International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7467
***Article Received**: 20 May 2023 **Revised:** 20 July 2023 **Accepted:** 02 August 2023*

A Machine Learning Model to Identify Fake Data from Social Media using Sentiment Features

Anand R[1], J Alamelu Mangai[2]
[1]Department of Computer Science and Engineering,
BMS Institute of Technology and Management, Yelahanka,
Bangalore-560064, India,
anandor@bmsit.in
[2]Department of Computer Science and Engineering,
Presidency University,
Bangalore-560064, India,
alamelu.jothidurai@presidencyuniversity.in

Abstract— The exponential growth in the use of social media is leading to sharing of information among each other through which the spreading of fake news is common these days. online social networking is the main source for fake news. The most popular social media are Twitter and Facebook, through which the majority of the news reaches the public. This study is aim to try different classification algorithms in comparing with Dataset. For our experiment purpose the dataset used is Real or Fake News dataset which is extracted from Kaggle, which comprises 30Mb of twitter data. The two major classification algorithms used are Naive Bayes and Logistic Regression classification algorithm. The algorithms result in Accuracy score 82.48%, AUC 1.0 and kappa score 0.64 and Accuracy score 91.16%, AUC 0.91 and kappa score 0.82 respectively for the given dataset. The two different classification algorithms are successfully checked with the given dataset. The sentimental analysis is the other way of identification of fake data problem which can be implemented to know the positive and negative sentiment in the given twits. VADER feature is the one of the feature extraction which can be tried with the dataset to find out fake and real data.

Keywords-AUC,RAC,Kaggle,Twitter,Facebook.

I. INTRODUCTION

In our society the fake news is distribution faster than the real news which might be from the social media. The fake news is the biggest threat in our society. In earlier days there was only way is to spread news is newspaper so there was very less percentage of spreading fake news. But now a day the technology has been improved a lot so that the fake news can be spread with in few seconds through the finger tips. We all know that the society is completely depended on the technology so called social media where lots of communication happens so it is the easy way to spread the fake news faster through applications like Facebook, twitter, WhatsApp, Instagram, YouTube and many more other fake news applications. Still these days there is no meaning for Fake news. In our words, news which is intentionally or unintentionally sped through social media to bring impact on one's life, family, reputed companies, nation, culture, religion etc. which is verifiably false.

The news article which is circulating through the social media which is carrying false information or misinformation or disinformation and misleading society. There is a need to find the various forms of false news sources which are spreading in societal media. Many recognition methods available which are time and resource consuming, it also not reliable due to many subjective judgment, etc[4].The machine learning techniques are the more powerful and advanced methods in finding the anti-social behavior in public media.

The paper presents different ML models used to identify fake news from the given twitter text. The main focus is to focus on the title of the twitter message and to decide is it sufficient to decide if the message is fake or real. The paper prepared as follows Segment 2 provides through literature survey of the given problem. Overview of ML classification algorithm used is provided in Segment 3. Segment 4 Describes about implementation and model development with results summarization, Section 5 is about Scope and Future work followed by reference section.

II. LITERATURE SURVEY

It is tough to understand the slang, abbreviations, emotions, acronyms and hyperlinks. [1]. Text preprocessing is the important part of NLP which is used to understand the linguistic barriers. [7]. NLP process the data using natural language with accuracy. [8]. Classification algorithms is a part of NLP; it uses input text known as corpus. [10]. The technique of normalization is for fixing words in data text for to construct accurate sentence as per grammar. [5] Stemming functions are used to reduce the word into their simple form. [6]. Stop word

ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7467
***Article Received**: 20 May 2023 **Revised:** 20 July 2023 **Accepted:** 02 August 2023*

eliminator tool get cleared of expressions that appear regularly [7]. Using sentiment enhances the correctness of classification algorithms.

Different classification procedures are used for the sentimental examination twitters. The specified data was divided in to two portions training and testing.

In sentiment study feature collection plays a vital part which helps in improving the classification algorithm. [8].TF-IDF to find occurrence of term in the article and IDF is for checking the distribution of the term in the document. [9]. Using TF-IDF is used to estimate all weights of all features in the given manuscript and as a end result, distinct out the term with the utmost weight [12].

