IJCSN - International Journal of Computer Science and Network, Volume 7, Issue 6, December 2018
ISSN (Online) : 2277-5420
www.IJCSN.org
Impact Factor: 1.5

Developing POS Level Emotion-Based Features on Satire Detection

[1] Pyae Phyo Thu; [2] Than Nwe Aung

[1] University of Computer Studies, Mandalay
Mandalay, Myanmar

[2] University of Computer Studies, Mandalay
Mandalay, Myanmar

Abstract - Due to the implicit traits embedded in the language, handling figurative languages appear to be the most trending topics in public opinion mining and social multimedia sentiment analysis. Failures in recognition of these languages can lead to the misrepresentation of actual sentiments, attitudes or opinions person or community try to expose. Satire is a more alive form of figurative communication which intends to criticize someone's behavior and ridicule it. This work proposes the POS level emotion-based features by using the emotion lexicon SenticNet and VADER. It is approached as a classification problem by applying a supervised machine learning algorithm: Random Forest. The system can tackle the problem of high bias error in both long text and short text datasets with 83% to 89% accuracy whereas the BOW gives high accuracy but cannot handle the problem of high bias error in satirical language processing.

Keywords – *Satire Detection, POS Level Features, Emotion-Based Features, SenticNet, VADER.*

1. Introduction

In order to stimulate individuals to become an active reader, people use implicit attributes while expressing their ideas and feelings. Unlike literal language, figurative language, also known as implicit language, is any figure of speech that uses the word deviate from its actual intentions. Satire is a more alive form of figurative communication. It can be found in various outlets namely literature, television, the internet, social media, comics, and cartoons. The major role of satire is to criticize someone's behavior and ridicule it.

Satire loses its meaning when the public do not understand the real intentions hidden in its ironic dimension. It can flip the polarity of the desired attitudes or opinions person or community try to expose [9]. That is why the ability to detect such ironic language as satire is of great importance in natural language processing (NLP), computational linguistics, and intelligent text processing.

Many researchers have analyzed satirical language from lexical, grammatical and psycholinguistic perspectives. However, according to the implicit nature of satirical figurative language and ambiguous nature of emotion embedded in the language, no other works have been done from the emotional point of view. So it stays remain as the challenging task to detect satire from its emotion.

The major goal of the proposed work is to provide the solution to the emotion-based satirical figurative language processing. This work is implemented as the following objectives:

- To identify the emotion-based features that can tackle the problem of the ambiguous nature of emotion in satire detection
- To observe the features that achieve not only in short text but also in long text

The rest of the paper is organized as follows. Section II discusses the existing works on satirical figurative language processing. Section III discusses the data collection, feature extraction and selection of the proposed system. The experiments and outcome of the proposed system are described in section IV and concludes it in section V.

2. Related Work

Satire detection has attracted scholarly minds for a long time. There are many studies which observe its origins, functions, motivations, forms, plots and typical devices used in satirical language. The phenomenon of literary satire has been thoroughly examined, but the computerized approach to satire detection hardly applies. Pioneer work which attempted to model satire was published in [7]. This approach involved querying the Web using standard text classification functions - binary feature weighting and bi-normal feature scaling, and lexical features - profanity and slang. They pointed out that a combination of SVM and BNS scaling provides high accuracy and low recall, achieving a high F score with validity feature.

A set of language-independent features such as frequency of the rare word, ambiguity, Part-of-Speech, synonyms, sentiments, and characters were proposed in [12]. Due to the lack of vocabulary and semantics forms, these frequency-based features outperformed the word baseline by recognizing the existence of satire in Spanish, English, and Italian.

Lexical, grammatical and semantic features of Amazon's satirical and non-satirical reviews were proposed in [5]. The corpus contained 700 Amazon reviews of which 375 are satirical reviews. Using a one-way multivariate analysis (MANOVA), various set of linguistic indices were examined. The results indicated that the lexical sophistication is an important predictor of satire detection in product reviews. The model achieved 71% accuracy in leave-one-out cross-validation (LOOCV).

In 2017, a set of unigrams, semantic, psycholinguistic, and statistical features were proposed by using ensemble feature selection methods in order to detect satirical, sarcastic and ironic content in news and customer reviews [3]. They reinforced the belief that satire and irony share common psycholinguistic characteristics such as correctness, meaningfulness, imagery, familiarity and proved that satire is iron subtypes across several experiments.

