# ISSN 2395-1621



# Implementation on Fake News Detection using Machine Learning Algorithm

Prasad Vijay Kothavale, Swapnil Manohar Kangane, Prasad D. Shinde, Suhas Ramesh Awale

> ko.prasad28@gmail.com swapnilkangane62@gmail.com prasaddshinde2804@gmail.com suhas123@gmail.com

BE Computer Engineering,

Anantarao Pawar College of Engineering.

# ABSTRACT

## ARTICLE INFO

Machine learning analysis deals with identifying and classifying opinions or sentiments expressed in source text. Social media is generating a vast amount of sentiment rich data in the form of tweets, status updates, blog posts etc. Sentiment analysis of this user generated data is very useful in knowing the opinion of the crowd. Twitter sentiment analysis is difficult compared to general sentiment analysis due to the presence of slang words and misspellings. Knowledge base approach and Machine learning approach are the two strategies used for analyzing sentiments from the text. Public and private opinion about a wide variety of subjects are expressed and spread continually via numerous social media. Twitter is one of the social media that is gaining popularity. Twitter offers organizations a fast and effective way to analyze customers' perspectives toward the critical to success in the market place. Developing a program for analysis to identify and detect fake news, is an approach to be used to computationally measure customers' perceptions. This system uses knowledge base including various patterns for news along with multiple strategies to detect the sentiment expressed in a news article and if a news is genuine or not. Various machine learning and knowledge base approaches are used to compare patterns and apply strategies.

Article History Received: 12<sup>th</sup> June 2021 Received in revised form : 12<sup>th</sup> June 2021 Accepted: 14<sup>th</sup> June 2021 Published online : 14<sup>th</sup> June 2021

Keywords: Machine Learning, Fake News Detection, Decision Tree, Random Forest, LSTM, Text Mining.

#### I. INTRODUCTION

Fake news is where individuals or organizations intentionally publish hoaxes, propaganda and other misinformation and present it as factual. This can include blog and social media posts and fake online media releases. After detecting multiple classes tweets/news will be checked for genuinely.

Twitter has emerged as a major micro-blogging website, having over 100 million users generating over 500 million tweets every day. With such large audience, Twitter has consistently attracted users to convey their opinions and perspective about any issue, brand, company or any other topic of interest. Due to this reason, Twitter is used as an informative source by many organizations, institutions and companies. On Twitter, users are allowed to share their opinions in the form of tweets, using only 140 characters. This leads to people compacting their statements by using slang, abbreviations, emoticons, short forms etc. Along with this, people convey their opinions by using sarcasm and polysemy. Hence it is justified to term the Twitter language as unstructured. In order to extract sentiment from tweets, sentiment analysis is used. The results from this can be used in many areas like analyzing and monitoring changes of sentiment with an event, sentiments regarding a particular brand or release of a particular product, analyzing public view of government policies etc.

A lot of research has been done on Twitter data in order to classify the tweets and analyze the results. In this project we aim to predict the sentiments from tweets by checking the polarity of tweets as positive, negative or irrelevant. Sentiment analysis is a process of deriving sentiment of a particular statement or sentence. It's a classification technique which derives opinion from the tweets and www.ierjournal.org

formulates a sentiment and on the basis of which, sentiment classification is performed. Sentiments are subjective to the topic of interest. We are required to formulate that what kind of features will decide for the sentiment it embodies. In the programming model, sentiment we refer to, is class of entities that the person performing sentiment analysis wants to find in the tweets. The dimension of the sentiment class is crucial factor in deciding the efficiency of the model. For example, we can have two-class tweet sentiment classification (positive and negative) or three class tweet sentiment classification (positive, negative and irrelevant). Sentiment analysis approaches can be broadly categorized in two classes - lexicon based and machine learning based. Lexicon based approach is unsupervised as it proposes to perform analysis using lexicons and a scoring method to evaluate opinions. Whereas machine learning approach involves use of feature extraction and training the model using feature set and some dataset.

