



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

EMOTION TWEETS CLASSIFICATION USING MACHINE LEARNING

K. b.v.s.prasad¹, s.adinarayana M.E(CSE),PHD Professor², M. bhargavi³, K.siva teja reddy⁴, K. Pawan⁵

¹Department of Computer Science and Engineering, Raghu Institute Of Technology, Visakhapatnam, AP, India

Abstract - Twitter is a Social verbal exchange Platform which permits customers to label their tweets in extraordinary codecs public can effortlessly explicit their perspectives and thoughts on a extensive form of cluster on subjects thru social networking websites. Sentiment evaluation is a Study comes beneath Natural Language Processing It allows to discover the real Sentiment hidden in a text. In these studies, we attention on Finding the Sentiments for Twitter statistics set. It has greater exacting because of its unstructured statistics, constrained size, use of slangs, misspells, emojis, abbreviations, prevent words. Most of the researchers dealt with diverse sorts of methods of sentiment evaluation and evaluate their results the use of diverse methods had been underexplored inside the literature Hence, in our studies we strive to carry out sentiment evaluation on twitter statistics the use of herbal language processing. Here we are taking two data sets for testing and training. In train data set we have label means two values 0 and 1. It is to indicate racist and non-racist tweets. In this research we are making the data as a structured manner with nlp techniques like tokenization, stemming, lemmatize etc and we have used svm classifier for prediction of the data and we need to find the confusion matrix and accuracy. Compares its performance with other well-known classifiers. Comparison results shows that SVM classifier better predicts the outcome.

Keywords: Twitter, Natural Language Processing, Naïve Bayes classifier support vector machines, random forest classifier, machine learning Techniques

1. Introduction

Internet is the only the arena is evolving in a special guidelines and special approaches on their perceptions in recent times most of these are completed thru online classes, evaluation web sites many humans depend on consumer generated context on any product to a extremely good quantity with regards to carry out any favored action. When humans need to Purchase a product thru Internet, they'll search for opinions if it changed into top they had been inquisitive about that unique product internet site thru on line, earlier than making up a decision. Some evaluation is to be completed on the ones opinions so the very last end result says whether or not the product is purchase or not. There are wonderful sentiment evaluation strategies which are to be had with many apps for exceptional domains, like in enterprise to get comments for merchandise from customers. Machine mastering and Knowledge base strategies are strategies which can be specifically used for sentiment analysis. In the case of Knowledge base technique this calls for a huge database with prebuilt feelings and an green and powerful information illustration for figuring out sentiments. In the case of Machine mastering technique doesn't require any prebuilt set of feelings uses a education set so that it will evolve a sentiment classifier which classifies sentiments in tweets and there through system mastering technique is greater less complicated than knowledgebase technique.

Classification is a way used to carry out the class on wonderful records units into wonderful classes, these strategies are essentially divided into classes unsupervised and supervised mastering and allow us to take an instance in supervised mastering technique there can be a trainer which makes the laptop to examine from the labeled records this is given to that these records set is within the shape of bi-magnificence i.e, figuring out climate the announcement is bad or fine is indicated with the aid of using 1 and 0. In unsupervised mastering laptop learns with the unlabelled records and that is utilized in grouping the records as an instance in cluster analysis.

2. Literature Survey

Go and L. Huang (2009) [1] proposed a sentiment analysis solution for Twitter data using remote monitoring. In this solution, the training data consists of tweets containing pictograms that act as labels. They build the model using Naive Bayes, MaxEnt, and Support Vector Machine (SVM). Neethu and RajasreeR [2] performed a tweet sentiment assessment based primarily on the use of key system control strategies, primarily based on a few specific domains. To handle typos and slang, a function vector is generated using nb, SVM, most entropy classifiers and sets for accuracy checking. Huy Nguyen and Bac Le [3] created a version of Twitter's sentiment survey on system use. Learn strategies by applying feature set to add precision bigrams, unigrams, object-oriented functionality, and more. Sayali P. Nazare, Prasad S. Nar, Akshay S. Phate, Ingle [4] generated a data set via Twitter API key and accumulated all tweets related to the Blue Whale game theme, Vector machine support, maximum entropy classifier and set the SVM and Nb classifiers are applied when using MATLAB's built-in functions. Mejova [5] considering the evaluate of the sentiment evaluation i.e., they given a clean description approximately the phrase sentiment evaluation and novelty techniques in clean way which are used to carry out evaluation on emoticoncharacter textual content via way of means of overlaying all of the difficulties which are confronted all through the evaluation phase of product or textual content Boiy, Erik, Hens, Pieter, Deschacht, Koen & Moens, MarieFrancine, MarieFrancine [6] underline the usefulness of the classifier and summarize some methods of evaluation through the method of calculating the precision of each method, with promising results and The difficult situation we face when accumulating information from the World Wide Web and Niu, Zhen, Zelong Yin, Xiangyu Kong [7] reproduce the type of textual content that you experience when using the internet offers. We have developed three algorithms, but more efficient advancements for overall efficiency decoration. It also provides new technology based primarily on shipbuilding sorters.