VADER is a rule based emotion analyzer that has been educated on societal media text. It is very similar to Text Blob which is a python collection for natural language processing, which takes text as an input and returns polarity and objectivity as output. In the present scenario users direct their viewpoints smoothly and data is produced in fraction of seconds. Drawing perceptions from such data is vital for to make efficient decision. All these are some elementary demonstration; new copies can be used with the different datasets.

In the paper the current information is poised i.e identical amount of confident and undesirable twitters are used for classification. This may not be correct in all scenario, prototypical should receipts attention of such cases,also no sentimentality which also need to be addressed.

III. METHODOLOGY

3.1 Dataset

On further to data collection from twitter, the dataset utilized is the Real or Fake News dataset from Kaggle, which comprises 30Mb of data. The '0' indicates negative tweets and '1' indicates positive tweets. Dataset collected is precarious to the prototypical competence thus it is divided into a two parts train and examination set. The exercise set is the main feature using our prototypical is skilled and additional helps in illustration decision.

As it mentioned in Fig:1, the dataset contains equal spreading of positive and negative twitters. Title column represents the title of the tweet and text gives the detailed tweet followed by label.

This dataset is all about Real or Fake News or Text dataset.
Here are only 4 columns.

number:
title:
text:
label:

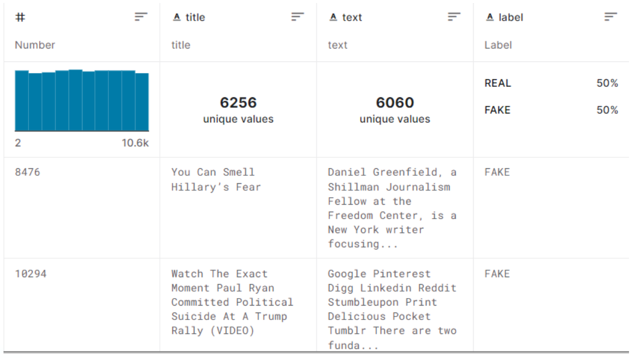

Fig.1. Real or Fake News Dataset [24].

3.2 Preprocessing

Converting raw data in to clean data is very much necessary in the preprocessing stage. The data preprocessed is transformation of the raw data in to clean data before it is given to algorithm for further process [6].

Need for Data Pre-processing

To achieve the good result from any ML model, the data format is very important. Different machine learning models requires the data in specified format.

Another important thing is the data set to be prepared in such that one or more machine learning or deep learning methods are to be allowed to execute on the dataset, and best is selected.

Data preprocessing is the process of making the data more machine understandable. It is a important step in ML. The raw data what we get is changed to machine understandable form by removing all unwanted noise from the data. This lead to extraction of proper feature from data, which results to more accurate results.

List of steps for data pre-processing:

Stage 1: Eliminate duplicate tweets and make it case sensitive and if required change the case also of text: Case-sensitive study differentiates among binary cases of the similar term built on the background of the idiom. It is vital for a capable examination to evade giving the prototypical such glitches.

Stage 2: Discontinue disputes are detached if they take no influence on the tweet's communication meaning (for sample and, or, still, etc.).

Stage 3: Elimination of Tweet-precise features: All explanation terms and hyperlinks take stood reformed to basic labels or have been removed entirely.

Stage 4: Elimination of distinct fonts and symbols: By removing distinct fonts can assistance in joining binary relations that were before believed to be dissimilar.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7467
***Article Received**: 20 May 2023 **Revised:** 20 July 2023 **Accepted:** 02 August 2023*

Stage 5: Generate a monolingual dictionary to eliminate needless relations and punctuation characters from information.

Stage 6: Jargon and acronyms are being prolonged, along with modification of spellings.

Stage 7: Tag a part of speech: It denotes to the procedure of allocating a label to each period in the contribution

script and classifying it as nouns, verbs, adjectives, and so on. Taggers are capable at recognizing unambiguous characters.

3.2.1 Information Quality Assessment

The data quality is very important for any ML model. Poor quality of data leads to poor performance of the model. The amount of success rate in any of machine learning method is purely based on quality of data used.