#### **Problem Statement:**

News is fastest and easiest way to connect to the world and also to express one. But many people express false sentiments on tweets/news to get attention and to spread it viral. To check sentiment class as well as polarity and further the genuinely of news is essential.

**II. PROPOSED SYSTEM** 





The central outline of the proposed algorithm is as follows.

- First of all news are fetched from news using API
- Then the news are preprocessed
- Then the features of news will be extracted
- Polarity of news will be found
- News will be classified in following classes,
  - o Fun
  - Happiness
  - o Love
  - Neutral
  - $\circ$  Anger and etc.
  - Patterns and strategies will be applied
- Check if news was genuine or not

### **III. ALGORITHM DETAILS**

Collection of tweets with a specific keyword, e.g. "#ParisAttacks" or "Brussels". The Twitter API only allows the collection of such tweets within a ten-day window. For this reason this step must start as soon as an event happens or a fake tweet begins. First of all tweets are fetched from twitter by using Twitt4j API, also particular tweets can be searched. Tweets are then classified according to sentiment classes such as happy, sadness, disgust, anger, etc. Then the tweets are preprocessed such as tokenization, removal of symbols and other things will take place here. Then the features of tweets will be extracted based on keywords to detect the sentiment class. Polarity of tweet will be found based on Stanford NLP algorithm or Open Apache NLP algorithm. After NLP pattern matching as shown in Table 1 will be applied to it and then strategies will be applied. In this method we will collect all three results and if any both of them return true then that tweet will not be genuine.

Sentiment classes for tweet classification are as follows,

- 1. Love/ Happy
- 2. Neutral
- 3. Anger
- 4. Hate
- 5. Sadness
- 6. Surprise
- 7. Disgust

To achieve classification we will use database of respective words of each classes and compare tweet texts with database to detect class. TF-IDF algorithm along with Decision tree will be used for this purpose. For achieving sentiment classification following feature extractions will be used,

- 1. Sentiment-related features
- 2. Punctuation features
- 3. Syntactic and stylistic features
- 4. Semantic features
- 5. Top words
- 6. Pattern-related features

Where tis the tweet, PW and NW are the total score of positive words and that of negative words as returned by SentiScore[1]. In case the tweet does not contain any emotional word, p is set to 0.

(a) Manual analysis of tweets and search for fake tweets. In this step we filter out all the irrelevant tweets. For example, if we collected tweets containing the keyword "Brussels" (due to the unfortunate Brussels attacks), we ignore tweets talking about holidays in Brussels.

(b) Collection of more tweets relevant to the story with keywords that we missed in the beginning of Step 1 (this step is optional). For example, while searching for fake tweets we might come across tweets talking about another fake tweet. We add the keyword that describes this new fake tweet in our tweet collection.

(c) Categories tweets into fake tweets. Group all tweets referring to the same fake tweet.

(d) Identify all the unique users involved in a fake tweet. This set of users will be used in Steps 1 to 2. Collect users' most recent 400 tweets, posted before the start of the fake tweet. This step is required because we aim to examine the users' past behavior and sentiment, e.g. whether users' writing style or sentiment changes during the fake tweet, and whether these features are significant for the model. To the best of our knowledge, this set of features is considered for the first time in the academic literature in building a fake tweet classifier.

1. Collect users' followers (friends). This data is essential for making the propagation graph.

2. Collect users' information, including user's registration date and time, description, whether account is verified or not etc.

Pattern	Description	
atti_cnt	No of users who have favored in weibo	
cmt_cnt	No of users who have commented in weibo	
repo_cnt	No of users who have reposted in weibo	
sent_score	Sentiment score of weibo	
pic_cnt	No of pictures posted in weibo	
tag_cnt	No of #tag in weibo	
mention_cnt	No of @mentioned in weibo	
smiley-cnt	No of smileys in weibo	
qm_cnt	No of question marks in weibo	
fp_cnt	No of first person in the weibo	
Length	Length of the Weibo	
is_rt	Whether the weibo is repost	
Hour	Hour the weibo was posted	
Source	How the Weibo was posted	

**Table 1** Patterns for tweets analysis

The system operates on tweets and tweets are fetched by using Twitt4j API using Java as programming platform. The steps involved in working are as follows,