Dey, Lopamudra&Chakraborty [10] have amassed 2 units of dataset they may be film critiques and inn critiques via way of means of the use of classifiers naïve Bayes and KNN and Their aim is to test which classifier offers higher consequences . The experimental consequences suggests that the NB classifier offers respectable overall performance withinside the case of film critiques dataset and on thinking about inn critiques dataset each classifiers suggests approximate consequences. Finally, NB classifier is higher for film critiques type. Yelena , Leena A Deshpande, M.R. Narasinga Rao [11] is developed the importancvariances in the sematic analysis and the information is retrained the version conqer to the glide this is identified. They have used 2 strategies weight-primarily based totally capabilities and nGram that identifies the unnamed labels withinside the textual content which improves its accuracy and Pak and Paroubek(2010) [8] proposed a model to classify the tweets as objective, positive and negative By collecting tweets using Twitter API they have created a twitter corpus Taboada is an LB method [9] uses sentiment analysis with opinion words and match them with the data to determine polarity. They assigns sentiment scores to the opinion words describing how Positive, Negative and Objective the words contained in the dictionary

T. Sajana, M.R.Narasingarao[12] they have got completed survey at the detection of sickness after which the prediction of malaria sickness the use of diverse system mastering strategies, Image Processing strategies and They have found that system mastering strategies are frequently relevant for crucial prognosis of malaria. F. Neri, C. Aliprandi, F. Capeci, Cuadros and audience, correlation evaluation of Facebook with measurable information that is to be had in public. Kavya Suppala Narasinga Rao[13] evolved a supervised set of rules this is primarily based totally at the probabilistic technique to categorise the textual content to a selected elegance i.e wonderful or poor. G. Vinodhini and R.Chandrasekaran[14] completed sentimental evaluation and opinion mining the use of strategies like Naïve Bayes,Maximum entropy, and Support vector system to categorise critiques. Abdullah alseedi and Mohammed zubair[15] khan had been proposed social media sentimental evaluation the use of system mastering strategies like SVM and Bayesian algorithms and approaches. Lin and Kolcz [16] proposed incorporating more than one classifier into large-scale twitter information. They tried to educate logistic regression (LR) classifiers from the hashed 4-grams as capabilities. The schooling dataset various from one to a hundred million of examples with ensembles of three to fortyone classifiers. The best overall performance become obtained when the quantity of classifiers become 21 and the quantity of times become a hundred million, attaining a type accuracy of 0.81. Sharma et all. [17] were proposed an unsupervised document-primarily based totally-sentiment evaluation system able to determine the sentiment orientation of textual content documents primarily based totally on their polarities and subjectivities .

Kahn et all. [18] Propose a Twitter Poll Framework (TOM) for emotional tweets. The proposed combination scheme includes SentiWordNet ratings, emoji ratings, and an improved polar classifier. It is not just a perception technique. The test was performed using six data sets that tested the proposed rule set to perform 83 mean harmonic implications.3% 0.

PoWei Liang et al. [19] We designed a framework called "Minor Opinion" that automatically investigates and detects the emotions of social media posts. tweets with subtitles were shuffled for the classification task, and in this framework, messages with emotion were extracted (tweets without comments were removed) and polarity was determined (that is, is it good or bad) To achieve this, the evaluator tweeted "opinion" and "no opinion" using the NB classifier with a single histogram and the results of test confirmed the relevance of the framework for emotional assessment in real microblogging applications [20]. Awais Muhammad and Samin Tanzila (2012) argue that the use of the Internet has made sand a global village. The use of the Internet has shortened the distance and brought people closer together.

Chou and Chou (2000) [21] observed that banks across the arena are seeing a large market for banking potential due to the phenomenal explosion of digital commerce. Financial institutions want to design and implement robust networks to provide green services to their customers. Cao et al [22] (2015) propose that social media has the capability to construct employees' social capital to definitely impact their information integration.

