f) \cdot P(c)} \quad (4)$$

where,

$P(f,c)$ - Joint Probability Distribution function

$P(f)$ - Marginal Probability Distribution of 'f'

$P(c)$ - Marginal Probability Distribution of 'c'

c - Positive and negative classes

- **Chi-square** - Chi-square measures observe the count and the expected count and deals with analysing the deviations between them.

$$\lambda^2(f, c) = \frac{N(AD-CB)^2}{(A+C)(B+D)(A+B)(C+D)} \quad (5)$$

Where,

A, B, C, D - Frequencies that represent the presence or



absence of feature in the sample

A - Count of samples in which feature 'f' and 'c' occurred together

$$N=A+B+C+D$$

V. PERFORMANCE ANALYSIS RESULTS

The screenshots show how the sentiment analysis is done using Neuro Fuzzy Classification algorithm (Figure to Figure 9).

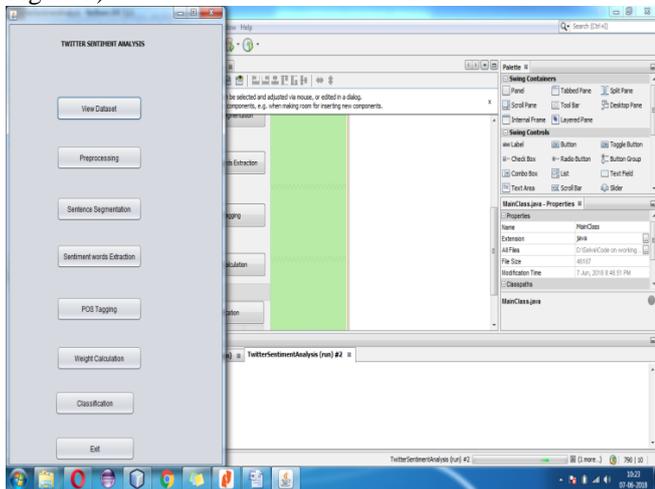


Figure 1: Main Screen

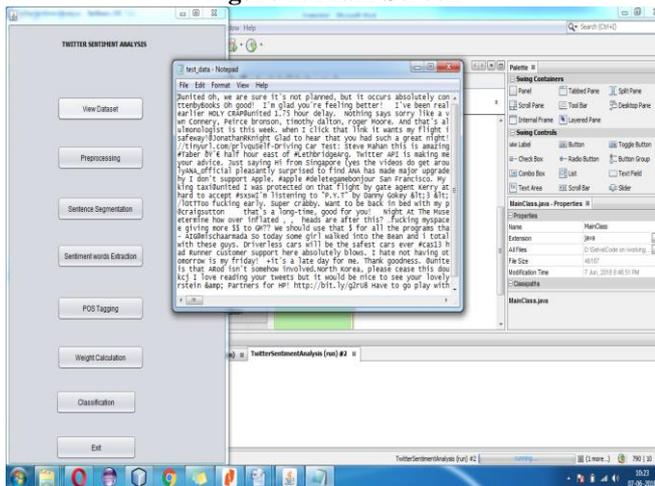


Figure 2: View Dataset

In Pre-processing phase, the symbols are removed and the words in the dataset are displayed (Figure 3).

