MACHINE LEARNING ALGORITHMS COMPARATIVE STUDY ON SUICIDAL TEXT

[1]K. Sai Teja Reddy, [2]Dr. M. V. Rathnamma
[1]Student, [2]Associate Professor
[1]National Institute of Foundry and Forge Technology, [2]KSRM College of Engineering
[1]k.saitejareddy2001@gmail.com, [2]rathnamma@ksrmce.ac.in

Abstract --Machine learning is a branch of artificial intelligence that allows computers to recognize and exploit patterns in data to make predictions. Machine learning has become grown in popularity in recent years, particularly in the field of computer science, as it has been linked to a wide range of applications, including fraud detection, structured data, and recommender systems. Raw data, but vital client information, is constantly generated on social networking sites. However, when this data is mined using diverse methodologies, such as machine learning techniques, it becomes increasingly valuable. Moreover, a few studies showed a significant link between social media and suicide, this client-generated data could be used to potentially save lives, particularly of vulnerable social media users. We intend to contribute to the study of suicide communication on social media with our investigation. We evaluated the performance of the three machine learning algorithms: Naive Bayes, Random Forest, and Support Vector Machine by recognizing and classifying suicide-related information from Twitter.

Keywords: **Text classification, Machine Learning, SVM, NB, Prism Social media, Suicide**

Introduction

A social network, such as Facebook or Twitter, is a type of social media platform that provides a web-based service that allows users to communicate with one another [1]. These stages have benefited the general public while also posing a major threat to vulnerable web users who are at risk of harming themselves as a result of the data they get, such as the spread of suicidal ideation [2]. A few studies have found a link between social media and suicide behaviour [3–5], and the World Health Organization reports that those groups make up the majority of social media users [8, 9].

Overall, this has sparked some growing worries [6, 10, 11] about the effect or influence of social media on these vulnerable consumers. However, when handled correctly, these stages can provide a wealth of information on people groups' daily routines and behaviours, which can be used to think about and comprehend suicide, as well as potentially intervene [3].

As indicated by [2], to help social media clientele who are suicidal, understand the correspondence of suicidal ideation. Studies [2, 6, 12] have demonstrated that it is practically inevitable that a person will seek non-proficient aid through social media rather than proficient help because of concerns about social stigma. Following that, this investigation is aimed to add to the ongoing research on suicide in social media by directing a gauge trial to measure the effectiveness of well-known machine classifiers in distinguishing suicide-related and non-suicide-related communication. Similarly, we use information control to draw various renditions of preparing information and investigate the impact of information control on performance.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

Machine learning is a subset of artificial intelligence that provides tools for recognizing and utilizing patterns for prediction from data [1–3]. According to [3], the use of machine learning has increased quickly in recent years, particularly in computer science, where it has been applied to a variety of areas such as misrepresentation detection, drug design, and web search and recommender systems. Furthermore, one of the most common tasks in machine learning is classification [3–5], which involves determining the class of a hidden example.

The classification task has been performed using a variety of machine learning algorithms, one of which being Random Forest.RF is a classification determine learning calculation that was created [6–8], and while it is less well-known than other machine learning algorithms such as Decision Tree and Naive Bayes, known to be straightforward [7, 8].Furthermore, this calculation employs the different and-vanquish learning approach [9], which is based on the consecutive covering guideline [9]: a standard is established (the overcome stage), which accurately predicts the worth of the objective quality (for example, class); the covered occurrences are removed (the different stage), and the interaction is repeated until all examples are covered.

As a result, we will evaluate the categorization performance of this study using three calculations on brief text identifying with suicide correspondences on social media.

Implementation Methodology

Text Classification and Evaluation

Classification is one of the most well-known machine learning tasks, and it has been used in a variety of fields, including sentiment analysis. Furthermore, they stated that classification may be divided into two types: binary classification, which consists of two categories (classes) for information instances, and multi-class classification, which consists of many categories for instances.

