ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 6, Issue 1, January 2017

An Efficient Approach for Sarcasm Recognition on Twitterusing Pattern-Based Method

Khan ShehlaKulsum, Prof.S.G. Vaidya

***Abstract* - Sarcasm is widely used on social media to make a remark that means the opposite of what you actually want to say, or to make criticism, or to hurt somebody's feelings. Sarcasm transforms the polarity of the statement into its opposite. However, identification of sarcasm is harder even for human-beings due to its intrinsic ambiguous nature.Sentiment Analysis systems are concerned with the automatic identification of polarity of the contents, but they fail to take into account sarcastic contents. Misinterpreting sarcastic utterance can lower their performance. Therefore automatic recognition of sarcastic contents on data collected from social networks can highly improve the performance of Sentiment Analysis systems as well as many Natural Language Processing (NLP) applications.In this paper, we propose an automated system for detection of sarcasm on Twitter by using features related to sentiments, punctuations,semantic and patterns. These features cover different types of sarcasm associated with wit, whimper and evasion. Machine learning algorithms are used to classify whether a given tweet is sarcastic or non-sarcastic.We study the impact of these features on the efficiency and performance of machine learning to detect sarcasm and assess its significance for the classification.**

***Index Terms* -Sarcasm Detection, Sentiment analysis, Machine Learning, Twitter, Hashtag.**

I. INTRODUCTION

One of the rapidly growing areas of Natural Language Processing (NLP) is sarcastic sentiment analysis with research being conducted to find efficient ways to analyze sentiments in the written text to identify sarcasm, irony or humor on social media. Sarcasm is a sophisticated form of speechthat is used in micro-blogging websites such as twitter and other social networks. Sarcasm is usually used forthe purpose of criticism or mockery. Sarcasm transforms the polarity of a positive or negative statement into its opposite. The recognition of sarcasm present in the context can highly improve the performance and accuracy of Sentiment Analysis system [4] and Natural Language Processing (NLP) applications.During last few years researches have been carried out in the field of sentiment analysis of the textual data available on Online Social Media (OSN). The data that is available on Twitter, Facebook, Amazon, etc. is used by many enterprises and organizations to study the sentiments and opinions of people towards popular products [1], movies reviews [2], political events [3], future trends in the stock-market, etc.

Sarcasm is an activity that is intended to make another person feel stupid or to show them that you are annoyed by an act of speaking or writing in a way that is opposite of what you actually meant.Sarcasm is used by the speaker to convey his message in an implicit way. Sentiment analysis is a process of identifying and aggregating the feelings and opinions expressed by people on social media towards a particular topic. Sentiment Analysis systems automatically determinepolarity of the contents as positive, neutral or negative, but they fail to take into account sarcastic contents which result in the lower performance of these systems.Liebrecht et al. [6] argued how the presence of sarcasm can switch the polarity of the text. D. Maynard and M. Greenwood[7] suggested certain rules to help to determine the polarity of tweets when the sarcastic utterance is discovered.

Sarcasm is considered one of the most difficult problems in Sentiment Analysis. Sentiment Analysis identifies opinions or feelings expressed by the Internet users towards a specific movie review, political event, popular product etc. which helps in decision making, enhance the product and services, marketing and customer support, etc. Detection of sarcasm is important for the enhancement of sentiment analysis systems[6] [7]. Apart from this, failure to detect sarcasm can cause problems in some Natural Language Processing systems such as online review synthesis systems, dialogue systems, opinion oriented summarization[16], etc.because of inability to detect sarcastic comments.

Sarcasmis defined as an activity in which the person saysopposite of what hemeans or speaks in a way that is intended to make someone else feel stupid or show them that he is angry".Sarcasm can be broadly categorized into wit, whimper, and evasion.

- Sarcasm is used as wit for the purpose of being funny. In this case, the person tends to exaggerate or employ a special form of speech or use a different tone so as to make it easy to recognize. Often this is expressed by making use of exclamation marks, question marks, capital letter words, sarcastic emoticons like :P.
- Sarcasm is used as whimper to show how annoyed or angry the person is. Here, the person makes use of exaggeration or employs very positive expressions to describe a negative situation.
- Sarcasm is used as evasion in those situations when the person wants to avoid giving a clear respond and makes use of sarcasm by using unusual words, intricate sentences, etc.