- 1. Authenticate
- 2. Fetch Tweets
- 3. Preprocess (Tokenization, removal of special symbols and URLs)
- 4. Read Tweet features such as mention counts, tag counts, smiley, URL, etc.
- 5. Perform NLP to detect polarity of preprocessed tweet.
- 6. Match patterns from read features of tweets as shown in table 1.
- 7. Apply strategies such as time of tweet, username, location and compare this info in the know database.
- 8. If polarity is negative and any of the pattern matches or polarity is positive but patterns are matching and strategies returns true then tweet is not genuine

The presented mathematical model is based on set theory of modeling a system where full set of system is considered as set 'S'. This is the simplest way of presenting system in mathematical forms and is very easy to understand.

#### **IV. RESULT AND DISCUSSION**

In the proposed system, Based on the query, we will fetch tweets from twitter account using twitter API. The fetched tweets will be subjected for preprocessing. We will then apply the various patterns and strategic algorithms as well as few machine learning algorithms for NLP for supervise the data. The algorithms result i.e. the sentiment and influence will be represented in graphical manner (pie charts/bar charts). The proposed system is more practical than the existing one. This is as a result of we'll be ready to shrewdness the statistics determined from the illustration of the result will have a sway in a very specific field furthermore influence of negativity spread by fake tweets.

Many of the existing systems as discussed in literature focuses on only NLP algorithms and resulting in sentiment or polarity of the tweet. Our system extends these results with applying various pattern matching as shown in table 1 along with strategies to detect if the tweets were genuine or not so that other users will know how much to trust such tweets. The system also can be used in predicting anonymous or fake profiles on twitter which can be helpful further to mitigate such false tweets.

Comparative results of existing and proposed system is as following Table 2,

Parameters	Existing System	Proposed System
Sentiment Analysis	Yes	Yes
Polarity Detection	Somewhat	Yes
Classification	Somewhat	Yes
Pattern Matching	No	Yes
Fake Tweets	No	Yes
Graphical Analysis	No	Yes
User Alerts	No	Yes

#### Table 2: Comparative Results

With reference to table 2, it is clear that we overcome various problems in existing system and our approach works efficiently.

www.ierjournal.org



Fig.2: Pie Chart For Polarity of Tweets

As shown in Fig.2. The system will detects the polarity of tweets fetched from API. Polarity checking helps classification of sentiments. Also following Fig.3 shows the sentiment classification results in various classes,\



Fig.3: Sentiment Classification of tweets

Finally our system checks how many tweets along of fetched tweets were genuine and results achieved are shown in Fig.4.



Fig.4: Tweet Genuienity

#### V. CONCLUSION

The Developed web application Using ensemble learning, various ML classifier, applying multi-stage classification

strategy, and finally with a deep LSTM model, we got a optimal solution for this stance detection competition. Detect fake Image by using ELA and deep learning.

#### **VI. REFERENCES**

[1] A. Pak and P. Paroubek. "Twitter as a Corpus for Sentiment Analysis and Opinion Mining". In Proceedings of the Seventh Conference on International Language Resources and Evaluation, 2010, pp.1320–1326.

[2] R. Parikh and M. Movassate, "Sentiment Analysis of UserGenerated Twitter Updates using Various Classification Techniques", CS224N Final Report, 2009

[3] A. Go, R. Bhayani, L.Huang. "Twitter Sentiment Classification Using Distant Supervision". Stanford University, Technical Paper ,2009

[4] J. Read. "Using emoticons to reduce dependency in machine learning techniques for sentiment classification".In Proceedings of ACL-05, 43nd Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2005

[5] L. Barbosa, J. Feng. "Robust Sentiment Detection on Twitter from Biased and Noisy Data". COLING 2010: Poster Volume, pp. 36-44.

[6] A. Bifet and E. Frank, "Sentiment Knowledge Discovery in Twitter Streaming Data", In Proceedings of the 13<sup>th</sup> International Conference on Discovery Science, Berlin, Germany: Springer,2010, pp. 1–15.