3. Methodology Framework

Sentiment analysis is a machine learning tool that analyzes text. By training machines we can get the emotions in text, machines using the tools automatically learn how to detect sentiment without human input. The proposed framework is shown from the Figure 1

Machine learning steps are:

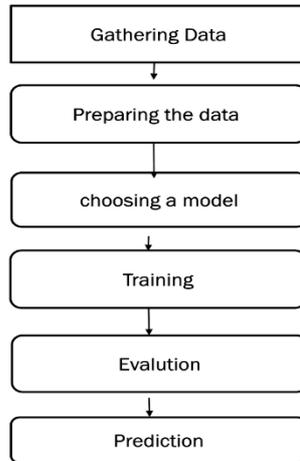


Figure 1 Machine learning Framework

The First Stage of this Framework is Gathering Data. Here data Gathering is the first step in machine learning and it is the foundation of machine learning. Here we gather the data base of our requirement for example from companies we can gather data like from banks, insurance data, showrooms data etc. Some top most companies provide API'S (Application programming interface we can collect data from api's also. And from webscraping we can gather data .web scraping is an automatic method to collect huge amounts of data from websites and we store it as csv file and we use it in our model.

Table 1 Features the first 6 rows in the Twitter data set

Sl.NO	Id	Label	tweet
0	1	0	@user when a father is dysfunctional and is so selfish he drags his kids into his dysfunction. #run
1	2	0	@user @user thanks for #lyft credit i can't use cause they don't offer wheelchair vans in pdx. #disapointed #getthanked
2	3	0	bihday your majesty
3	4	0	#model i love u take with u all the time in urđ±!!! đđđđđ!đ!đ!
4	5	0	factsguide: society now #motivation
5	6	1	[2/2] huge fan fare and big talking before they leave. #allshowandnogo

The second stage of the framework is to Preparing the Data from the Tabel 1 we have raw data to convert raw data to a meaningful data we perform data analysis, data analysis means replacing null values, spaces, symbols with meaningful values This is called

as Exploratory Data Analysis or data cleaning. After data cleaning we perform Data visualization means representing data in form of graphs like heapmaps, barplot, countplot etc.

Table 2 Features the Exploratory Data Analysis on the data set

Sl.NO	Id	Label	tweet
0	1	0	when a father is dysfunctional and is so selfish he drags his kids into his dysfunction. #run
1	2	0	thanks for #lyft credit i can't use cause they don't offer wheelchair vans in pdx. #disapointed #getthanked
2	3	0	bihday your majesty
3	4	0	#model i love u take with u all the time in ur
4	5	0	factsguide: society now #motivation
5	6	1	huge fan fare and big talking before they leave. #allshowandnogo

Here in the Table 2 we eliminated the null values and punctuations and emojis by using the numpy library in python. The next stage is the Choosing a model. We need to select a model once we done with the data centric steps. There are different models which are developed by data scientists but in our model we have used SVM(support vector machines), Random forest and Naïve Bayes .

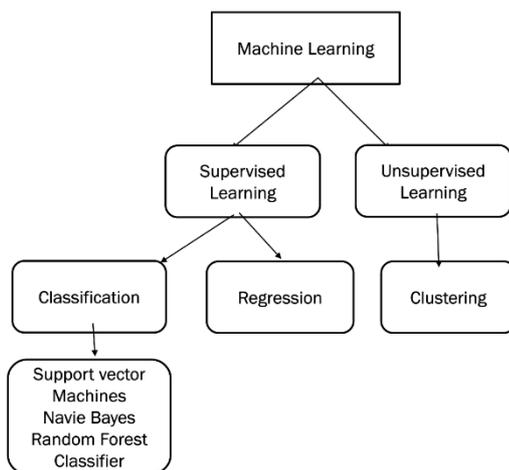


Figure 2 Features the Selection Of the classification Model

SVM (Support vector Machine): SVM is one of the supervised system gaining knowledge of set of rules. It is used for class and regression demanding situations however broadly speaking used for class problems. In this SVM set of rules we plot every statistics object as a factor in n-dimensional area with the fee of every characteristic being the fee of a specific coordinate. Then we carry out class through locating the hyper aircraft that differentiates the 2 instructions very well.

Random Forest: It is likewise one of the supervised system gaining knowledge of set of rules used for each class and regression. This set of rules creates choice bushes on statistics samples and after receives the prediction from every of them and subsequently selects pleasant answer through the usage of voting.

Naïve Bayes: Naïve bayes is the supervised system gaining knowledge of set of rules it's miles primarily based totally at the bayes theorem. Naive bayes classifier is one of the easy and handiest class algorithms which allows in constructing the short system gaining knowledge of fashions that makes short predictions.