In addition, some researchers stated that satirical cues are often reflected in paragraphs and not throughout the document [2]. They proposed linguistic features at the paragraph level that reveal satire by incorporating a neural network and an attention mechanism. Data analysis with paragraph-level language features gave the scores of 90.56 F1, which is about 1% more than document level language features. The combination of these two characteristics leads to the most promising results and the score of 91.11 F1 which is about 5% to 8% higher than that of the prior art.

In this way, the best of the best clues has proposed for satirical analysis. However, most of the satirical language processing were focused on statistical analysis and lexical information which can only be applied to long text datasets but does not work well in short text datasets. Later, although many features for satire analysis that could work well with both short sentences and long sentences have proposed, no research has been conducted on emotional perspectives. Some researchers also noted that satire recognition with emotion is difficult because of its ironic dimension [12].

3. Emotion-Based Satire Detection

In our previous works on emotion-based satire detection [4, 14, 15, 16], the system mainly focused on the document level emotion. Hence, it does not perform well in recognition of satirical data. In this work, POS level emotion-based features are proposed in order to tackle the problem of ambiguous nature of emotion.

3.1 Data Collection and Preprocessing

As shown in Fig. 1, the first part of the proposed work is data collection and text preprocessing. The data in this process has two forms, a short text and a long text. Since Twitter is a very popular microblogging service, users can use short messages up to 280 characters long and many users post their daily activities via Twitter. Thus, it becomes the main source of social media analysis. In this work, these tweets were collected from the list of satirical and non-satirical twitter accounts used in [12]. After cleaning and preprocessing the tweets by using tweet pre-processing tool TwitPre [13], about 12,000 tweets were collected as shown in Table 1.

In addition to the short text twitter dataset, long text dataset namely NR dataset is obtained from the publicly available corpus proposed in [6]. In this dataset, about 1,705 real news articles and 1,706 satire news articles were randomly collected by scraping various news websites: Reuters, CNET, CBS News, Daily Currant, Daily Mash, Dandy Goat, Empire News, The Lapine, Lightly Braised Turnip, Mouthfrog, News Biscuit, News Thump, Satire Wire, National Report and The Spoof. The details of the datasets used in proposed emotion-based satire detection can be seen as shown in Table 1.

Table 1: Datasets used in Proposed Emotion-Based Satire Detection

Dataset	***Satire***	***NonSatire***	***Total***
News Articles (NR)	1,706	1,705	3,411
Twitter Dataset (TAB)	5,947	6,059	1,2006

IJCSN - International Journal of Computer Science and Network, Volume 7, Issue 1, January 2018
ISSN (Online) : 2277-5420
www.IJCSN.org
Impact Factor: 1.5