In accordance with this, we divided the data into two datasets for binary and multi-class classification in our study. The Suicide and Flippant classes are in the binary dataset, whereas the Suicide, Flippant, and Non-Suicide classes are in the multi-class dataset. The Non-suicide category encompasses all classes that are unrelated to suicidal ideation or make a passing reference to suicide (for example Suicide or Flippant). As a result, the classes crusade, support, remembering, reports, and others all fit under the non-suicide category in this scenario.

Data Collection and Annotation

We acquired tweets using vocabulary terms from well-known social media websites to gather material including suicidal ideation, as well as names of deceased persons as search catchphrases to get tweets about suicide. Master human annotators compiled an aggregate of 8,000 tweets and divided them into seven categories.

An online communication that mentions suicide or contains indicators of suicidal contemplations can appear in a variety of formats, and not all of them are necessary for suicide prevention. The significance of a text is first determined using a clinical definition of suicide. It can either be related to the term, notice suicide in an unexpected way (in hyperboles or non-clinical contexts, such as suicide terrorism), or be unrelated to the definition. Only texts that match the definition are further commented on.

Feature Preparation

We used the text of the tweets to train and evaluate machine classifiers to distinguish between suicidal thoughts and other types of suicide-related communication, such as flippant references to suicide. The following feature sets were derived from the text:

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

Parts of Speech (POS) and other language structure aspects, such as the most frequently used words and phrases, are used to express lexical properties of the sentences. These are common traits found in most text mining projects. POS also detects references to oneself and others. These phrases have been identified as clear inside suicidal communication in previous studies.

Features that elicit an emotional response Characteristics that convey emotion are viable and passionate features, as well as the degrees of phrases were employed in the text. These were incorporated because of the extremely emotive nature of the task. Suicidal communication is characterized by emotions such as dread, anger, and general aggressiveness [1].

Parts of Speech We utilized the Stanford POS Tagger8 to assign a POS name to each word in a Tweet. Things (singular, plural, appropriate), action words (indicating tenses such as present, past, and present participle), first versus third person references, descriptors and modifiers (near, standout), pronouns (personal, possessive), and other labels (conjunctions, determiners, cardinal numbers, images, and contributions) are all examples of models. We considered the frequency of each POS in a Tweet as a component for each.

Other Structural Features We considered nullifications in the phrase (absolute number), the specific use of a first-person pronoun (singular or plural), and external communication elements such as the consideration of a URL in a tweet or a notice graphic as examples (showing a re-tweet or answer).

General Lexical Domains These characteristics indicate lexical categories such as "home," "religion," "brain research," "human science," and so on. WordNet Domains labels were often used to eliminate them.

Affective Lexical Domains These are a collection of categories that are directly linked to domains that represent 'successful' concepts. These include concepts representing states of mind, circumstances evoking feelings, or enthusiastic reactions such as joy, anger, grief, sadness, enthusiasm, surprise, love, hate, and happiness; yet more explicit sub-categories such as friendliness, aggressiveness, awful attitude, distress, and anxiety; and alternate extremes such as good regrettable concern, negative fear, good bad tension, hesitance, self-indulgence, self-belittling and confidence. These are incredibly appropriate for the language we're looking at in this study.

Sentiment ScoreSentiWordNet10 assigns a score to each word that ranges from nothing to one for both energy and pessimism. The total number of words in a Tweet was used as features.

Words The most frequently used words and n-grams are found among the (first 100) unigrams, bigrams, and trigrams in the preparation set.

Keyword list We also included all 62 keywords found in the Web structure text that were used in the pre-sifting search (for example, 'sleeping and never waking', 'don't have any desire to try any longer', 'end everything', 'does not merit carrying on with', 'my life is futile', 'commit suicide', 'to experience any more', 'need to end it', 'need to vanish', 'need to bite the dust', etc.). Each of the search phrases was added as a separate feature, with a single global binary component reflecting the consideration of any of them in a Tweet.