3.2.2 Train/Validation/Test Split Based Data

To test the machine learning model three types of data set used three sets. Example of information used to fine-tune the parameter is called training. Set of data used to fine-tune the prototypical is called validation set. Test set is to check the presentation of model. In majority of studies for false update discovery have separated their dataset into train, validate and test, in some trainings have used only the train, and test sets [13], [14]. The percentages of information divided 60:20:20, 70:30, and 80:20 are precise shared in false update discovery.

3.2.3 Tokenization, Stemming and Lemmatization

Tokenization is a common technique which breaks the text in to words. Performing tokenization is very common in text data processing. The removal of derivational suffixes is a part of stemming; by this we can obtain other word. Lemmatization is a normalization process used for text, it generates root from inflated words [15]. Some sample inflection finishes are s, bat, and bats. Stemming and stop-word removal process is time consuming, but it causes the small difference in the results. Not all researchers may use all three techniques. The improvement in results may be due to adopting some additional pre-processing, such as stemming and stop words removal.

3.3 Term Vectorising

Term vectorization is charting of words to slant of paths.TF-IDF and Bag of Words(BoW) methods are very frequently used in false news detection [16]. In TF-IDF, the value upsurges uniformly amount of periods the word appears in the manuscript, and stable by occurrence of the word in the body. The semantic sagacity of the terms is gone in its challenge to interpret it to vectors [17]. The BoW is to compute the regularity tally of each term within the manuscript, which is used to harvest a numeric demonstration of the term. The disadvantage of this process is contextual information is lost. The latest models like (GloVe) and Word2vec are used on fake news discovery which are pre-trained models. The advantage of this prototypes is their capability to train with huge dataset [18]. Table gives the summary of the NLP methods and term path prototypes used in deep knowledge.

Table 1: Advantage and Disadvantage of Word Vector Models.

Method	**Advantage**	**Disadvantage**
TF-IDF	Information about more significant and less significant words.	Slow for larger dataset, semantics, co-occurrences in different documents.
Bag-of-Words	Ease of implementation	Ignores ordering and semantic relations among words
Word2Vec	The background data preserved, the size of the implanting path is very minor.	Not very efficient with unfamiliar words. Sub-words not represented.
Doc2Vec	Faster than Word2Vec,Regardless of its length.	Not good for short documents
GloVe	Unlike other methods, it does not rely on native indicators.	Global statistics are used
BERT	Identify and capture contextual meaning in a sentence	Compute intensive at inference time.

3.4 Feature Extraction

Extraction of feature and selection are the normally used in text mining [19], [20]. Fake news discovery focuses on social context features [21]. Text features comprises the writing style and emotion. The spread network contains rich information like comments, responses and tweets that show the way of data movement, it also gives information about user profile and interaction. It is very vital to select correct feature extraction algorithm because the feature reduction contains an incredible effect on the text classification results. Some of the common algorithms are Term Frequency-Inverse Document Frequency(TF-IDF), Information gain (IG), Principal Component Analysis(PCA) and Chi-Square Statistics(CHI). With the feature extraction the accomplishment percentage is greater.

Formula for Term Frequency as follows:

$$tf(t,d) = \frac{f_{t,d}}{\sum_{t\epsilon d} f_{t,d}} \quad (1)$$

Where t is the number of times term in the document d.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7467
***Article Received**: 20 May 2023 **Revised:** 20 July 2023 **Accepted:** 02 August 2023*

Formula for Inverse Term Frequency as follows:

$$idf(t, D) = \log \frac{N}{|\{d\epsilon D: t\epsilon d\}|} \tag{2}$$

Where *t* is the number of times term appears in document *D*, and *N* total number of documents.

The tf-idf calculated as follows:

$$\text{Tfidf}(t,d,D)=\text{tf}(t,\text{idf}(t,D)) \tag{3}$$

One more important feature can be considered is emotions of the people. Fake data can also contain people's emotions [22]. Although they are totally different, by analyzing sentiment could improve the Fake News detection. To achieve this sentiment analyzer required, VADER(Valence Aware Dictionary and Sentiment Reasoner) is one of the tool[23].It is publicly and performs better than other tools like LIWC,GI,WordNet,and SentiWordNet.

The Scores given by VADER are as follows: -

1.How negative the tone of a text.

2.How positive the text tone is.

3.How neutral tone is.

4.How compound/mixed it is with compare to other values.

The value range from -1 to 1, if some classifiers which are not working with negative numbers it is changed 0 to 2 range by adding 1 for it.