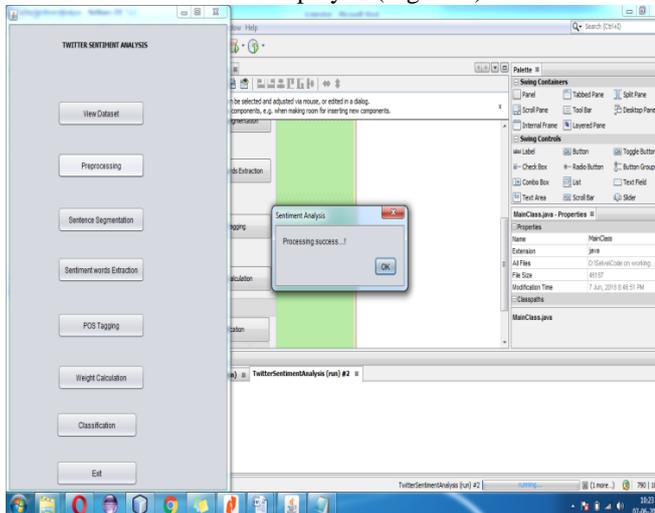


Figure 3: Pre-processing

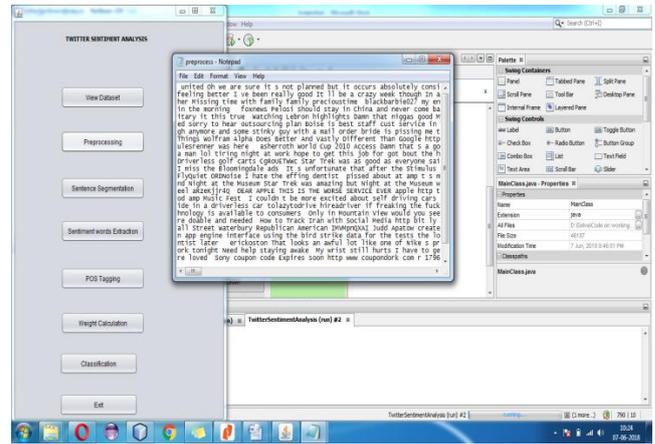


Figure 4: Output of Pre-processing

Sentence Segmentation

The sentences are extracted and the words in the dataset are segmented. The whitespace is taken as separator and the sentence is segmented.

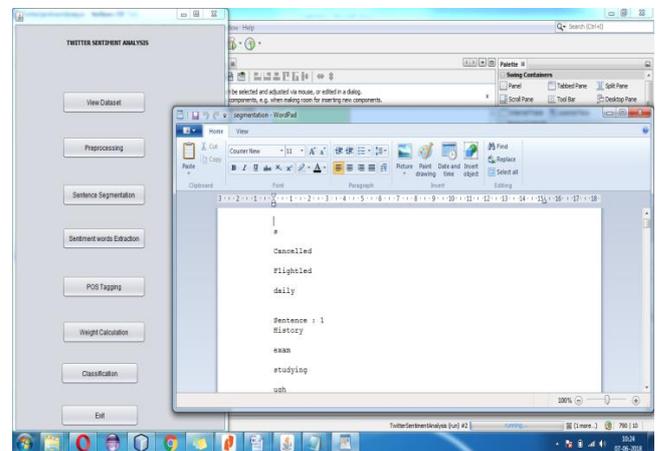


Figure 5: Sentence Segmentation

Sentence words Extraction

The words in the sentence are extracted. The important words which are required for sentiment analysis are taken. The stop words which are present in the sentence are removed and the words in the sentence are filtered.

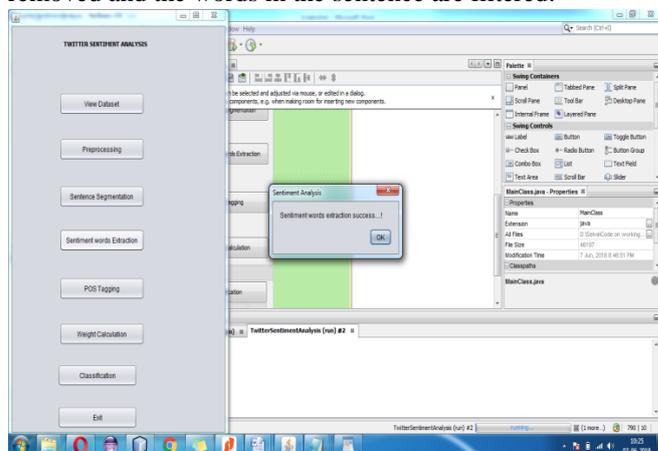


Figure 6: Word Extraction

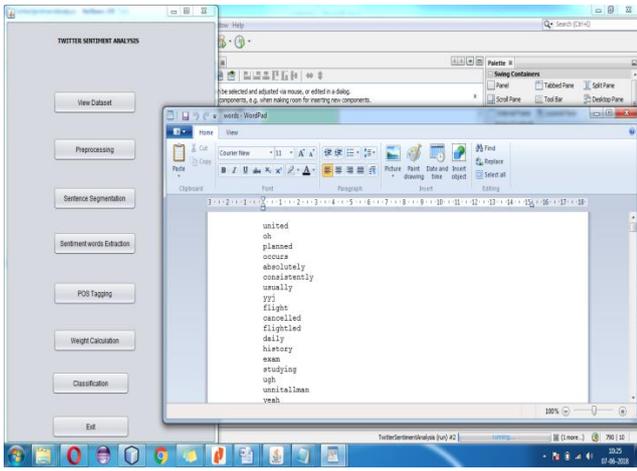


Figure 7: Word Filtering

POS Tagging

POS tagging means finding out the verb, noun and adjective for each words in the sentence. Taggers dictionary dataset is used for finding verb, noun and adjective.