Twitter is a micro-blogging service which allows its user to read and post short text messages (maximum message length of 140 characters) called tweets. Twitter is most widely used microblogging platform by the people all around the world to report the events happing in real-time, share their emotions, and express their attitude or opinion towards a particular topic. On Twitter, around 500 million tweets are sent per day and the number of active users is estimated to be around 317 million.In this paper, we explore the sarcasm present in feeling and its detection on Twitter. The problem in understanding the sentiments of users for performing the analysis is due to the informal language used in tweets and limitation of 140 characters per tweets. In addition to this, the presence of sarcasm complicates the problem even further.

We aim to propose a pattern-based approach for recognition of sarcastic tweets on Twitter. Sarcasm relating to evasion, wit, and whimper are detected by making use of different components of the tweets and do not make use of already assembledtweeter knowledge-base as proposed by Rajadesingan et al. [8].For determining sarcastic utterance on twitter using classification technique, we first create the dataset that contains a collection of tweets including sarcastic as well as non-sarcastic tweets. The data that is collected is not in the proper format; therefore this data needs to be pre-processed. This is followed by feature extraction where we extract features relating to sentiments, punctuations, and common sarcastic patterns. Once the features are extracted then we perform the classification using machine learning techniques. Finally, the system is able to determine whether the given tweet is sarcastic or non-sarcastic based on this classification.Listed below are important contributions of our paper:

1) Our system efficiently detects sarcasm and classify the tweets depending on whether they sarcastic or not.
2) We also assess the added value of features related to sentiments, punctuation, syntax, and pattern in terms of precision on sarcasm detection process.

II. RELATED WORK

Researchers are giving more thought to the automatic analysis of opinions and sentimentson Twitter for last few years and also lots of papers that address the issue of tweets classification have been published. In general, we broadly classify the approaches for sarcasm detection into following three categories rule-based approach, semi-supervised approach, supervised approaches. In the supervised approaches,there are mainly three sets of feature which includes n-gram based features, sentiment-based features,and sarcastic pattern-based features. The contextual features are added to enhance the classification.

In general, people use sarcasm and irony in their daily communication. Sarcasm has also been studied in fields of psychological, neurobiological perspectives and linguistics to understand human behavior [8]. The automatic detection of sarcasm in the textual literature is considered a problem that is hard to resolute [6] and few studies have been done in this field.

The linguistic analysis offers some understandings of the characteristics of sarcasm. In General, the tweet that contains a positive sentiment might later contain one of the following types of arguments: explicit markers, exclamations, and intensifiers. The latter two types of words induce hyperbole.Hyperbole is a combination of different features like interjections, intensifiers, punctuations marks, quotes, etc. The hyperbolic text has greater probability of being categorized as sarcastic.Bharti et al [15] made two proposals for detection of sarcasm on Twitter. The first approach was a parsing-based lexical generation algorithm which recognizes sarcasm as a contradiction between positive sentiment and negative situation as well as a contradiction negative sentiment and positive situations. The second approach suggested an algorithm that determines the presence of sarcasm in tweets by identifying the use of interjections such as wow, yay, oh, aha etc. that are present when the tweet starts.

Sriram et al. [9] proposed a method for short text classification on Twitter to improve information filtering. They used non-context features in the text such as the occurrence of sentiment words, slangs, phrases related to time event and the user information to classify tweets into a predefined set of classes which includes events, opinions, etc.Akcora et al. [10] suggested a system that is able to identify breaking point in the public opinion by recognizing the emotive-pattern and the word-pattern that is present onTwitter datawhich is used to determinehow the public opinion variesover the time.

O.Tsuret al. [4] proposed an approach for recognition of sarcastic sentences on online product. They used a semi-supervised algorithm for identification of sarcasm called SASI that was used to detect sarcastic utterance in product reviews. It consists of two main stages namely, semi-supervised approach for pattern acquisition which followed by classification of sarcasm. They found that there exist certain features using which sarcastic utterances can be characterized. The training set that was used consisted of a small set of sentences. These sentences were manually assigned score 1 to 5 where score 1 stands for absence of sarcasm and score 5 stands for very sarcastic sentence. Syntactic and pattern based features were extracted that were used to build feature vectors and then are utilized construct the model that assigns scores to the unlabelled examples.