3.5 Classifiers

3.5.1 Logistic Regression

Because we classify of the texts based on a broad variety of characteristic, a outcome (true or false) is used because it offers an easy approach to divide issues into a basic class or many classes, which is useful when dealing with complex situations. We adjusted hidden layers to get the finest outcome for every unique datasets and assessed many values before achieving the greatest accuracy in the LR model, which was then used to train the LR model.

Following is a mathematical equation 1 description of how the logistic regression hypothesis function calculated.

$$h_{\emptyset}(x) = \frac{1}{1+e^{-(\beta_0+\beta_1 X)}} \tag{4}$$

Using a sigmoid function, logistic regression can convert the output to its likelihood value; the aim is to decrease the cost function as much as possible to get the highest possible probability. As previously shown, the cost function may be calculated by using equation 2.

$$Cost(h_{\emptyset}(x), y) = \begin{cases} \log\left(h_{\emptyset}(x)\right), & y = 1, \\ -\log\left(1 - h_{\emptyset}(x)\right), & y = 0. \end{cases} \tag{5}$$

Following is the pseudocode for Logistic Regression algorithm:

Input: Training Data

Begin

1.For i=1 to k

2.For each training data instance di.

3.Set the target value for the regression to

$$Zi = \frac{y_i - P(1|d_j)}{[P(1|d_j)(1-P(1|d_j))]}$$

4.Weight of instance initialization dj to

$$[P(1|d_j)\left(1 - P(1|d_j)\right)]$$

5.Finalize a f(j) to the data with class value (Zj) and weight (wj)

Classical label decision

6.Assign (class label:1)if Pid >0.5 ,otherwise (class label:2)

End

3.5.2 Naïve Bayes

Naive Bayes classification algorithm are a set of managed learning processes grounded on smearing Bayes' theorem with the "naive" supposition of qualified individuality among each duo of features given the value of the class variable.

By thinking about the same, not unusual to place properties among junk mail messages (i.e., an inappropriate message which can be warned uninvited) & information items which aren't true, like

• Spelling errors

• Manipulation of opinion on a few things that influences the reader's opinion is frequently construed.

• Resemblance of a limited collection of phrases used as junk mail messages has similarity when compared to specific junk mail messages from a syntactic standpoint. False info pages and spam email communications show the same linguistic similarities when searched from a grammatical standpoint.

• Info each false information & junk mail messages aren't trusted.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7467
***Article Received**: 20 May 2023 **Revised:** 20 July 2023 **Accepted:** 02 August 2023*

Mathematically described:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)} \qquad (6)$$

Following is the pseudocode for Naïve Bayes algorithm:

Input:

Training dataset T,

F=(f1,f2,f3,…..fn) values of the changeable in testing dataset.

Output:

A class of challenging dataset.

Step:

1.Read the training dataset T;

2.Calculate the mean and standard deviation of the predictor variable in each class;

3.Repeat.

Calculate the possibility of fi using the gauss density equation in each class;

Until the possibility of all predictor variables (f1,f2,f3….fn)has been calculated.

4.Calculate the same for each class;

3.6 Performance Measure

The main objective to measure the enactment of machine learning classification models on the given dataset. Model has been trained on entire text from different twitter data and then evaluated. The model trained for to notice the fake data by using the title of the articles. The research was accompanied on a PC equipped with a Intel Core i7 with 3.60GHz with 4.00 GB RAM.

The performance is calculated by following commonly used metrics used in classification task.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (7)$$

$$Precision = \frac{TP}{TP+FP} \qquad (8)$$

$$Recall = \frac{TP}{TP+FN} \qquad (9)$$

$$F1\ score = 2 * \frac{Precision*Recall}{Precision+Recall} \qquad (10)$$

where:

- TP (True Positive): when predicted fake news as fake news,
- TN (True Negative): when predicted true news as true news,
- FN (False Negative): when predicted true news as fake news,
- FP (False Positive): when predicted fake news as true news.

IV. RESULTS AND DISCUSSIONS

For the fake news detection, our focus is to find out recall of the model, as to detect most of the positive examples. To reduce False negative percentage, the F1 metric could be beneficial. It is the blend of both precision and recall metrics. The model efficiency also unhurried and also Cohen-Kappa score calculated. Table 3 reviews the outcomes obtained on the dataset holding twitter news. When associating the classification algorithm outcomes on label text considered and we estimate the similar performance for the twitter news body.