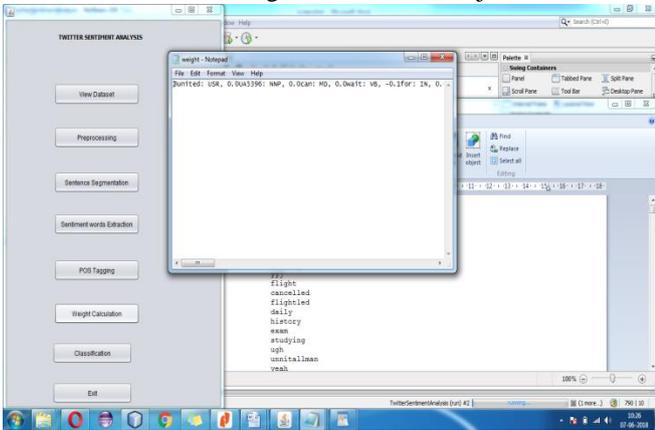


Figure 8: POS Tagging

Weight Calculation and Classification

For weight calculation and classification, K-Nearest Neighbour (KNN) and proposed Weighted Neuro Fuzzy Classifier algorithm is used. The weight of each sentence is classified as positive or negative.

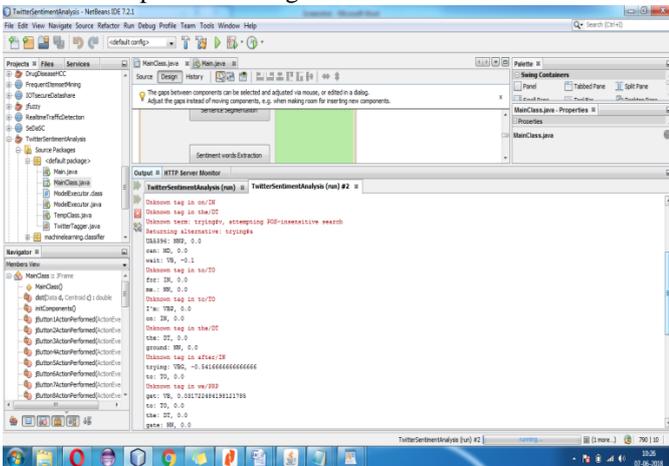


Figure 9: Weight Calculation & Classification

For positive negative classification, positive-negative words dataset is used. The words in the input sentence are compared with the positive negative dataset and classification is done to obtain the result of sentiment analysis.

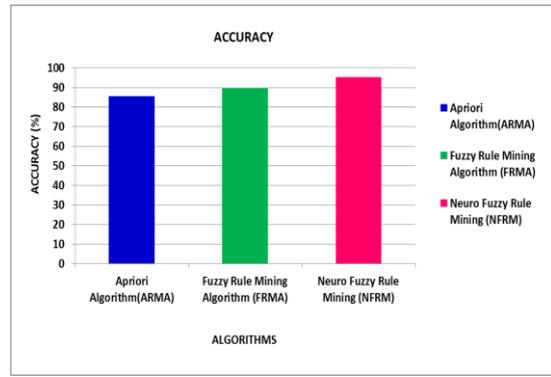


Figure 10: Accuracy

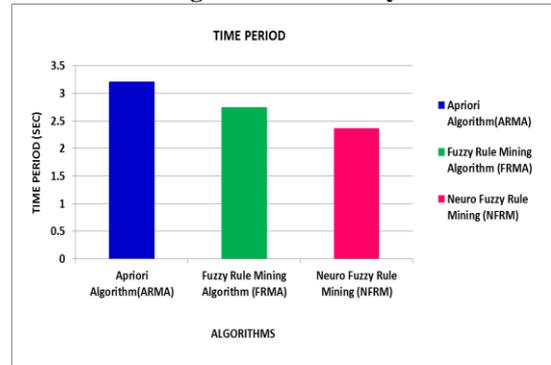


Figure 11: Time Period

The performance in terms of Accuracy, Time period, Precision and Recall are shown below (Fig. 10 to Fig. 13). Neuro Fuzzy Rule Mining (NFRM) algorithm outperforms Apriori Algorithm (ARMA) and Fuzzy Rule Mining Algorithm (FRMA). FRMA yields better accuracy, precision and recall and involves the least time.