After selection We need to Train the Model. In the training model we select the model and train the data and we use fit() and predict() to the data. Here in this stage bulk learning is done

Model Evaluation Techniques These are the performance measures of the model. After training a model we have to check for the performance measures . ME techniques are:

Confusion matrix: It is $N \times N$ matrix used for evaluating the performance of classification model here N is no of target classes. The matrix compares the actual target values with those predicted by machine learning model.

Precision score: Precision quantifies the no of positive class predictions that actually belong to positive class.

F1 score: It provides a single score that balances both the concerns of precisions and recall in one number.

Recall: It quantifies the number of positive class predictions made out of all positive examples in dataset.

The last Stage is the Prediction of our model. It predicts the outcome of the model whether the words are positive or negative.

4. Tweet_Classify_Algorithm

Consider a training data set D consists of documents which belongs to different classes say class P and Q .

1. Calculate prior probability of classes P and Q as

Class P = number of objects of class P / total number of objects.

Class Q = number of objects of class Q / total number of objects.

2. Now calculate the total number of word frequencies of both classes P and Q i.e., n_i as

n_a = the total number of word frequency of class P .

n_b = the total number of word frequency of class Q .

3. Calculate the conditional probability of keyword occurrence for given class $P_b(\text{word1} / \text{class } P) = \text{wordcount} / n_i(P)$
4. Uniform distributions are to be performed in order to avoid zero frequency
5. Now a new document M is classified based on the probability for both classes P and Q $P_b(M/W)$.

a) Find $P_b(P / W) = P_b(P) * P_b(\text{word1}/\text{class } P) * P_b(\text{word2}/ \text{class } P) \dots * P_b(\text{wordn} / \text{class } P)$.

b) Find $P_b(Q / W) = P_b(Q) * P_b(\text{word1}/\text{class } Q) * P_b(\text{word2}/ \text{class } Q) \dots * P_b(\text{wordn} / \text{class } Q)$.

6. After calculating probability for both classes P and Q , the class with higher probability is the one the new document M assigned.

5. Implementation

5.1 In the Proposed work we need to figure out how a sentiment is extracted from a tweet/text in the twitter data set .In our work we have taken two data sets train and test data sets which will useful to train and test the data .Our aim is to preprocess the data and then analyze the data and check the accuracy of the system.

To implement this work we have used Google cola notebook. We used specific libraries to implement the model the libraries are pandas, numpy, re, seaborn , matplotlib etc. Pandas is a software program library written for the Python programming language for records manipulation and analysis. In particular, it gives records systems and operations for manipulating numerical tables and time series. NumPy is an open-supply numerical Python library, NumPy incorporates a multi-dimensional array and matrix records systems. It may be utilized to carry out some of mathematical operations on arrays which includes trigonometric, statistical, and algebraic routines. Seaborn is a library in Python used for Preparing graphics, Seaborn is a records visualization library constructed on pinnacle of matplotlib and carefully included with pandas records systems in Python. Visualization is the principal a part of Seaborn which allows in exploration and knowledge of records.

5.2 Data Collection

From the Kaggle we have taken the data sets. In this paper, the sentiment analysis is done on Twitter data. The two datasets collected here contains 49159 tweets these tweets are collected based on the situation on all topics. We can clearly observe in the table 3 There are different attributes in the database such as item-id, label, tweets, but sentiment text has been considered for our proposed research in training dataset

Table 3 Features the first 10 rows in the Twitter data set

Sl.NO	Id	Label	tweet
0	1	0	@user when a father is dysfunctional and is so selfish he drags his kids into his dysfunction. #run
1	2	0	@user @user thanks for #lyft credit i can't use cause they don't offer wheelchair vans in pdx. #disapointed #getthanked
2	3	0	bihday your majesty
3	4	0	#model i love u take with u all the time in urđ±!!! đđđđđ!đ!đ!
4	5	0	factsguide: society now #motivation
5	6	1	[2/2] huge fan fare and big talking before they leave. #allshowandnogo
6	7	0	@user camping tomorrow @user @user @user @user @user @user @user dannyâ!
7	8	1	the next school year is the year for exams.đ¯ can't think about that đ #school #exams #hate #imagine #actorslife #revolutionschool #girl

8	9	0	we won!!! love the land!!! #allin #cavs #champions #cleveland #clevelandcavaliers â!
9	10	0	@user @user welcome here ! i'm it's so #gr8 !

5.3 Labelling on Training Data

The first attribute item-id contains the id of the tweet, the second attribute label represents the boolean value (1 or 0) i.e., the sentiment tweet is taken as 0 and tweet without any sentiment is declared as 1, and the third attribute is sentiment text represents the text or tweet based on all situations either containing sentiment or not In test dataset there are two fields id and the text/tweets. From the Figure 3 the blue indicates test tweets and yellow indicates train tweets. We can see length of train and test data sets from the figure 3. we need to count the length by using str.len() for both the data sets then we have plotted the histogram using matplotlib library and we are using 20 bins to plot the histogram and using plt.show() will show the whole histogram.