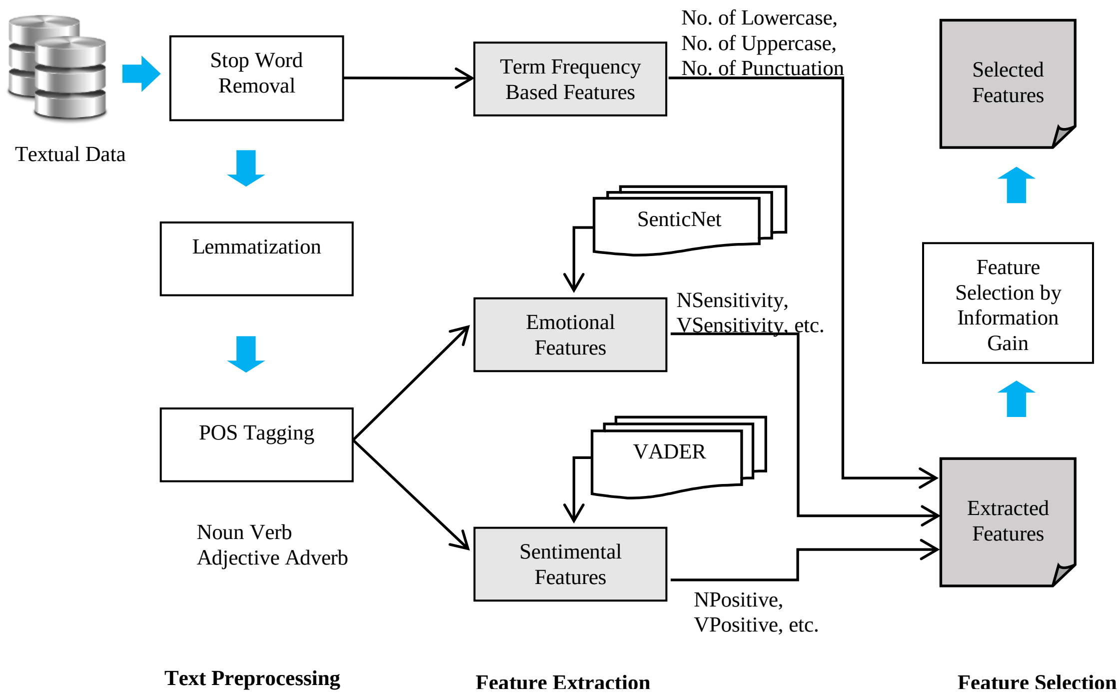

Fig. 1 Architecture of the Proposed Emotion-Based Feature Extraction.

3.2 Features Extraction

This study proposes the POS level emotion-based features towards the recognition of satirical figurative language in natural language processing. The former: BOW is used as the baseline and the latter: Emotional, Sentimental, and Term Frequency-Based Features are used in the proposed system.

- Bag-of-Word (BOW) features - The bag-of-words model (BOW) is a simplified word-based feature extraction model used for natural language processing and information retrieval (IR). In this work, the lemmatized words bigrams skips 1 gram, 2 grams and 3 grams are used as text classification baseline. The highest 1,000 most frequent word occurrences are used in this work and the rest are filtered out.

- Emotional features – After lemmatization and POS tagging using NLTK Word Net POS tagger and lemmatizer, the words tagged with four POS level: Noun, Verb, Adjective, and Adverb are extracted. It is motivated by the work in [1]. For each POS level word, four basic emotions of SenticNet lexicon [11] sensitivity, attention, pleasantness, and aptitude are extracted. Then, the total emotional values of every word in each level are used as the features. In total, 16 emotional features are used. For example, after lemmatization and POS tagging the satirical tweets "Stationary motorway traffic to be charged parking fees", the Noun words *fee, motorway, traffic*, the Verb words *charged*, *parking*, the Adjective words *stationary* are extracted. There is no Adverb word. The emotional values of each word are extracted by using the emotion lexicon SenticNet. As a result, the emotional values of Noun words *fee*, *motorway*, *and traffic* are fee: {'sensitivity': '-0.84', 'attention': '0', 'pleasantness': '-0.83', 'aptitude': '0', motorway: {'sensitivity': '0', 'attention': '-0.98', 'pleasantness': '0.559', 'aptitude': '0.501', and traffic: {'sensitivity': '0.989', 'attention': '0', 'pleasantness': '-0.98', 'aptitude': '-0.93'}. After summation of corresponding emotional values, the emotional features of Noun level are NSensitivity (0.149), NAttention (-0.98), NPleasantness (-1.251), and NAptitude (-0.429). In this way, the emotional features of remaining POS levels such as VSensitivity, VAttention, VPleasantness, VAptitude, AdjSensitivity, AdjAttention, AdjPleasantness, AdjAptitude, AdvSensitivity, AdvAttention, AdvPleasantness, and AdvAptitude are computed. In order to compare the performance of the proposed POS level features, the document level features are also computed by summation of corresponding emotional values of every word in the text. For example, the document level emotional features of previous tweet is DocSensitivity (0.301), DocAttention (-1.63), DocPleasantness (-0.59) and DocAptitude (0.274).

- Sentimental features – Similar to the emotional features, 16 sentimental features are extracted by using sentiment lexicon VADER [10]. For example, for each POS level word, four sentimental values of VADER lexicon Positive, Negative, Neutral, and Compound are extracted. For Noun level words *fee, motorway, and traffic*, the extracted sentimental values are as follows. Fee: {'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0, motorway: {'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0, traffic: {'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0}. After summation of corresponding values, the resulted Noun level features are NNeg (0.0), NNeu (3.0), NPos (0.0) and NCompound (0.0). In this way, the remaining sentimental features V - Neg, Neu, Pos, Compound, Adj - Neg, Neu, Pos, Compound, and Adv - Neg, Neu, Pos, Compound are computed.

- Term Frequency-Based features – Inspired by the work in [8], term frequency-based features which measure the number of lowercase characters, the number of uppercase characters, and the number of punctuation words (i.e. ! " # $ % & ' ( ) * + , - . / : ; < = > ? @ [\ ] ^ _ ` { | } ~) are extracted. Punctuation words are used because it is very important in social networks to amplify the desired emotions. For example, an ellipse may be a sign of a satire or a full stop of a negative emotion. In total, there are 34 features in this term frequency-based feature sets.