Algorithms Used

Naïve Bayes

In nature, Naïve Bayes is probabilistic. It is used for classification in a variety of real-world applications. It's utilized in things like heart disease prediction, spam filtering, cancer disease classification, document segmentation, and online review sentiment prediction. The Bayes theory underpins the Naïve

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

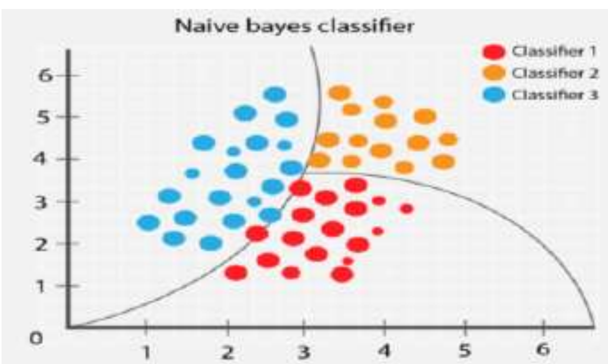
Bayes classification algorithm. Its features are self-contained, hence justifies the name Naïve. It does imply that changing the value of a feature has no direct impact on other features. Because it is probabilistic, this algorithm is found to be faster. It's also scalable, making it ideal for applications that require scalability.Bayes Rule and conditional probability are two related notions of Naïve Bayes.

In machine learning, Naive Bayes classifiers are a set of fundamental "probabilistic classifiers" based on Bayes' hypothesis with a solid (naive) freedom presumption between features.

$P(A|B)=P(B|A)P(A)/P(B)$

Using Bayesian probability terms, the above formula as bellow

Posterior = prior*likelihood/evidence



Random Forest

RF is another classification algorithm which follows ensemble classification approach. It's made up of a lot of DTs. Tin Kam Ho coined the name when it was originally established in 1995. RF combines Breiman's bagging concept with a random selection of features. Each of the decision trees in RF is a separate learner. When they are combined, they become random forest. Data exploration is one of the common approaches for which RF is widely used. Classification and Regression Tree is an example of a decision tree used in RF (CART). It divides the feature space into numerous regions using a recursive, top-down, and greedy technique.

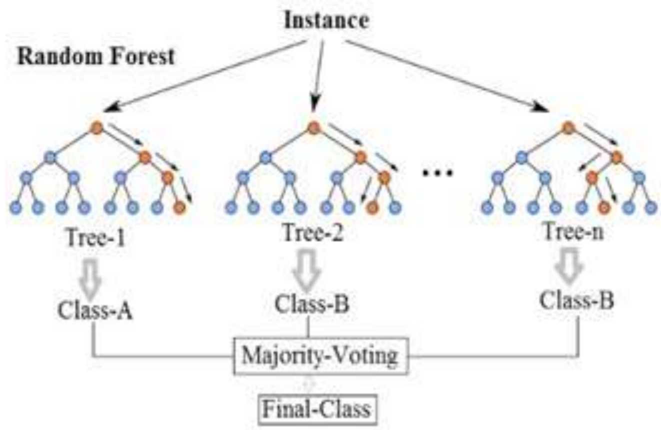


Figure 5: Random forest example

The RF generates numerous DTs for the specified instance, as seen in Figure 5. Each tree represents a distinct class. The final class will be decided by a majority vote. A training data subset is chosen from each tree. The stop condition is then checked. It comes to a halt after the condition is met, and it computes the forecast error. If the stop condition is not satisfied, the next split is built, which is then subjected to a series of operations, including variable subset selection and an iterative procedure to select the optimal split. The benefits of RF are numerous. For many datasets, it has a high level of accuracy. It is capable of effectively handling huge datasets.

Support Vector Machine

SVM is a frequently used machine learning algorithm in predicting outcomes for heart disease. Handwritten character recognition, cancer prediction, protein classification, image classification, and text labelling are just a few examples of real-world applications. It is a discriminative classifier that uses training data to predict class labels. It achieves the classification formally with a definition of hyperplane. Therefore, it can provide the largest minimum distance, also known as the maximum margin, for training samples.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

The margin of the data utilized for training can be minimized with this SVM. SVM's working method is depicted in Figure 2.