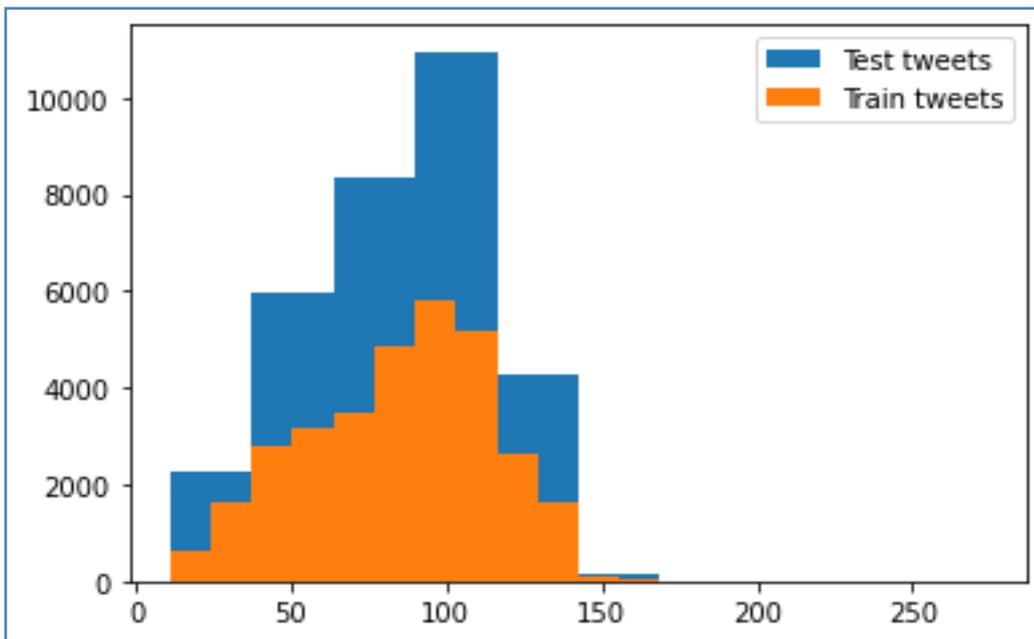


Figure 3 Graph plot representing length of train tweets and test tweets

5.4 Data Preprocessing

Data to be cleaned using regular expressions and we need to remove the similar patterns in the data and replacing the emojis by white spaces We need to combine the training and test tweets for implementing data preprocessing During data preprocessing, the stop-words like it consists of less than or equal to three letters have to be removed Tokenization is used to secure the sensitive data be changing the original data. To reduce the word to the root form, stemming is implemented using porter stemmer.

Table 4 Actual and Predicted values of confusion matrix -SVM Classifier

	positive	negative
positive	True positive 9312	False Negative 374
negative	False Positive 55	True Negative 91

Table 5 Features Performance Measures OF SVM Classifier

Accuracy	Precision	Recall	F1 Score
0.955	0.96	0.99	0.98

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

$$\text{Accuracy} = \frac{(9312 + 91)}{(9312 + 91 + 55 + 374)} = 0.955$$

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

$$\text{precision} = \frac{9312}{9312 + 55} = 0.96$$

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

$$\text{Recall} = \frac{9312}{9312 + 374} = 0.99$$

$$\text{F1 score} = \frac{2 * \text{precision} * \text{recall}}{(\text{precision} + \text{recall})} = 0.98$$

6. Performance analysis

We have used other classifiers to test the accuracy of our model. The other classifiers are naive Bayes and Random Forest classifiers. now the accuracy of the classifiers we need to predict. but we need to train the models by fitting the data into the model and then we need to predict the accuracy. First, we need to implement confusion matrix. we need to import the naive Bayes and random forest classifiers from sklearn library and then we need to fit the data into these models and the we have the predicted values. we need to import confusion matrix from sklearn Metrics and we need to plot the confusion matrix using the matplotlib library.