3.3 Features Selection

Feature selection also called attribute selection or variable selection is the process of selecting a subset of most relevant features or variables for model building. There are four main reasons for doing this:

- To simplify the model by reducing the number of parameters
- To reduce training time
- To improve generalization and reduce over fitting
- To avoid curses of dimensionality

The central principle of using feature selection is that the data which contains duplicate or irrelevant features can be removed without incurring much loss of information. It is often used in areas with many features and relatively few samples or data points. In order to solve the problem of the ambiguous nature of emotion by selecting the most relevant features of the system, the proposed work used Information Gain (IG) to analyze the feature importance ranking and to select the relevant features set. The top 33 features are selected as the features for the proposed emotion-based detection (see Appendix 1).

3.4 Classification

The problem of emotion-based satirical language processing is approached as a classification problem by using supervised machine learning classifier Random Forest. It was proposed in [17] by ensembling the decision trees. It fits multiple decision trees on data by selecting N samples randomly, with replacement, from the training dataset. In order to fit multiple trees, random forest randomly selects F input features to split at each node of the decision tree. The value F is generally defined as log2(|D|) + 1, where |D| is the total number of samples. The small value of F yields less correlation among decision trees whereas the bigger value of F produces better classification accuracy. Then, the results of several decision trees are combined using bagging.

4. Experimental Results and Discussion

In this section, the performance of the proposed emotion-based satire detection is analyzed into two parts. The former shows the importance of proposed POS level emotion-based features whereas the latter compares the performance of the proposed features with word-based features (i.e. BOW model).

4.1 Feature Importance Analysis

In order to fully understand the contribution of each proposed feature, the information gain of each feature has been studied. Fig. 2 describes the scaled importance values of the five feature sets by averaging the importance of individual features on NR dataset. It is clear that the importance of features at the POS level is greater than that of document level features in both emotional and sentimental features. POS level emotional features and sentimental features are the second and third highest among the five feature sets whereas the term frequency-based features have the highest ranking among the five feature sets. On the other hand, emotional features and sentimental features at the document level are less important than the other three features.

IJCSN - International Journal of Computer Science and Network, Volume 7, Issue 6, December 2018
ISSN (Online) : 2277-5420
www.IJCSN.org
Impact Factor: 1.5

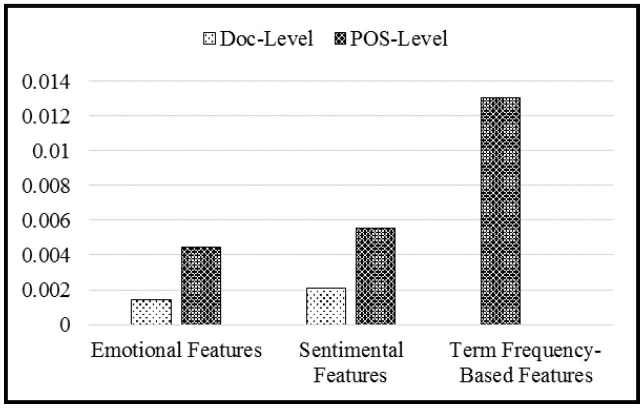

Fig. 2 Feature Importance Comparison of Five Feature Sets on NR Dataset

Fig. 3 illustrates the feature importance rankings of five feature sets on short text Twitter dataset. Similar to the long text analysis, the importance of emotional and sentimental features provided at POS level is superior to the importance of document level features. Unlike the long text analysis, the emotional features are higher than that of the sentimental features.

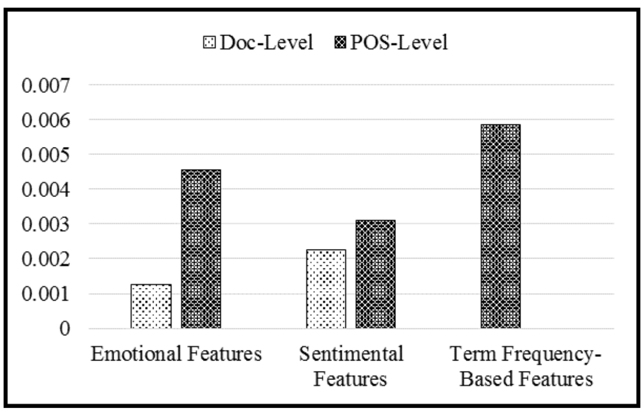

Fig. 3 Feature Importance Comparison of Five Feature Sets on TAB Dataset

This is due to the nature of the short text data. Since the length of the text is limited to at most 280 characters, the number of available words in short text is restricted to 40 words. Satire detection using the sentiment extracted from this words is not adequate, and its sentiment offers the surface level emotion which is not sufficient for satire detection. Thus, POS level emotional features are more important than the sentimental features in short text data.