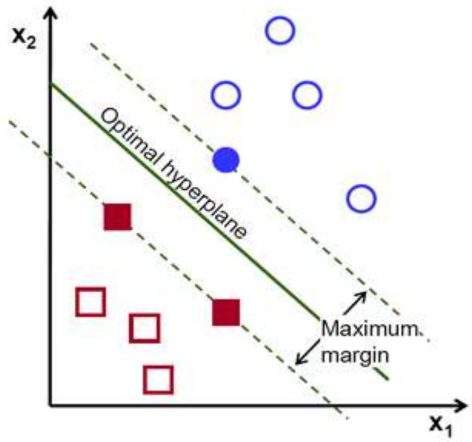
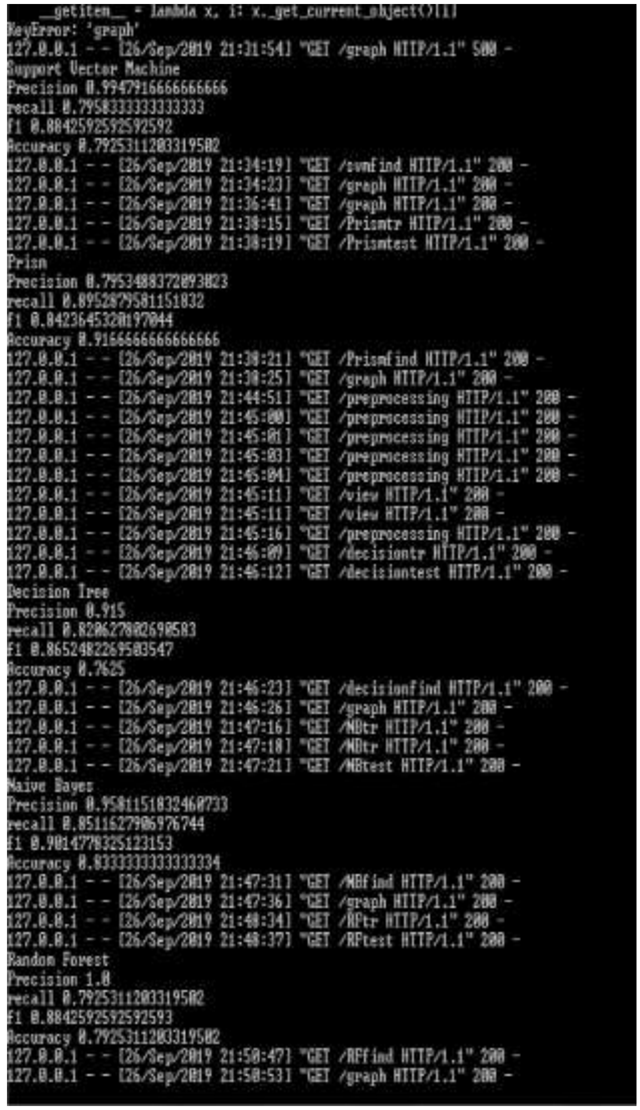


Figure 2: Illustrates how hyperplane is formed for discrimination

Result Analysis and Performance

The dataset was then used to evaluate the performance of three machine classifiers in categorizing suicide-related text: NB, RF, and SVM. These classifiers were chosen due to their widespread use as well as their properties: SVM functions admirably with text that are short and informal [2, 12], NB settles on classification choice dependent on the probability of feature event [2], while RF assist with diminishing the quantity of false adverse outcomes [2].