Table 2: Sum up the performance of the models for TF-IDF feature.

	Logistic Regression	Naive Bayes
Accuracy score	**0.91**	**0.82**
Precision score	**0.93**	**0.82**
Recall score	**0.89**	**0.85**
F1 Score	**0.91**	**0.83**
Cohen-Kappa Score	**0.82**	**0.64**

A like regular of trials was done on the heading manuscripts. Table 2 sum up the presentation of the classification models**.** As we can see from the outcomes, logistic regression model attained very decent performance. LR model succeeded to identify good amount of false newscast trainings in the test set, while still retained the FP ratio at a sensible rate. Figure 2 and Figure 3 shows the more detailed presentation of the model using a confusion matrix. The matrix for both classification algorithms were listed. It shows, that all optimistic examples (fake news articles) were appropriately categorized by the prototypical. The AUC score: 0.911 and 1.0 is recorded for logistic regression and naive bayes classification algorithm.

Confusion matrix:

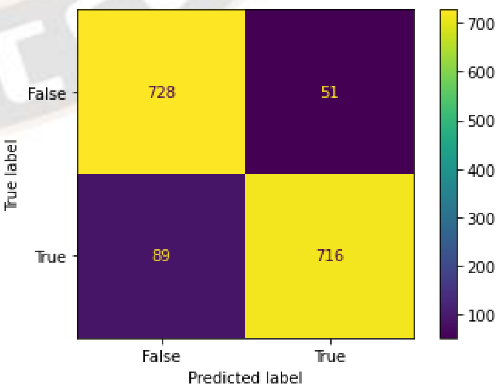

Figure 2: Confusion Matrix Logistic Regression

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7467
***Article Received**: 20 May 2023 **Revised:** 20 July 2023 **Accepted:** 02 August 2023*

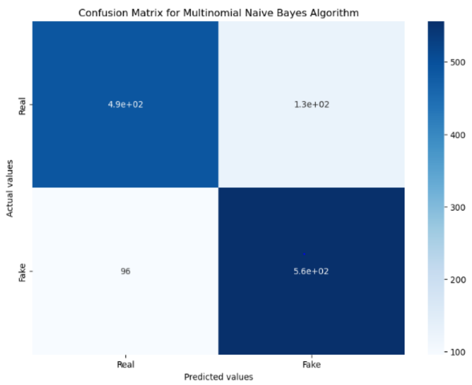

Figure 3: Confusion Matrix Naive Bayes

The results of TF-IDF and VADER feature are listed as shown in the table 3. The table contains all the values which is generated for the given dataset. The accuracy, precision, recall and F1 for the given data is been calculated. The values in the table is for title of the twitter dataset. The summarized performance of both classification algorithms are tabulated with adding VADER feature. The observation is that Logistic Regression algorithm gives more accuracy result compare to Naïve Bayes classification algorithm.

Table 3: Summarizes the performance of the models for TF-IDF and TF-IDF+VADER feature for Title of the dataset.

Classifier	Feature Extraction	Accuracy	Precision	Recall	F1-Score
Naïve Bayes	TF-IDF	**0.68**	**0.67**	**0.77**	**0.716**
	TF-IDF + VADER	**0.68**	**0.67**	**0.77**	**0.716**
Logistic Regression	TF-IDF	**0.8256**	**0.85**	**0.80**	**0.824**
	TF-IDF + VADER	**0.8295**	**0.86**	**0.81**	**0.838**

The results what we got for the TF-IDF and VADER feature is listed in the table 4. The table gives the result summary of the performance for the text data in the Twitter. The dataset twitter text is processed and the accuracy, precision, recall and F1 score calculated for the text. As the observation from table Logistic Regression give more precise results compare to NB when tested with text of the dataset.

Table 4: Summarizes the performance of the models for TF-IDF and TF-IDF+VADER feature for Text of the dataset.