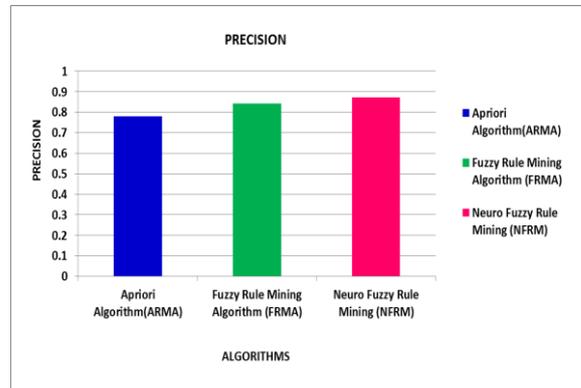


Figure 12: Precision

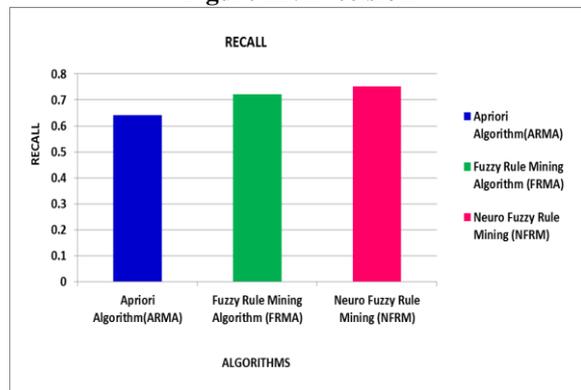


Figure 13: Recall



VI. CONCLUSION

In this paper, features are selected using Term Frequency - Inverse Document Frequency (TF-IDF). Neuro Fuzzy Classification algorithm is proposed for Traffic Sentiment Analysis (TSA). The proposed Neuro Fuzzy Rule Mining (NFRM) algorithm outdoes the Apriori Algorithm (ARMA) and Fuzzy Rule Mining Algorithm (FRMA) in terms of accuracy, precision, recall and time.

REFERENCES

1. Medhat, W., Hassan, A., Korashy, H. 2014. Sentiment analysis algorithms and applications: A survey, Ain Shams Eng.
2. B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Retrieval*, vol. 2, no. 1/2, pp. 1–135, Jan. 2008.
3. V. Hatzivassiloglou and K. R. McKeown, "Predicting the semantic orientation of adjectives," in *Proc. 8th Conf. Eur. Chapter Assoc. Comput. Linguist.*, 1997, pp. 174–181.
4. B. Liu, M. Hu, and J. Cheng, "Opinion observer: Analyzing and comparing opinions on the Web," in *Proc. 14th Int. Conf. World Wide Web*, 2005, pp. 342–351.
5. T. Nasukawa and J. Yi, "Sentiment analysis: Capturing favorability using natural language processing," in *Proc. 2nd Int. Conf. Knowl. Capture*, 2003, pp. 70–77.
6. A.-M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in *Natural Language Processing and Text Mining*. New York, NY, USA: Springer-Verlag, 2007, pp. 9–28.
7. E. Riloff, S. Patwardhan, and J. Wiebe, "Feature subsumption for opinion analysis," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2006, pp. 440–448.
8. B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Retrieval*, vol. 2, no. 1/2, pp. 1–135, Jan. 2008.
9. C. L. Zhang, D. Zeng, J. X. Li, F. Y. Wang, and W. L. Zuo, "Sentiment analysis of Chinese documents: From sentence to document level," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 60, no. 12, pp. 2474–2487, Dec. 2009.
10. B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: Sentiment classification using machine learning techniques," in *Proc. ACL Conf. Empirical Methods Natural Lang. Process.*, 2002, vol. 10, pp. 79–86.
11. P. D. Turney, "Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews," in *Proc. 40th Annu. Meet. Assoc. Comput. Linguist.*, 2002, pp. 417–424.
12. B. K. Tsou, R. W. Yuen, O. Y. Kwong, T. La, and W. L. Wong, "Polarity classification of celebrity coverage in the Chinese press," in *Proc. Int. Conf. Intell. Anal.*, 2005, pp. 137–142.
13. K. Bloom, N. Garg, and S. Argamon, "Extracting appraisal expressions," in *Proc. HLT-NAACL*, 2007, pp. 308–315.
14. J. Wiebe, T. Wilson, R. Bruce, M. Bell, and M. Martin, "Learning subjective language," *Comput. Linguist.*, vol. 30, no. 3, pp. 277–308, Sep. 2004.
15. R. W. Yuen, T. Y. Chan, T. B. Lai, O. Kwong, and B. K. Tsou, "Morpheme-based derivation of bipolar semantic orientation of Chinese words," in *Proc. 20th Int. Conf. Comput. Linguist.*, 2004, pp. 1008–1014.
16. J. Kamps, M. Marx, R. J. Mokken, and M. De Rijke, "Using WordNet to measure semantic orientations of adjectives," in *Proc. Int. Conf. Lang. Resourc. Eval.*, 2004, pp. 1115–1118.
17. M. Hu and B. Liu, "Opinion feature extraction using class sequential rules," presented at the AAAI Spring