The performance aftereffects of singular classifiers are shown in the following area, for example, the standard classification measure scores: Precision (P), Recall (R), F-measure (F), and Accuracy (A), for both binary and multi-class datasets.



```
    __getitem__ = lambda x, i: x._get_current_object()[i]
KeyError: 'graph'
127.0.0.1 - - [26/Sep/2019 21:31:54] "GET /graph HTTP/1.1" 500 -
Support Vector Machine
Precision 0.9947916666666666
recall 0.7958333333333333
f1 0.8842592592592592
Accuracy 0.7925311203319502
127.0.0.1 - - [26/Sep/2019 21:34:19] "GET /svmfind HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:34:23] "GET /graph HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:36:41] "GET /graph HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:38:15] "GET /Prismtr HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:38:19] "GET /Prismtest HTTP/1.1" 200 -
Prism
Precision 0.7953488372093023
recall 0.8952879581151832
f1 0.8423645320197044
Accuracy 0.9166666666666666
127.0.0.1 - - [26/Sep/2019 21:38:21] "GET /Prismfind HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:38:25] "GET /graph HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:44:51] "GET /preprocessing HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:45:00] "GET /preprocessing HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:45:01] "GET /preprocessing HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:45:03] "GET /preprocessing HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:45:04] "GET /preprocessing HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:45:11] "GET /view HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:45:11] "GET /view HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:45:16] "GET /preprocessing HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:46:09] "GET /decisiontr HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:46:12] "GET /decisiontest HTTP/1.1" 200 -
Decision Tree
Precision 0.915
recall 0.8206278026905830
f1 0.8652482269503547
Accuracy 0.7625
127.0.0.1 - - [26/Sep/2019 21:46:23] "GET /decisionfind HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:46:26] "GET /graph HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:47:16] "GET /NBtr HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:47:18] "GET /NBtr HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:47:21] "GET /NBtest HTTP/1.1" 200 -
Naive Bayes
Precision 0.9581151832460733
recall 0.8511627906976744
f1 0.9014778325123153
Accuracy 0.8333333333333334
127.0.0.1 - - [26/Sep/2019 21:47:31] "GET /NBfind HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:47:36] "GET /graph HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:48:34] "GET /RFtr HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:48:37] "GET /RFtest HTTP/1.1" 200 -
Random Forest
Precision 1.0
recall 0.7925311203319502
f1 0.8842592592592593
Accuracy 0.7925311203319502
127.0.0.1 - - [26/Sep/2019 21:50:47] "GET /RFfind HTTP/1.1" 200 -
127.0.0.1 - - [26/Sep/2019 21:50:53] "GET /graph HTTP/1.1" 200 -
```