Classifier	Feature Extraction	Accuracy	Precision	Recall	F1-Score
Naïve Bayes	TF-IDF	**0.82**	**0.82**	**0.85**	**0.834**
	TF-IDF + VADER	**0.82**	**0.82**	**0.85**	**0.834**
Logistic Regression	TF-IDF	**0.91**	**0.94**	**0.90**	**0.919**
	TF-IDF + VADER	**0.91**	**0.94**	**0.91**	**0.924**

The table 5. Is the result for whole dataset which we got after checking the performance of both the classification algorithms with two feature The table value clearly shows the different between the performance when it checked with the dataset. The feature tried with algorithms are TF-IDF and combining VADER with TF-IDF.

When VADER combined with TF-IDF the accuracy is 0.77 and F1 score is 0.768 which is slightly low compare to logistic regression. The Dis-advantage of Naive Bayes is that it only uses a small number of training data to estimate the parameters necessary for classification. If the test data set has a unconditional variable of a classification that wasn't existing in the training data set, the Naive Bayes model will assign it zero possibility and won't be able to make any extrapolations in this respect.

Table 5: Summarizes the performance of the models for TF-IDF and TF-IDF+VADER feature for the dataset.

Classifier	Feature Extraction	Accuracy	Precision	Recall	F1-Score
Naïve Bayes	TF-IDF	**0.82**	**0.82**	**0.85**	**0.834**
	TF-IDF + VADER	**0.77**	**0.80**	**0.74**	**0.768**
Logistic Regression	TF-IDF	**0.91**	**0.94**	**0.90**	**0.919**
	TF-IDF + VADER	**0.92**	**0.93**	**0.92**	**0.924**

V. CONCLUSION

The current work is about the comparing the two different classification algorithms. The dataset used for the both algorithms are same. The dataset has checked for the performance, AUC and ROC graphs. The dataset is taken for the Kaggle repository and its of size 30Mb.The dataset is consisting of different tweets which is extracted from twitter social media. The two (Naive Bayes and Logistic Regression) classification algorithms compared and got Accuracy Score 82.48 and 91.16. The kappa score is 0.64 and 0.82 respectively. The sentiment VADER feature also combined with TF-IDF and checked for the performance, the accuracy of 0.77 and 0.92 resulted for NB and LR classifiers. In the current study the dataset is balanced, which contain both positive and negative data, the work can also be extended to the dataset which contain the neutral meaning for the data. Through sentiment analysis the optimistic and destructive precision is measured for different ML algorithms for a given data.

REFERENCES

[1] W. Medhat , Hassan A, Korashy H (2014) Sentiment analysis algorithms and applications: a survey. Ain Shams Eng J. 5:1093–1113.

[2] A. Abbasi, et al., "Selecting Attributes for Sentiment Classification Using Feature Relation Networks," IEEE

Transactions on Knowledge and Data Engineering, vol. 23, pp. 447-462, 2011.

[3] L. Mandloi and R. Patel, "Twitter Sentiments Analysis Using Machine Learninig Methods," 2020 International Conference for Emerging Technology(INCET), 2020, pp. 1-5, doi: 10.1109/INCET49848.2020.9154183

[4] R. H. H. Aziz and N. Dimililer, "Twitter Sentiment Analysis using an Ensemble Weighted Majority Vote Classifier," 2020 International Conference on Advanced Science and Engineering (ICOASE), 2020, pp. 103-109, doi: 10.1109/ICOASE51841.2020.9436590

[5] C. Catal, U. Sevim, and B. Diri, "Practical Development of an Eclipse-Based Software Fault Prediction Tool Using Naive Bayes Algorithm," Expert Systems with Applications, vol. 38, pp. 2347- 2353, 2011

[6] Prof. Bhushan Thakre, Dr. R.M Thakre. (2017). Analysis of Modified Current Controller and its Implementation in Automotive LED. International Journal of New Practices in Management and Engineering, 6(04), 01 - 06. https://doi.org/10.17762/ijnpme.v6i04.60

[7] Sampurnanand Dwivedi, Vipul Singhal. (2023). A Study of Responsive Image Denoising Algorithm . International Journal of Intelligent Systems and Applications in Engineering, 11(3s), 286–291. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2686

[8] M. Mhatre, D. Phondekar, P. Kadam, A. Chawathe, and K. Ghag, "Dimensionality reduction for sentiment analysis using pre-processing techniques," 2017 International Conference on Computing Methodologies and Communication (ICCMC), pp. 16-21, 2017.