Similarly, the individual features of each feature set are also ranked according to their important score. Topmost POS level and document level features of the proposed emotion-based satire detection are reported in Table 2. The formation used in this table is <FeatureSets>_<Level> <Features>. *SN* represents the emotional features extracted using SenticNet. *Senti* represents the sentimental features extracted using VADER. *TF* represents the term frequency-based features and *Doc* represents the document level features. Meanwhile, *N, V, Adj,* and *Adv* represent the POS level – Noun, Verb, Adjective, and Adverb respectively.

Table 2: Top 10 Features of POS Level and Document Level Feature Extraction

NR Dataset	***TAB Dataset***
• TF_'	• TF_Lower
• TF_"	• Senti_DocNeu
• TF_)	• Senti_NNeu
• TF_Lower	• SN_AdvAttention
• Senti_DocCompound	• TF_'
• SN_DocAptitude	• TF_Upper
• TF_(	• SN_NSensitivity
• Senti_VCompound	• SN_NAttention
• TF_Upper	• TF_.
• Senti_DocNeu	• SN_NPleasantness

In long text NR dataset, POS level sentimental features and term frequency-based features appear more common than the other features. On the other hand, document level features are not the most common one in long text analysis.

In short text dataset TAB, the number of uppercase and lowercase letters of the term frequency-based features stand as the most important features due to the rich of punctuation words in tweets. In contrast to long text datasets, emotional features appear more frequently than the sentimental features.

Meanwhile, the document level emotional features do not appear frequently in the top-ranked lists.

These results also inspire that the POS level features are more important than the document level features in emotion-based satire detection.

4.2 Performance of each Feature Extraction Model

In this subsection, the performance of the proposed POS level emotion-based features are compared with the performance of the word-based features (BOW). Experiments are set up by using 10-Fold cross-validation.

Table 3: Classification Results of Proposed Emotion-Based Features

Dataset	*Accuracy*	*Precision*	*Recall*	*Specificity*	*F1*
NR	89%	0.897	0.895	0.934	0.88
TAB	83%	0.834	0.837	0.796	0.83

Table 4: Classification Results of Word-Based Feature Extraction (BOW Model)

Dataset	*Accuracy*	*Precision*	*Recall*	*Specificity*	*F1*
NR	80%	0.799	0.798	0.732	0.851
TAB	66%	0.658	0.659	0.372	0.659

As shown in Table 3 and Table 4, the performance of the proposed emotion-based features outperforms the performance of the word-based features (BOW) in both short text and long text data. In long text NR dataset, the proposed features are about 9% better than the word-based features. In short text TAB dataset, since the word-based features mainly depend on the number of word occurrences, its performance significantly drops to 66%. As a worst case, it offers 0.372 specificity value. Thus, the BOW model misclassifies most of the non-satirical data. This shows that BOW model cannot handle the problem of high bias error in short text data whereas the proposed emotion-based features can handle the problem of high bias error in both short text and long text data.

4.3 Performance of each Feature Extraction Model on Existing Efforts

This subsection describes the performance of each feature extraction model on the datasets used in the existing efforts. As summarized in Table 5, there are two types of data distribution: balanced and imbalanced dataset.

Table 5: Dataset Description of Existing Efforts

Dataset	*Detection*	*Data Type*	*Satire*	*Non-Satire*	*Total*
PR2014	Satire	Product Review	375	375	750
TABACL	Satire	Tweets	16,596	16,596	33,192
NR2009	Satire	News Articles	233	4,000	4,233
PRRavi	Irony	Product Review	499	2,498	2,997
FNRRavi	Fake News	News Articles	393	1,272	1,665
TABFY	Satire	Tweets	3,103	33,756	36,859

In PR2014 and TABACL datasets, the number of positive data (i.e. satirical data) and the number of negative data (i.e. non-satirical data) are equal. This type of dataset is known as the balanced dataset. These two datasets are classified by using Random Forest classifier.