Figure: Console window results for used algorithms

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

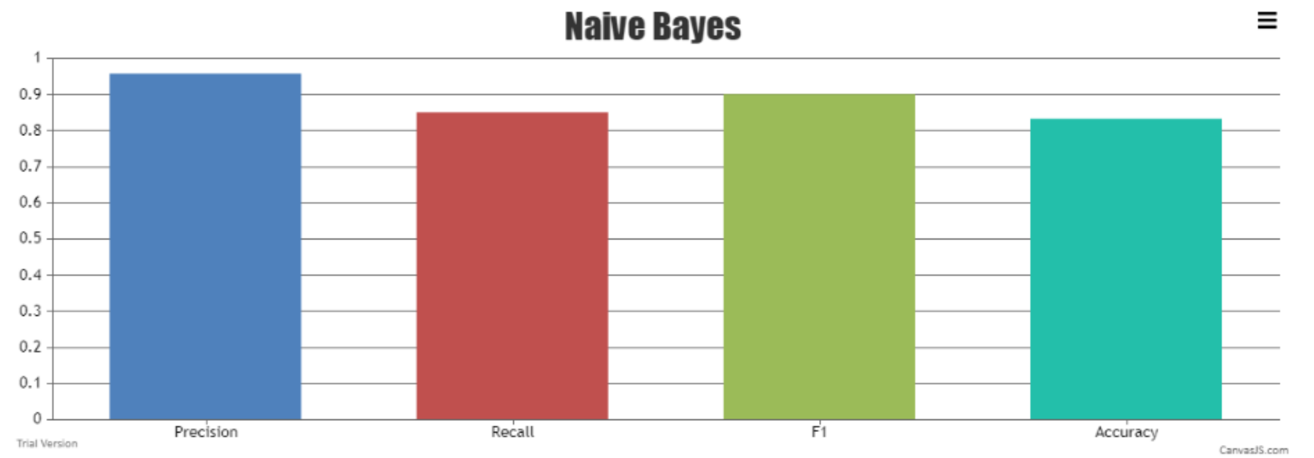


Figure: Performance of Naïve Bayes

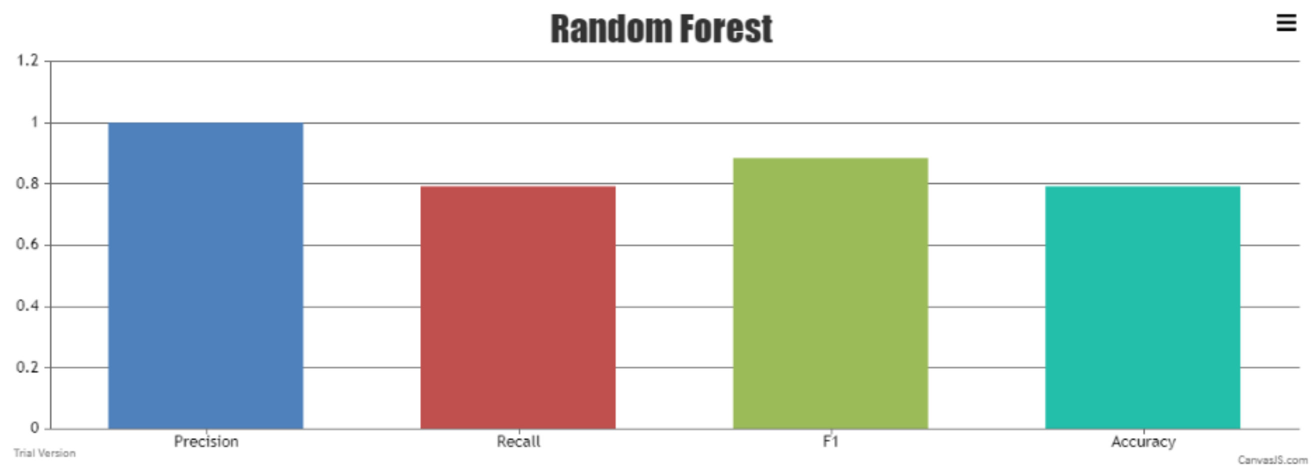


Figure: Performance of Random Forest

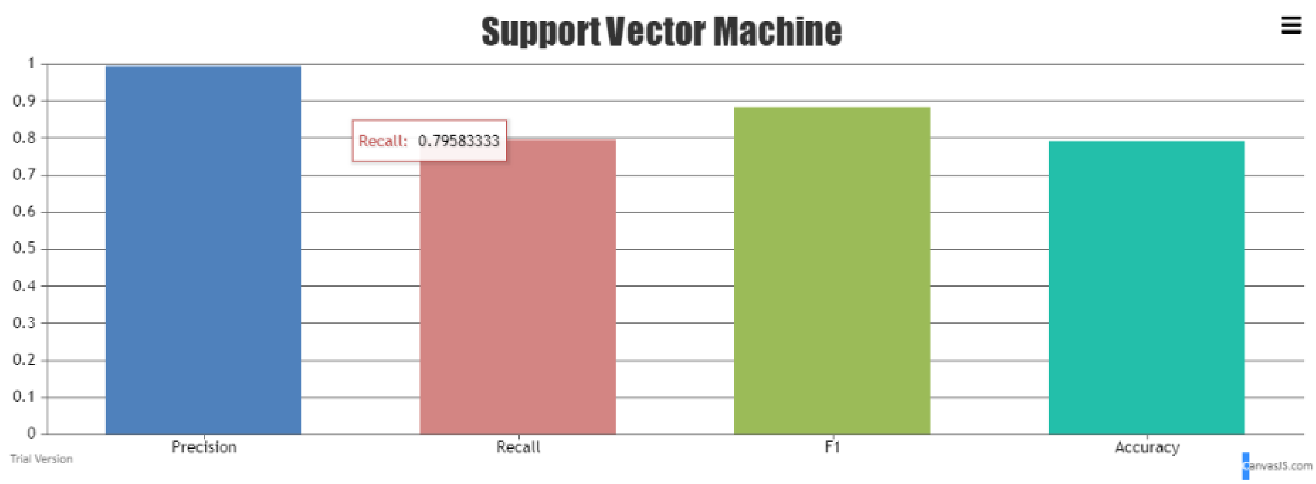


Figure: Performance of SVM

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

Conclusion

Machine learning approaches are becoming increasingly popular, and some algorithms, both well-known and less well-known, have exhibited varied characteristics when applied to short informal text. The study's purpose was to compare the classification performance of the three machine learning calculations to other widely used and well-established machine learning methodologies. According to the study's findings, the Naive calculation outperformed the other algorithms when applied to a variety of datasets of short texts. We will expand on this research by looking into the concept of multi-task learning, which is thought to produce superior results since it is more accurate and resilient, as well as other feature selection methods that have benefits such as lower dimensionality and increased classification performance.