The sentiment analysis on Twitter has been an area of research for last few years which aims to identify the opinion of users. In his work, Rajadesingan et al.[8] highlights the limitations of sentiment analysis tools when sarcasm is present in the contents. They used behavioral modeling framework called SCUBA.SCUBA determines the likelihood of the user being a sarcastic person by analysing the past tweetsof that user. This behavioral modeling approach for detection of sarcasm utilizes features which capture the different form of sarcasm.This supervised learning framework uses these features along with categorised data to decidewhether the tweet contains sarcasm or not.One of the shortcomings of this approach is that when prior knowledge of the user is not available with the system then its performance decreases. Another shortcoming is that the knowledge base grows very fast and the training has to be redone every time the new tweets get collected.

Riloff et al. [14] suggested an approach for identifyingthe particular form of sarcasm where a positive-emotion contrasts with a negative-circumstance. They collected tweets that contain sarcastic contents and developed a bootstrapping algorithm. The algorithm starts with a single seed word “love” that detect and learn expressions automatically when positive sentiment contrasts with negative situations. But the problem with this approach is that majority of sarcastic tweets on Twitter do not belong to this class of sarcasm. Another issue is thatthis approach depends on the presence of the all probable negativesituations in the dataset that is used for training, which reduces itsperformance when dealing with new tweets.

Davidov et al. [5] in his paper recommended a semi-supervised algorithm to identifywhether the contents are sarcastic in nature or not by determining sarcastic pattern used in text. Their approach used two data sets, one from Twitter while another fromAmazon. Their approach was based on recognizing the sarcastic pattern in written text. They extracted syntactic and pattern-based feature sets to be used in the feature vector. Such patterns exist on social media. The words are classified into high-frequency words and content words. The pattern is defined as an arrangement of high-frequency words and spaces for content words. They used a sample of 90 phrases that was classified as sarcastic and 90 phrases that were classified as non-sarcastic. Their system achieved an F1 score of 0.78 on a product review test. This approach might be misleading if the dataset used for training is not big enough or not balanced. Another issue is that it treats context word in the same way irrespective of their grammatical functions. The efficiency is reduced since these patterns do not consider emotional contents nor do they make the distinction between sentimental words and non-sentimental words.

Sarcastic tweets on Twitter are often marked explicitly by hashtag #sarcasm so that it becomes easy to understand sarcastic utterance in short text messages in its literal meaning. These hashtags are considered to be the digital additional linguistic equal for the non-verbal expressions that are often used by people in their communication to convey sarcasm. Liebrecht et al. [6] discussed how the presence of sarcasm can switch the polarity.They developed a system that that is capable of recognizing sarcastic tweets. Their system was trained on seventy-eight thousand tweets that are marked by hashtag #sarcasm.The classification algorithm made use of Balanced Winnow classification which is implemented in Linguistic Classification System.

Maynard et al. [7] showed that performance of sentiment analysis can be highly improved if we could recognize the presence of sarcasm within the sarcastic statements. Theyproposed a rule-based approach for sarcasm detection and suggested a set of rules to determine the polarity of the tweet when sarcasm is present depending on the contents of the hashtag and the sentiment of the tweet.This approach relied on the hashtags provided by users to determine sarcasm in their tweets. They tokenized the hashtagto determine positive as well asnegative words present in the hashtag and then used this information for detecting the sentiments in the written context. When sarcastic statements are found thenthe polarity of an opinion was simply reversed.The shortcoming of this approach was that it fails to handle sarcasm when it is not mentioned in the hashtag.

Dong et al. [11] proposed adaptive recursive neural network. This classification framework makes use of textual context and syntactic construct to transmit the sentimentalities of words on the way to the target. It performs the computation in bottom-up manner. Vector representations are computed recursively and the sentiment label is predicted according to this vector.

Campbell et al. [12]determined whether there exist certain conditions to make sense of sarcastic irony and studied the contextualcontents that indicatesarcasm. They suggested that for identifying sarcasm we need four entitiesnamely allusion to failed belief, pragmatic deviousness, adverse tension and existence of a victim, as well as stylistic constituents.

Tepperman et al. [13] studied the presenceof “yeah right!” expression to determine whether the contents are sarcastic. They proposed an approachthat makes use of prosodic, contextual and spectral hintsto determinethe presence of sarcasm in spoken dialogues.This approach is efficient in recognizing whether a specific expression is sarcastic, but the drawback with this approach was that in the absence of such components it is very difficult to identify sarcasm.