- Symposium Computational Approaches Analyzing Weblogs, Palo Alto, CA, USA, 2006, Paper AAAI-CAAW-06.
18. W. Che, Z. Li, and T. Liu, "LTP: A Chinese language technology platform," in *Proc. 23rd Int. Conf. Comput. Linguist., Demo.*, 2010, pp. 13–16.
19. D. Zeng, D. Wei, M. Chau, and F. Wang, "Chinese word segmentation for terrorism-related contents," in *Intelligence and Security Informatics*. New York, NY, USA: Springer-Verlag, 2008, pp. 1–13.
20. L.-W. Ku, Y.-T. Liang, and H.-H. Chen, "Opinion extraction, summarization and tracking in news and blog corpora," in *Proc. AAAI Spring Symp., Com-put. Approaches Anal. Weblogs*, pp. 100 107, 2006.
21. A.-M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in *Natural Language Processing and Text Mining*. New York, NY, USA: Springer-Verlag, 2007, pp. 9–28.
22. Mullen, T., & Collier, N. 2004. Sentiment Analysis using Support Vector Machines with Diverse Information Sources. *Proc. of 9th EMNLP*, pp. 412-418.
23. Kudo, T., & Matsumoto, Y. 2004. A Boosting Algorithm for Classification of Semi-Structured Text. In *EMNLP*, Vol. 4, pp. 301-308.
24. Kaur, A., & Gupta, V. 2013. A survey on sentiment analysis and opinion mining techniques. *Journal of Emerging Technologies in Web Intelligence*, 5(4), 367-371.
25. Maynard, D., & Funk, A. 2011. Automatic detection of political opinions in tweets. In: *Proceedings of the 8th international conference on the semantic web, ESWC'11*, p. 88-99.
26. Syed, A.Z., Aslam, M., Martinez-Enriquez, A.M. 2014. Associating targets with SentiUnits: a step forward in sentiment analysis of Urdu text. *Artificial Intelligence Review*, 41(4), pp. 535-561.
27. Abdulla, N. A., Ahmed, N. A., Shehab, M. A., Al-Ayyoub, M., Al-Kabi, M. N., & Al-rifai, S. 2014. Towards improving the lexicon-based approach for arabic sentiment analysis. *International Journal of Information Technology and Web Engineering (IJITWE)*, 9(3), 55-71.
28. Jianping Cao, Ke Zeng, Hui Wang, Jiajun Cheng, Fengcai Qiao, Ding Wen, and Yanqing Gao, *Web-Based Traffic Sentiment Analysis: Methods and Applications* IEEE Transactions on Intelligent Transportation Systems, Vol. 15, No. 2, April 2014.
29. "ICTCLAS," 2011. [Online]. Available: <http://ictclas.nlp.ir.org/>
30. D. Zeng, D. Wei, M. Chau, and F. Wang, "Chinese word segmentation for terrorism-related contents," in *Intelligence and Security Informatics*. New York, NY, USA: Springer-Verlag, 2008, pp. 1–13.
31. E. Riloff, J. Wiebe, and W. Phillips, "Exploiting subjectivity classification to improve information extraction," in *Proc. Nat. Conf. Artif. Intell.*, 2005, pp. 1106–1111.
32. K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," in *Proc. 12th Int. Conf. World Wide Web*, 2003, pp. 519–528.
33. C. Whitelaw, N. Garg, and S. Argamon, "Using appraisal groups for sentiment analysis," in *Proc. 14th ACM Int. Conf. Inf. Knowl. Manage.*, 2005, pp. 625–631.
34. N. Kobayashi, K. Inui, Y. Matsumoto, K. Tateishi, and T. Fukushima, "Collecting evaluative expressions for

- opinion extraction,” in Natural Language Processing - IJCNLP 2004, K. Y. Su, J. Tsujii, J. H. Lee, and O.Y.Kwong, Eds. Berlin,Germany: Springer-Verlag, 2005, pp. 596–605.
35. Liu, B., 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1), pp.1-167.
 36. Pang, B., Lee, L. and Vaithyanathan, S., 2002, July. Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10* (pp. 79-86). Association for Computational Linguistics.
 37. Mouthami, K., Devi, K.N. and Bhaskaran, V.M., 2013, February. Sentiment analysis and classification based on textual reviews. In *Information Communication and Embedded Systems (ICICES), 2013 International Conference on* (pp. 271-276). IEEE.
 38. Bhadane, C., Dalal, H. and Doshi, H., 2015. Sentiment analysis: Measuring opinions. *Procedia Computer Science*, 45, pp.808-814.