However, the number of negative data is actually different from the number of positive data because of the rare phenomena of figurative languages compared to literal languages in the daily stream. This type of dataset is known as the imbalanced or class-imbalanced dataset. Since the size of the positive and negative data used in this domain is very different, it is essential to balance the dataset before driving the classification. Thus, Balanced Bagging classifier (i.e. an ensemble bagging classifier with additional balancing) is used instead of random forest classifier.

Table 6: Classification Results of BOW Model on Existing Efforts

Dataset	*Accuracy*	*Precision*	*Recall*	*Specificity*	*F1*
PR2014	71%	0.703	0.711	-	0.70
TABACL	62%	0.615	0.616	0.316	0.61
NR2009	57%	0.567	0.567	0.576	0.56
PRRavi	83%	0.834	0.834	0.869	0.83
FNRRavi	92%	0.922	0.922	0.918	0.92
TABFY	93%	0.931	0.931	0.936	0.93

Table 7: Classification Results of Proposed Emotion-Based Detection on Existing Efforts

Dataset	*Accuracy*	*Precision*	*Recall*	*Specificity*	*F1*
PR2014	73%	0.741	0.741	-	0.74
TABACL	72%	0.723	0.725	0.723	0.72
NR2009	75%	0.750	0.766	0.794	0.76
PRRavi	94%	0.942	0.942	0.971	0.94
FNRRavi	95%	0.948	0.947	0.935	0.94
TABFY	93%	0.932	0.931	0.938	0.93

Table 6 and Table 7 describes the performance of each feature extraction model on the datasets used in the existing efforts. In the PR2014 dataset used in [5], the proposed emotion-based detection is about 2% higher than the word-based feature extraction model. Meanwhile, the BOW model offer only 0.709 F1. In TABACL used in [12], the proposed emotion-based detection is about 10%

superior to the BOW model. Hence, the word-based feature extraction models do not perform as in proposed emotion-based detection.

In NR2009 dataset [7], the best performance provided by the proposed emotion-based detection is 0.768 F1. Meanwhile, the BOW offer only 0.567 F1. In PRRavi dataset [3], the proposed emotion-based detection offers 0.942 F1 whereas the word-based model offer only 0.834 F1. In FNRRavi dataset [3], BOW model gives 0.922 F1 whereas the proposed emotion-based detection offers the highest F1 0.947 which is about 3% higher than the word-based model. Meanwhile, the proposed system achieves 0.932 F1 in TABFY dataset [2] whereas the word-based model offers 0.931 F1.

It shows that the performance of the proposed POS level emotion-based features surpasses the performance of the word-based feature extraction model.

5. Conclusions and Future Work

In conclusion, an approach to detect satirical figurative language from an emotional point of view has proposed. It contributes in two ways: (1) although the ironic dimension of the language causes difficulty in the detection of satire from emotion, emotion-based satirical language processing is proposed using the emotion lexicon SenticNet and VADER. (2) Emotions in figurative languages are ambiguous and often lead to high bias errors, but a series of experiments are performed to prove that the proposed POS level emotion-based features can tackle to problem of ambiguous nature of emotion.

However, since the proposed POS level emotion-based features are extracted by using thesaurus-based approach (i.e. SenticNet and VADER), the performance of the emotion lexicon limits the performance of the proposed system. In some cases, non-satirical data are incorrectly classified as satire data especially in a simple negative sentence.

Thus, relying solely on the emotion of targeted text is not enough, it requires additional features to distinguish between satirical and non-satirical language. Hence, the contextual emotions embedded in the document and the contextual emotion between conversations should be considered in the future analysis.