III. PROPOSED APPROACH

The aim is to classify the given tweet as sarcastic or non-sarcastic. Therefore, we construct a learning set which is the collection of tweets including both sarcastic and non-sarcastic in nature. Then we extract features relating to sentiments, punctuations, semantics and pattern. We extract these features such that they utilize various components present in the tweet as well as covers different forms of sarcasm.Finally classification is done by using machine learning algorithms. The tweets are collected by querying Twitters Streaming API. These tweets are manually checked and then they are annotated with different levels of sarcasm and are used during the experiments.

Data

As mentioned above the dataset consists of tweets collected by usingstreaming API provided by Twitter. Sarcastic tweets were collected by querying the API for tweets that containedhashtag

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 6, Issue 1, January 2017

#sarcasm. Non-sarcastic tweets were collected on various issues that have some emotional contents.

Data Pre-processing

The tweets may contain simple text as well as references to URL's, other Twitter users (using @<user>) or a content-tag called hashtags, for example #sarcasm, #ihateyou,etc. Therefore it is important to first clean such text data before we could extract the features from it. For sarcastic tweets, we discard all tweets that contain https-address referring to a photo that contains sarcasm. We also remove the tweets written in other languages or those tweets whose length is less than 3 words as well as the duplicate tweets.

In total we prepared three sets as given below:

1) Training set: This set contains a collection of sarcastic and non-sarcastic tweets which is used for training our model. We collected around six thousand tweets out of which three thousand tweets are sarcastic and other three thousand are non-sarcastic. We manually classified them with scores ranging from 1 to 6 where score 1 means highly non-sarcastic and score 6 means highly sarcastic.
2) Optimization set: This set is used to optimize the parameters for the feature set and is used during the experiment process. As mentioned above the sarcastic tweets are collected by querying the Twitter's streaming API. Here we removed very short tweets or those written in other languages or those containing URL's referring to photos.
3) Test set: The tweets in this set are manually classified depending on whether they are sarcastic or non-sarcastic. This is used to calculate the performance of our system.

System model

The System-model that is used to determine the sarcastic utterance in the tweet is shown below.

Twitter is used by people to express the views and opinion about a particular topic or to share their feelings and thoughts. These tweets are collected and processed. The NPL classifier is used to perform tokenization and part of speech tagging. Features relating to sentiment, punctuation, and pattern are extracted. Finally, the classifier is used to perform the classification on tweets to determine whether the tweet is sarcastic or non-sarcastic.

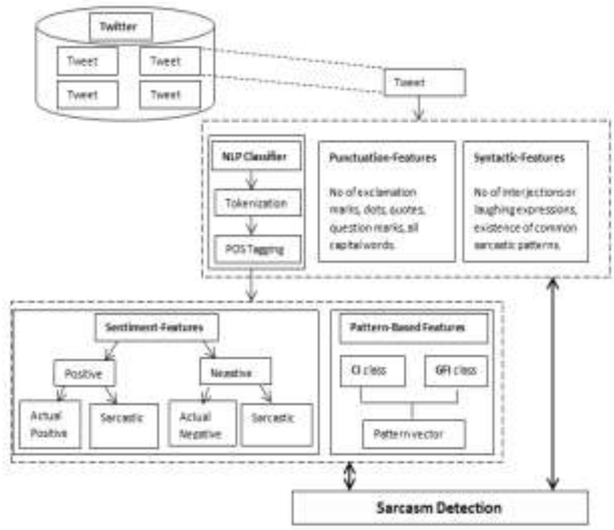


Figure1: System model for detection of sarcasm on Twitter

NLP Classifier

The classifier for natural language processing is a tool that takes in the data-items and categorizes them into one of the k-classes. In natural language processing (NLP) we perform various tasks such as tokenization, part of speech tagging (PoS), lemmatization, etc.

Tokenization

Tokenization is the process of splitting the given sequence of text into pieces called tokens. Tokenization generally occurs at the word level. A token is a sequence of characters in the given document that are grouped together to form a useful semantic unit which can later be used for processing. Tokens can consist of the sequence of alphabets, or it can be alphanumeric. In lexical analysis, tokenization breaks the stream of text into words, phrases, or other meaningful componentsreferred to as tokens. Tokenization outputs a list of tokens that becomes input for further analysis.Some of the heuristics used bytokenizerare as follows (i) the whitespace characters such as a space or line break are used to separate the tokens, (ii) tokens can be composed of contiguous strings of alphabetic characters or numbers or alphanumeric characters, (iii) tokens may or may not include punctuations and whitespaces.