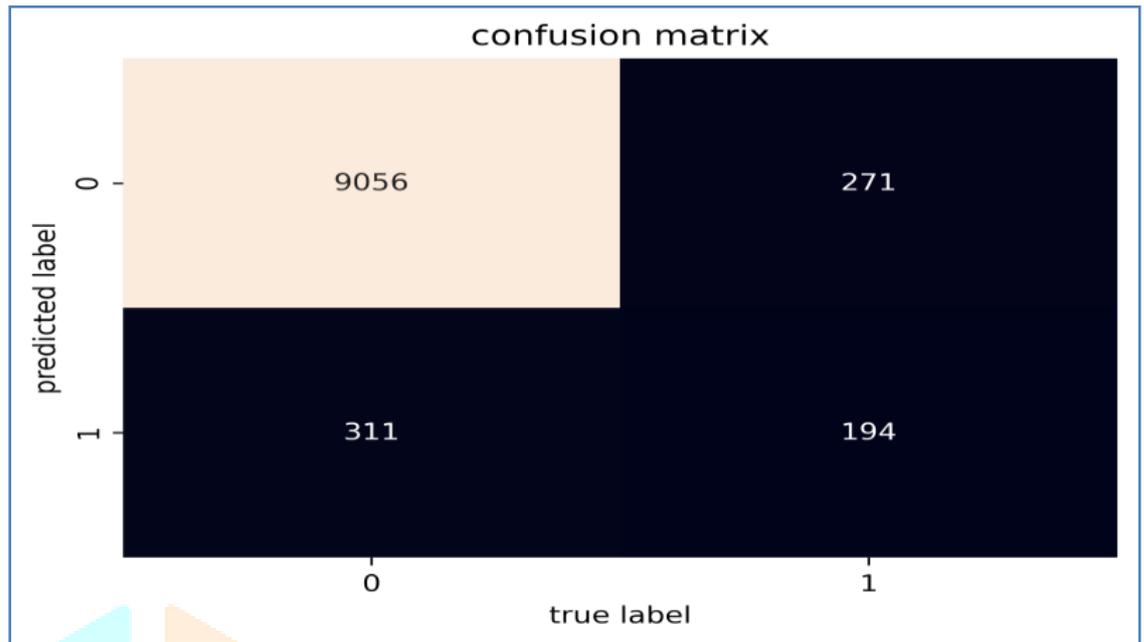


Figure 7 Confusion matrix for naive Bayes classifier

Confusion matrix for Naïve bayes classifier in Figure 7 is N*N matrix used for evaluating the performance of classification model here N is no of target classes The matrix compares the actual target values with those predicted by machine learning model.

Table 6 Actual and Predicted values of confusion matrix Naïve bayes Classifier

	Positive	Negative
Positive	True Positive 9056	False Negative 271
Negative	False Positive 311	True Negative 194

Table 7 Performance Measures Of Naïve bayes Classifier

Accuracy	Precision	Recall	F1 Score
0.94	0.97	0.97	0.97

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

$$\text{Accuracy} = \frac{(9056 + 194)}{9832} = 0.94$$

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

$$\text{precision} = \frac{9056}{9327} = 0.97$$

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

$$\text{Recall} = \frac{9056}{9367} * 100\% = 0.97$$

$$\text{F1 score} = \frac{2 * \text{precision} * \text{recall}}{(\text{precision} + \text{recall})} = 0.98$$

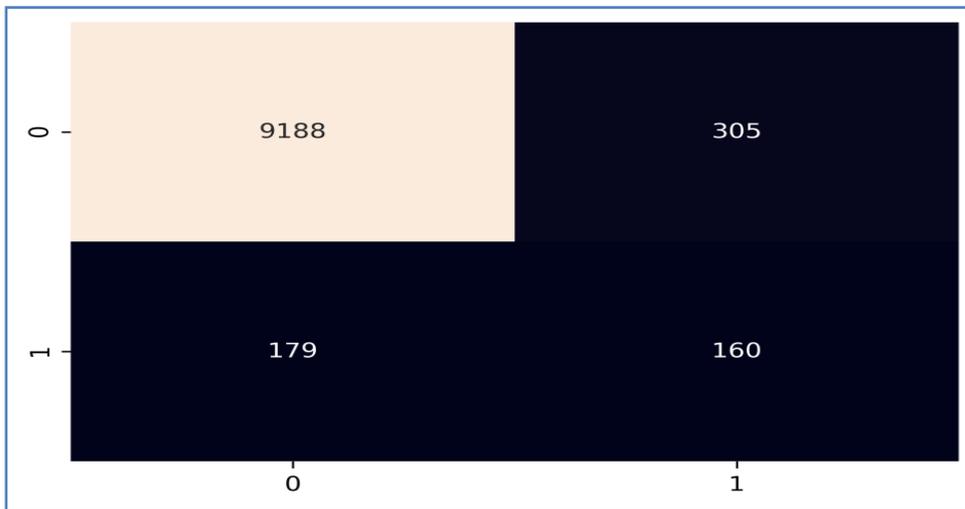


Figure 8 Confusion matrix for Random forest classifier

Confusion matrix for Random forest classifier in Figure 8 is N*N matrix used for evaluating the performance of classification model here N is no of target classes The matrix compares the actual target values with those predicted by machine learning model