Appendix

Appendix 1: Selected Features used in the Proposed Emotion-Based Detection

NR Dataset	***TAB Dataset***
• TF_'	• TF_Lower
• TF_"	• Senti_NNeu
• SN_AdvSensitivity	• TF_'
• Senti_AdvNeu	• SN_AdvAttention
• TF_Lower	• TF_Upper
• TF_)	• SN_NPleasantness
• TF_-	• SN_NAptitude
• Senti_VCompound	• TF_.
• Senti_VNeu	• SN_NAttention
• TF_(	• SN_NSensitivity
• Senti_AdjNeu	• SN_AdjAptitude
• TF_Upper	• Senti_AdjNeu
• Senti_NCompound	• SN_VAptitude
• SN_NAttention	• SN_AdjPleasantness
• SN_AdjAptitude	• SN_VAttention
• SN_AdvAptitude	• Senti_VCompound
• Senti_VNeg	• Senti_VNeg
• TF_$	• Senti_VNeu
• Senti_AdvNeg	• SN_VPleasantness
• Senti_NNeu	• SN_AdvSensitivity
• Senti_AdjCompound	• SN_AdvAptitude
• SN_NAptitude	• SN_AdjAttention
• Senti_AdjNeg	• Senti_AdvNeu
• TF_\	• Senti_NCompound
• Senti_AdjPos	• SN_VSensitivity
• SN_AdjSensitivity	• SN_AdvPleasantness
• TF_?	• TF_{
• SN_AdvPleasantness	• TF_,
• SN_VPleasantness	• TF_*
• Senti_AdvCompound	• SN_AdjSensitivity
• SN_VAptitude	• TF_:
• TF_}	• Senti_AdvCompound
• SN_AdvAttention	• Senti_AdjPos

References

[1] A. Reyes, P. Rosso, and D. Buscaldi, “From humor recognition to irony detection: The figurative language of social media”. Data & Knowledge Engineering, Vol. 74, 2012, pp.1-12.

[2] F. Yang, A. Mukherjee, and E. Dragut, “Satirical News Detection and Analysis using Attention Mechanism and Linguistic Features”, 2017, arXiv preprint arXiv:1709.01189.

[3] K. Ravi, and V. Ravi, “A novel automatic satire and irony detection using ensembled feature selection and data mining”. Knowledge-Based Systems, Vol. 120, 2017, pp.15-33.

[4] P.P. Thu, and T.N. Aung, “Implementation of Emotional Features on Satire Detection”, International Journal of Networked and Distributed Computing, Vol. 6(2), 2018, pp. 78-87.

[5] S. Skalicky, and S. Crossley, “A statistical analysis of satirical Amazon.com product reviews”. The European Journal of Humour Research, Vol. 2(3), 2015, pp.66-85.

[6] A. Frain, and S. Wubben, “SatiricLR: a Language Resource of Satirical News Articles”. In LREC. 2016.

[7] C. Burfoot, and T. Baldwin, “Automatic satire detection: Are you having a laugh?” In Proceedings of the ACL-IJCNLP 2009 conference short papers, pp. 161-164.
[8] D. Davidov, O. Tsur, and A. Rappoport, “Semi-supervised recognition of sarcastic sentences in twitter and amazon”. In Proceedings of the fourteenth conference on computational natural language learning, 2010, pp. 107-116.
[9] D.G. Maynard, and M.A. Greenwood, “Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis”. In LREC 2014 Proceedings. ELRA.
[10] C.H.E. Gilbert, “Vader: A parsimonious rule-based model for sentiment analysis of social media text”. In Eighth International Conference on Weblogs and Social Media (ICWSM-2014).
[11] E. Cambria, S. Poria, D. Hazarika, and K. Kwok, “SenticNet 5: discovering conceptual primitives for sentiment analysis by means of context embeddings”. In Proceedings of AAAI. 2018.
[12] F. Barbieri, F. Ronzano, and H. Saggion, “Do We Criticise (and Laugh) in the Same Way? Automatic Detection of Multi-Lingual Satirical News in Twitter”. In IJCAI 2015, pp. 1215-1221.
[13] P.P. Thu, and N. Nwe, “TwitPre: Tweets Preprocessing Tool for Social Media Analysis”, Proceedings of the 15th International Conference on Computer Applications (ICCA 2017), February 16-17, 2017, Yangon, Myanmar, pp. 123-129.
[14] P.P. Thu, and N. Nwe, “Impact analysis of emotion in figurative language”. In Computer and Information Science (ICIS), 2017 IEEE/ACIS 16th International Conference on (pp. 209-214). IEEE.
[15] P.P. Thu, and N. Nwe, “Implementation of emotional features on satire detection”. In Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 2017 18th IEEE/ACIS International Conference on (pp. 149-154). IEEE.
[16] P.P. Thu, and T.N. Aung, “Effective analysis of emotion-based satire detection model on various machine learning algorithms”. In Consumer Electronics (GCCE), 2017 IEEE 6th Global Conference, pp. 1-5.
[17] T.K. Ho, “Random decision forests”. In Document analysis and recognition, 1995. Proceedings of the third international conference, Vol. 1, pp. 278-282.