Part-Of-Speech (POS) Tagging

In computational linguistics,Part-Of-Speech tagging is also called as word-category disambiguation. Part-Of-Speech tagger reads the text and assigns parts of speech such as noun, verb, adjective, adverb, conjunction, interjection, etc. to each token based on both its definition and its context which is constructed depending on its relationship with adjacent and related words in the sentence. Most applications in computational science require further fine-grained POS tagging, for example,nouns can be distinguished as the plural, possessive, and singular forms.It uses NN for singular common nouns, NNS for plural common nouns and NP for singular proper nouns. The algorithms used by POS tagger falls into one of the two categories namely rule-based and stochastic.

Feature Extraction

We extract features related to sentiment, punctuations, semantics and pattern which are explained in detail below.

Sentiment based features

The most common form of sarcasm that is used on social media is called whimper in which the user employs positive expression to describe a negative situation and vice versa. Here sarcasm is expressed by using contradictory sentiments for example, “I love being ignored all the time”. Therefore the system should be able to determine whether the positive sentiment is actually a positive emotion or sarcastic. Similarly, a negative sentiment can be an actual negative emotion or can be sarcastic.

To detect the sentiment based features we determine any kind of contradiction that exists between the sentiments of words and other components in the tweet. These sentiment contents are extracted and counted. Two listsof words are maintained one for positive words and another for negative words using SentiStrength database where positive words have a score ranging from 1 to 5 where score 1 means almost positive and score 5 means extremely positive. Similarly negative words are also scored ranging from -1 to -5 where score -1 mean almost negative and score -5 means extremely negative. The emotional contents of adjectives (POS tags - JJ, JJR, JJS), adverbs (POS tags – RB, RBR, RBS) and verbs (POS tags – VB, VBD, VBG, VBN, VBP, VBZ) have higher emotional contents than the noun. Therefore, the positive and negative words that have above-mentioned POS-tags associated with them are counted again and form two more features that represent highly emotional positive words and highly emotional negative words respectively. Emoticons are also used in tweets to express positive, negative and sarcastic feelings. These emoticons are used to extract features relating to sarcasm. In addition, the hashtags in the tweets also contain emotional contents. For example the tweet “Thanks for being there for me #ihateyou”, does not mean that the user is thankful but instead blames for not being there. Therefore features relating to positive and negative hashtags are also extracted.

Punctuation based features

The hyperbole, pragmatic and lexical features play a very important role in the detection of sarcasm in the textual data. Sarcasm is a sophisticated form of discourse in which the speaker make use of certain behavioral aspects such as facial gestures like rolling of eyes, low tons, or exaggeration. In written text, these behavioral aspects are translated into the use of punctuations, repetitions of vowels, use of all capital words, etc. Therefore to detect this kind of sarcastic utterance in tweets we count the number of exclamation marks, question marks, dots, quotes, all-capital words as well as vowels repeated more than twice (e.g. looooove). The feature is set to true if such a word exist otherwise it is set to false.

Pattern based features

To extract pattern based features we divide the wordsinto two classes namely “CI” and “GFI”. The classification of the words present in tweet into “CI” and “GFI” classes is done based on that word’s POS tag.The class “CI” contains the words of which the content is important and the class “GFI” contains the words for which grammatical

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 6, Issue 1, January 2017

functions is more important. The word is lemmatized if it belongs to the “CI” class and if the word belongs to “GFI” class then we replace this particular word with the expressions given inTable1.

Table1: POS-tags and expression that are used to replace the words that belong to “GFI” class.

GFI Class		
POS Tagging	**Meaning**	**Expression**
CD	Cardinal number	[CARDINAL]
FW	Foreign-word	[FOREIGNWORD]
LS	List-marker	[LISTMARKER]
MD	Modal	[MODAL]
NN,NNP, NNS	Noun	[NOUN]
PRP, PRP$	Pronoun	[INTERJECTION]
SYM	Symbol	[SYMBOL]
UH	interjection	[INTERJECTION]
WDT, WRB, WP	Wh-determiner	[WHDETERMINER]

Next, for each tweet, we generate a vector of words. Patterns are defined as ordered arrangement of words that are extracted from the learning set. We create N_f features as shown in Table 2.