Table 8 Confusion matrix Random for Forest Classifier

	Positive	Negative
Positive	True Positive 9188	False Negative 305
Negative	False Positive 179	True Negative 160

Table 9 Performance Measures Of Random Forest Classifier

Accuracy	Precision	Recall	F1 Score
0.95	0.97	0.98	0.97

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

$$\text{Accuracy} = (9188 / 9832) = 0.95$$

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

$$\text{precision} = (9188 / 9413) = 0.97$$

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

$$\text{Recall} = (9312 / 9367) = 0.98$$

$$\text{F1 score} = \frac{2 * \text{precision} * \text{recall}}{(\text{precision} + \text{recall})} = 0.97$$

The comparative analysis of our proposed work is done with SVM algorithm on the twitter data set The results of the proposed approach using Colab tool are presented in the Table 10.

TABLE 10 RESULTS OF SVM VERSUS NAÏVE BAYES VERSUS RANDOM FOREST ON ALL EVALUATION METRICS

Measures	SVM	Naïve bayes	Random Forest
Accuracy	0.955	0.94	0.95
Precision	0.96	0.97	0.97
Recall	0.99	0.97	0.98
F1 Score	0.99	0.97	0.97

The accuracy of the proposed approach is Increased from 94 to 95.5. The precision value of the proposed approach is increased from 0.96 to 0.97 which helps in improving the prediction of the specific class. The F-score value of the proposed approach improved from 0.896 to 0.961. The recall value of the proposed approach is improved from 0.97 to 0.99. The recall of the proposed approach is increased from 0.976 to 0.99.

From Table 10 and Figure 9 we can see that our proposed svm approach had given a proper solution. The comparative analysis of the machine learning framework is done with SVM algorithm on the twitter data set. In the Figure 9 we can see the results of performance measures in a visual representation.

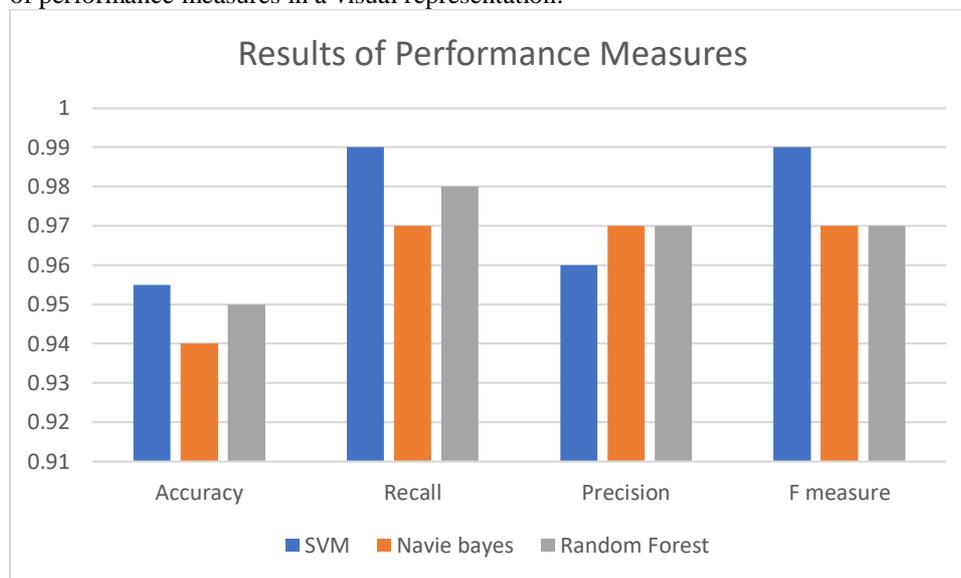


Figure 9 Test results on SVM and Navie bayes and Random Forest on twitter datasets.

Fig 9 shows the test results of F-Score, Recall, Accuracy, Precision and FP Rate on SVM and Navie bayes and Random Forest on twitter datasets.

7. Conclusion

In this paper developed a ML based model which performs sentiment analysis on Twitter data using Machine Learning Technique. The model that was proposed in this research built by using (NLTK) on the dataset containing tweets. The concept is used which contains both positive and negative words separately. The classification was done using different classifiers and calculated the probability of new input data and the tweet with the highest value is considered as either positive tweets or negative tweets. However, we choose an effective twitter feature dataset which enhances the accuracy of the classifier. As per the observation in the results, SVM model outperforms than the other compared models. This model can further enhance to any desired level if one wants to incorporate more features in the database.