References

[1] Y. Bas¸tanlar and M. O¨ zuysal, "Introduction to machine learning," in miRNomics: MicroRNA Biology and Computational Analysis. Springer, 2014, pp. 105–128.

[2] C. Robert, "Machine learning, a probabilistic perspective,"2014.

[3] P. Domingos, "A few useful things to know about machine learning," Communications of the ACM, vol. 55, no. 10,pp. 78–87, 2012.

[4] H. Liu, M. Cocea, A. Mohasseb, and M. Bader, "Transformation of discriminative single-task classification intogenerative multi-task classification in machine learning context," 9th International Conference on Advanced Computational Intelligence, ICACI 2017, pp. 66–73,2017.

[5] H. Liu, M. Cocea, and W. Ding, "Multi-task learning forintelligent data processing in granular computing context,"Granular Computing, p. in press, 2017. [Online].

[6] J. Cendrowska, "PRISM: An algorithm for inducing modular rules," International Journal of Man-Machine Studies, vol. 27, no. 4, pp. 349–370, 1987.

[7] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2016.

[8] M. Aburrous, M. A. Hossain, K. Dahal, and F. Thabtah, "Intelligent phishing detection system for e-banking using fuzzy data mining," Expert Systems with Applications, vol. 37, no. 12, pp. 7913–7921, 2010.

[9] J. F¨urnkranz, "Separate-and-conquer rule learning," Artificial Intelligence Review, vol. 13, no. 1, pp. 3–54, 1999.

[10] T. Trafalis and H. Ince, "Support vector machine for regression and applications to financial forecasting," Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium, no. May2016, pp. 348–353 vol.6, 2000.

[11] J. A. Cruz and D. S. Wishart, "Applications of machine learning in cancer prediction and prognosis," Cancer Informatics, vol. 2, pp. 59–77, 2006.

[12] G. D. Magoulas and A. Prentza, Machine Learning inMedical Applications. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 300–307.

[13] C. Catal and M. Nangir, "A sentiment classification modelbased on multiple classifiers," Applied Soft Computing Journal, vol. 50, pp. 135–141, 2017.

[14] S. Kiritchenko, X. Zhu, and S. M. Mohammad, "Sentiment analysis of short informal texts," Journal of Artificial Intelligence Research, vol. 50, pp. 723–762, 2014.

[15] K. Nigam, A. K. Mccallum, S. Thrun, and T. Mitchell,"Text Classification from Labeled and Unlabeled Documents using EM," Machine Learning, vol. 39, pp. 103–134, 2000.

[16] B. Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, andM. Demirbas, "Short text

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

classification in twitter to improve information filtering," in Proceedings of the 33rd international ACM SIGIR conference on Research and developmentin information retrieval, 2010, pp. 841–842.

[17] A. Abboute, Y. Boudjeriou, G. Entringer, J. Az´e,S. Bringay, and P. Poncelet, "Mining twitter for suicide prevention," in International Conference on Applications of Natural Language to Data Bases/Information Systems. Springer, 2014, pp. 250–253.

[18] L. Barbosa and J. Feng, "Robust sentiment detection on twitter from biased and noisy data," in Proceedings of the 23rd international conference on computational linguistics:posters, 2010, pp. 36–44.

[19] G. B. Colombo, P. Burnap, A. Hodorog, and J. Scourfield,"Analysing the connectivity and communication of suicidal users on twitter," Computer Communications, vol. 73,pp. 291–300, 2016.

[20] E. Haddi, X. Liu, and Y. Shi, "The role of text preprocessing in sentiment analysis" Procedia Computer Science, vol. 17, pp. 26–32, 2013.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731