Table2: Pattern Features for different levels of sarcasm and pattern lengths

		Length of Pattern		
		L_1	L_2	… L_N
Level of Sarcasm	1	f_{11}	f_{12}	… f_{1N}
	2	f_{21}	f_{22}	… f_{2N}
	:	:	:	: :
	6	f_{61}	f_{62}	… f_{6N}

Let f_{ij}be the feature where i is the level of sarcasm present in the tweet and j is the length. The feature f_{ij} denotes the degree of similarity of the given tweet to the patterns that have the level of sarcasm i and length j. Let t is the given tweet and p is the

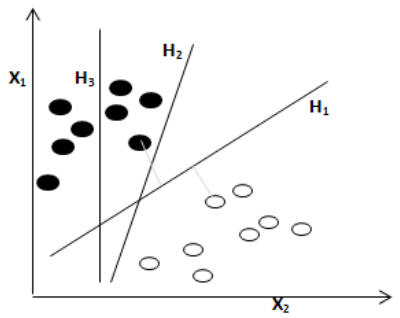
$$resm(p,t) = \begin{cases} 1 & \text{If the pattern is present in the same order in the given tweet.} \\ \alpha\ .\ n/N & \text{If n out of N words of the pattern appear in the same order in the given tweet.} \\ 0 & \text{If none of the words of our patterns appear in the given tweet.} \end{cases}$$

pattern in the training set, we compute the degree of similarity denoted by resm(p,t) as follows.

SVM Classifier

A Support Vector Machine (SVM) is a supervised learning model. SVM has associated training algorithms that are used to analyze the data. In machine learning, A Support Vector Machine (SVM) constructs a hyperplane in high dimensional space which is used for tasks such as classification, regression, etc. The SVM defines norms for looking for decision surface that is hugely far away from any point. The margin of the classifier is the distance from this decision surface to the closest data-point.

When given a set of learning data, in which each example is marked to show which class this example belongs to, then SVM learning algorithm builds a model that assigns the new example to one of these classes. In general, the SVM model represents the training examples as points in spaces that are mapped such that the training examples belonging to different classes are separated by a gap as wide as possible. When a new example is given then based on which side of the gap it fall in, the SVM predicts the class to which this example belongs to.



Here, H_1 separates the classes the maximum margin, H_3 does not separate the classes and H_2 separates them but only with a small margin.

Naïve Bayes Classifier

Naïve Bayes classifier is a probabilistic classifier that is based on application of Bayes’ theorem with strong hypothesis on independence between the features. The naïve Bayes classifier is based on the assumption that given the class variables, the value of a particular feature is independent of the values of any other feature.Suppose we are provided with a set that consists of classes, the this classifier determines to which class a given example belongs to. The model assigns class-labels to the given instance, represented as vector of feature values.

These class-labels are drawn from a finite set.Naïve Bayes requires only a small learning dataset to evaluate the parameters needed for classification.

Performance Measures

To evaluate the efficiency of the proposed approach,we use the performance measures such as accuracy, precision, and recall. The terms tp, fp, tn, fn denotes the number of true positives, false positive, true negative and false negative respectively.

- Accuracy: high accuracy is achieved when the results returned by the algorithm are significantly more relevant than irrelevant ones.

$$\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn}$$

- Precision:we can define precision for a particular class astp (true positives) divided by sum of tp (true positive) and fp (false positive) that is the over-all number of elements that belong to the positive-class.

$$\text{Precision} = \frac{tp}{tp + fp}$$

- Recall:we can define recall for a particular class as tp (true positives) divided by sum of tp (true positive) and fn (false negative) which is nothing but the over-all number of elements that essentially belong to the positive-class.

$$\text{Recall} = \frac{tp}{tp + fn}$$

IV. CONCLUSION

In this work, we proposed an automated system to determine the presence of sarcastic utterance in the given tweet. Given the set of tweets, the proposed system makes use of different components of the tweet to excerpt different feature-sets and employ machine learning algorithm to perform the classification. This system makes use of POS tags to select the common sarcastic patterns that characterize different levels of sarcasm present in the tweet. Since we used a small training set, therefore it does not cover all possible sarcastic patterns. The result can be further improved by using bigger training set.

In future work, we would like to study how the output of this systemthatrecognizes the presence of sarcastic utterance can be utilized to enrich the sentiment analysis system and NLP applications so that their performance can be further enhanced.