Références

1. L.Huang, Go, R. Bhayani. "Sentiment Assessment Using distant Supervision". Stanford University, Technical Paper, 2009
2. M. S. Neethu and R. Rajasree, "Sentiment assessment in twitter using tool analyzing techniques," 2013 (ICCCNT), Tiruchengode, 2013, pp. 1-5.
3. Le B., Nguyen H. (2015) "TSA Using ML Techniques". In: Le Thi Advances in Intelligent Systems and Computing, vol 358. Springer, Cham
4. Sayali P. Nazare, Prasad S. Nar, Akshay S. Phate, Prof. Dr. D. R. Ingle "Sentiment Analysis in Twitter" International Research Journal of Engineering and Technology (IRJET) Volume: 05, Jan-2018.
5. Mejova, Yelena. (2019) "Sentiment Analysis: An Overview" Go, Alec, Richa Bhayani, and Lei Huang. "Twitter sentiment class using far off supervision." CS224N Project Report, Stanford 1 (2009): 12.
6. Boiy, Erik, Hens, Pieter, Deschacht, Koen & Moens, Marie-Francine, Marie-Francine. (2007). Austria 13-15 June 2007 / Edited through manner of method of Leslie Chan and Bob Martens. pp. 349-360, 2007.
7. Xiangyu Kong, Niu, Zhen and Zelong Yin., "Sentiment class for micro weblog through manner of method of tool analyzing." In lawsuits of 2012 Fourth IC on Computational and Information Sciences, pp. 286-289. 2012
8. Nine Pak, Alexander, and Patrick Paroubek. "Twitter as a Corpus for Sentiment Analysis and Opinion Mining." LREc. Vol. 10. 2010.
9. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. "Lexicon based methods for sentiment analysis". Computational linguistics, 2011:37(2), 267-307.
10. Dey, Lopamudra & Chakraborty, Sanjay & Biswas, Anuraag & Bose, Beepa & Tiwari, Sweta. (2016). "Sentiment Analysis of Review Datasets Using Naïve Bayes and K-NN Classifier". International Journal of Information Engineering and Electronic Business. 8. 54-62. 10.5815/ijieeb.2016.04.07.
11. Leena A. Deshpande, M.R. Narasingarao "Addressing social reputation in twitter statistics using waft detection technique" Journal of Engineering Science and Technology Vol. 14, No. 2 (2019) 922 – 934.
12. Tiruveedhula, Sajana & ramanarasingarao, Manda. (2017) "Machine analyzing techniques for malaria disorder diagnosis - A review" Journal of Advanced Research in Dynamical and Control Systems. nine. 349-369
13. Kavya Suppala Narasinga Rao IJITEE, (2019). "Sentiment Analysis Using Naïve bayes classifier"
14. R. Chandrasekaran and G. Vinodhini, "Sentiment assessment and opinion mining: A survey," International Journal, vol. 2, no. 6, 2012.
15. Abdullah Alsaeedi and Mohammed Zubair Khan. (2019) "A test on sentimental assessment of social media statistics international Journal of Advanced Computer Science and Applications 10(2):361-374.
16. A. Kolcz and J. Lin, "Large-Format tool analyzing at twitter," in Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data, 2012: ACM, pp. 793-804.
17. R. Sharma, S. Nigam, and R. Jain, "Opinion mining of movie evaluations at file level," arXiv preprint arXiv:1408.3829, 2014.
18. F. H. Khan, S. Bashir, and U. Qamar, "TOM: Twitter opinion mining framework using hybrid class scheme," Decision Support Systems, vol. 57, pp. 245-257, 201.
19. P.-W. Liang and B.-R. Dai, "Opinion mining on social media statistics," in Mobile Data Management (MDM), 2013 IEEE 14th International Conference on, 2013, vol. 2: IEEE, pp. 91-96.
20. Awais Muhammad and Samin Tanzila (2012), "Advanced SWOT Analysis of E-Commerce", IJCSI International Journal of Computer generation Issues, Vol nine, Issue 2, No 2, pp. 569-574.
21. Chou, D., C. and Chou, A., Y., "The E-Commerce Revolution, A Guide to the Internet Revolution in Banking" statistics systems management, pp. 51-57, 2000
22. Cao, X., Guo, X., Liu, H., & Gu, J. (2015). The role of social media in supporting statistics integration: A social capital assessment. Information Systems Frontiers, 17(2), 351–362.