[9] G. Chowdhary, "Natural Language Processing," Annual Reviewof Information Science and Technology, vol. 37, no. 1, pp. 51-89 2020.

[10] A. Blom and S. Thorsen, "Automatic Twitter replies with Python," International conference "Dialog 2012".

[11] A. N. Rohman, E. Utami, and S. Raharjo, "Deteksi Kondisi Emosi pada Media Sosial Menggunakan Pendekatan Leksikon dan Natural Language Processing," Eksplora Inform., vol. 9, no. 1, pp. 70-76, 2019

[12] I. Riadi, S. Sunardi, and P. Widiandana, "Mobile Forensics for Cyberbullying Detection using Term Frequency - Inverse Document Frequency (TF-IDF)," J. Ilm. Tek. Elektro Komput. dan Inform., vol. 5, pp. 68-76, 2019.

[13] A. Abbasi, et al., "Selecting Attributes for Sentiment Classification Using Feature Relation Networks," IEEE Transactions on Knowledge and Data Engineering, vol. 23, pp. 447-462,2011.

[14] Peiman Barnaghi, John G. Breslin and Parsa Ghaffari, "Opinion Mining and Sentiment Polarity on Twitter and Correlation between Events and Sentiment", 2016 IEEE Second International Conference on Big Data Computing Service and Applications.

[15] A. Thota, P. Tilak, S. Ahluwalia, and N. Lohia, ''Fake news detection: A deep learning approach,'' SMU Data Sci. Rev., vol. 1, no. 3, p. 10, 2018.

[16] Z. Jin, J. Cao, H. Guo, Y. Zhang, and J. Luo, ''Multimodal fusion with recurrent neural networks for rumor detection on microblogs,'' in Proc. 25th ACM Int. Conf. Multimedia, Oct. 2017, pp. 795–816.

[17] S. M. Padnekar, G. S. Kumar, and P. Deepak, ''BiLSTM-autoencoder architecture for stance prediction,'' in Proc. Int. Conf. Data Sci. Eng. (ICDSE), Dec. 2020, pp. 1–5.

[18] Abdullah-All-Tanvir, E. M. Mahir, S. Akhter, and M. R. Huq, ''Detecting fake news using machine learning and deep learning algorithms,'' in Proc. 7th Int. Conf. Smart Comput. Commun. (ICSCC), Jun. 2019, pp. 1–5.

[19] Auma, G., Goldberg, R., Oliveira, A., Seo-joon, C., & Nakamura, E. Enhancing Sentiment Analysis Using Transfer Learning Techniques. Kuwait Journal of Machine Learning, 1(3). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/129

[20] R. R. Mandical, N. Mamatha, N. Shivakumar, R. Monica, and A. N. Krishna, ''Identification of fake news using machine learning,'' in Proc. IEEE Int. Conf. Electron., Comput. Commun. Technol. (CONECCT), Jul. 2020, pp. 1–6

[21] R. K. Kaliyar, A. Goswami, P. Narang, and S. Sinha, ''FNDNet— A deep convolutional neural network for fake news detection,'' Cognit. Syst. Res., vol. 61, pp. 32–44, Jun. 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389041720300085

[22] F. A. Ozbay and B. Alatas, ''Fake news detection within online social media using supervised artificial intelligence algorithms,'' Phys. A, Stat. Mech. Appl., vol. 540, Feb. 2020, Art. no. 123174.

[23] H. Reddy, N. Raj, M. Gala, and A. Basava, ''Text-mining-based fake news detection using ensemble methods,'' Int. J. Autom. Comput., vol. 17, pp. 1–12, Apr. 2020.

[24] K. Shu, S. Wang, and H. Liu, ''Understanding user profiles on social media for fake news detection,'' in Proc. IEEE Conf. Multimedia Inf. Process. Retr. (MIPR), Apr. 2018, pp. 430–435.

[25] V. Bakir and A. McStay, "Fake News and The Economy of Emotions," Digit. Journal., vol. 6, no. 2, pp. 154–175, 2017.

[26] C. J. C. J. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," in ICWSM, 2014, no. May, pp. 216–225.

[27] https://www.kaggle.com/code/yogeshpm/fake-or-real-news/input.