REFERENCES

[1] S. Homoceanu, M. Loster, C. Lofi, ‘‘Will I like it? Providing product overviews based on opinion excerpts,’’ in Proceedings of IEEE CEC, pages 26 - 33, (Sept. 2011).

[2] U. R. Hodeghatta, ‘‘Sentiment analysis of Hollywood movies on Twitter,’’ in Proceedings of IEEE/ACM ASONAM, pages 1401-1404, (Aug. 2013).

[3] Juan M. Soler, F. Cuartero, and Manuel Roblizo, ‘‘Twitter as a tool for predicting elections results,’’ in Proceedings of IEEE/ACM ASONAM, pages 1194 -1200, (Aug. 2012).

[4] O. Tsur, D. Davidov, and A. Rappoport, ‘‘ICWSM - A great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews,’’ in Proceedings of fourth Intl.Conf. on Weblogs and Social Media, pages 162 - 169, (May 2010).

[5]D. Davidov, OrenTsur, and Ari Rappoport, ‘‘Semisupervised recognition of sarcastic sentence in Twitter and Amazon,’’ in Proceedings of the 14th Conf. Computational Natural Language Learning, pages 107 - 116, (Jul.2010).

[6] C.Liebrecht, F. Kunneman, and A. Van Den Bosh, ‘‘The perfect solution for detecting sarcasm in tweets #not,’’ Association for Computational Linguistics, pages 29 - 37, (June 2013).

[7]D. Maynard and M. A. Greenwood, ‘‘Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis,’’ in the Proceeding of the 9th Int. Conf. Language Resource Evaluation, pages 4238 - 4243, (May 2014).

[8] A. Rajadesingan, R. Zafarani, and H. Liu, ‘‘Sarcasm detection on Twitter: A behavioral modeling approach,’’ in Proceedings of the 18th ACM Int. Conf. on Web Search Data Mining, pages 79 - 106, (Feb. 2015).

[9] B. Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, and M. Demirbas, ‘‘Short text classification in twitter to improve information filtering,’’ in the Proceedings of the 33rd Int. ACM SIGIR Conf. Research and Develop in Info. Retrieval, pages 841 - 842, (July 2010).

[10] C. G. Akcora, M. A. Bayir, M. Demirbas, and H. Ferhatosmanoglu,‘‘Identifying breakpoints in public opinion,’’ in Proceedings of the 1st Workshop on Social Media Analytics, pages 62 - 66, (July 2010).

[11] LiDong,FuruWei,C.Tan,D.Tang,M.Zhou,andKeXu,''Adaptivere cursive neural-network for targetdependent Twitter sentiment classification,'' in Proc. of the 52nd Annual Meeting on Assoc. for Computational Linguistics, vol. 2(Short paper) , pages 49 - 54, (Jun.2014).

[12] J. D. Campbell and A. Katz, ''Are there necessary conditions for inducing a sense of sarcastic irony?'' in Discourse Processes, vol. 49, pages 459 - 480, (Aug. 2012).

[13] J. Tepperman, D. Traum, and S. Narayanan, ''Yeah right: Sarcasm recognition for spoken dialogue systems,'' in the Proceedings of the InterSpeech 2006 - ICSLP, pages 1838 - 1841, (Sept. 2006).

[14] E. Riloff, A. Qadir, P. Surve, L. De Silva, N. Gilbert, and R.Huang,''Sarcasm as a contrast between positive sentiment and negative situation,'' in the Proceedings of the Conf. Empirical Methods Natural Language Processing, pages 704 - 714, (Oct. 2013).

[15] S. K. Bharti, K.S. Babu, and S. K. Jena, ''Parsing-Based sarcasm sentiment recognition in Twitter data,'' in Proceedings of IEEE/ACM (ASONAM), pages 1373 – 1380, (Aug. 2015).

[16]Minqing Hu and Bing Liu, ''Mining and summarizing customer reviews,'' in Proc. Of tenth ACM SIGKDD intl. conf. on knowledge discovery and data mining, pages 168-177 (2004).

Khan ShehlaKulsum,Department of Computer Science and Engineering,Shreeyash College of Engineering and Technology, Aurangabad, India.

Prof. S. G. Vaidya, Department of Computer Science and Engineering,Shreeyash College of Engineering and Technology, Aurangabad, India.