IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-7, Issue-3, March 2019 E-ISSN: 2347-2693

A Reckoning Analysis and Assessment of Different Supervised Machine Learning Algorithm for Breast Cancer Prediction

Pragati Prakash[1*], Nidhi Ekka[2], Manjit Jaiswal[3]

[1,2,3]Department of Computer Science & Engineering, Guru Ghasidas Vishwavidyalaya, Bilaspur, India

[*]*Corresponding Author: 25pragati05@gmail.com, Tel.: +91-8827305386*

DOI: https://doi.org/10.26438/ijcse/v7i3.8388 | Available online at: www.ijcseonline.org

Accepted: 09/Mar/2019, Published: 31/Mar/2019

Abstract— Throughout the 20th century, views about breast cancer have drastically changed. Breast cancer is the most common cancer in women worldwide, with nearly 1.7 million new cases diagnosed in 2012. This type of cancer is the second most common cancer overall.

Our work brings out comparison based on the performance of supervised machine learning algorithms on a binary classification task. The supervised machine learning algorithms which are taken into consideration in the following work are namely Support Vector Machine (SVM), Decision Tree (DT), K Nearest Neighbour (KNN), and Naïve Bayes (NB). This paper mostly focuses on detailed analysis and comparing the performance of above-mentioned algorithms on one binary classification task by analysing the Metrics such as Accuracy, Precision, Misclassification Rate, False Positive Rate, True Positive Rate and Specificity. The main part of the project is creating a useful tool for predicting breast cancer with high accuracy before getting ill or in the initial stage of the disease. In other words, we can anticipate the future for women diseases.

Keywords— Machine Learning, Breast Cancer, CART, Naive Bayes, K nearest neighbors, Support Vector Machine

I. INTRODUCTION

Supervised Learning is the machine learning methodology in which we aim to approximate a mapping function to map input values to the target values or output using training data which is already labelled. By learning the association between input and the given correct output, supervised learning will build a model that can predict the output value given input value. Supervised learning methodology can be divided into Regression and Classification problems.

Regression problems are those in which the output is a real-valued number such as 'gross revenue' and Classification problems are those in which the output is a category such as 'spam' and 'no spam'. Following are the major supervised machine learning algorithms: Linear Regression, Logistic Regression, Support Vector Machines, Naïve Bayes, Decision Trees, Linear Discriminant Analysis, K Nearest Neighbor algorithm, and Neural Networks.

In this paper, we have considered only the binary classification problem and for this purpose we evaluated the performance of Support Vector Machines, Decision Tree, K Nearest Neighbour, Naïve Bayes and Random Forest on one binary classification problem with six performance metrics: Accuracy, Precision, Misclassification Rate, False Positive Rate, True Positive Rate, and Specificity.

SVM algorithm is a non-probabilistic machine-learning algorithm which learns to build its model by classifying points in the feature space [7]. In this work, we have used a Radial Basis Function.

KNN algorithm which is non-parametric in nature will aim to locate the majority vote in a group of the k- closest neighbors but it does not depend on overall data structure and hence does not require training it explicitly [8]. Naive Bayes classifier is based on Bayes theorem and it simplifies learning by assuming that features are independent given each class [9].

Decision Tree is non-parametric based which is aimed to predict label for binary classification by building a tree-structure model. But the problem with Decision Tree is that the tree can grow complicatedly very deep with serious overfitting problems [10].

In section 2 we have Literature Survey of the different Machine learning algorithms, then in section 3, we discuss about detailed explanation and analysis of the above mentioned algorithms. In section 4, the experimental setup, the dataset used along with required pre-processing of datasets before the experiment and the performance metrics to be considered for evaluation. Finally, in section 5 we have comparison based results and conclusion.

II. LITRATURE SURVEY

A. Survey on CART Algorithm

In computer science, a decision tree is used in Decision tree learning for an item is from its observation to its target value conclusion. In decision analysis, a decision tree can be used to visually and explicitly represent decisions and decision making. In data mining, a decision tree describes data. As due to its advantages like simplicity in understanding and interpreting, the ability to handle both numerical and categorical data, little data preparation requirement, well performed with large datasets etc., it is widely used.

In [2], a study attempted to use additional data to classify different images through a rule-based classifier using the CART approach. This research studies the urban landscape dynamics for the city of Columbia, Missouri USA using multi-temporal and multi-date (1984, 92, and 2000) Landsat TM and ETM satellite imageries. This study assists urban planners in more effective urban planning and management strategies. In the medical field also CART has been used. Using CART approach, [3] proposes an automatic method to segment heart sounds in early diagnosis of any heart disease. By using the CART approach, they achieved 99.14% accuracy, 100% sensitivity and 98.28% specificity on the dataset used for the experiment. Prediction of Arthritis using data mining techniques was carried out in [6] using tool-WEKA on the dataset of arthritis. They also used the confusion matrix generated by the algorithm to determine which attribute will be the best predictor for a correct prediction. In [4], combined CART and fuzzy logic have been able to be implemented for classification as an intrusion detection system. Training, testing and validation of the model are done by using KDD Cup 1999 dataset that has been through the pre-processing and cleaning data process. Accuracy testing and validation are calculated by using the confusion matrix. From several tests performed, the best model is built from training 70%, the depth of tree 11 and node leaf minimum percentage 90% with an accuracy of 85, 68% and average time validation was 21, 92 second. Two segments of digital images collected by the sensors TM and AVIRIS were analysed and classified using CART to classify Remotely-Sensed Digital Images in [5].

B. Survey on Naïve Bayes Algorithm

This Classification is named after Thomas Bayes (1702-1761), who proposed the Bayes Theorem. In machine learning, naive Bayes classifiers are all about applying Bayes' theorem with strong (naive) independence assumptions between the features. The Naïve Bayes classifier greatly simply learning by assuming that features are independent given class [11]. This study demonstrates that the impact of the distribution entropy on the classification error, showing that low-entropy feature distributions yield good performance of naive Bayes. According to [12], the Naive Bayes algorithm is one of the most effective methods in the field of text classification, but only in the large training sample set can it get more accurate results. It is a study of Naïve Bayes classification algorithm based on the Poisson distribution model, which results in high classification accuracy even in the small sample set. Naïve Bayes can also be used to extract and classify web text content like in [13] using website users requests and extraction on web mining which found 93.86% accuracy with 260 keywords. Despite the strong assumption that all features are conditionally independent given the class, Naive Bayes performance is surprisingly well in many real-world applications [14]. [14] Introduced three novel optimization models for the naive Bayes classifier where both class probabilities and conditional probabilities are considered as variables. Later their performances were compared with the naive Bayes classifier, tree augmented naive Bayes [15], the SVM, C4.5 [16] and the nearest neighbor classifier. Their obtained results demonstrated an improved performance in naive Bayes classifier maintaining its structure. In Naive Bayes classifier, a strong correlation between attributes due to the conditional independence assumption is not always true in the real world. In Hidden Naive Bayes (HNB) algorithm [17], each attribute corresponds to a hidden parent which combines the influences of all other attributes. Even with too much test time on HD datasets cost if compared to other Bayesian algorithms, it gives significantly improved performance. A novel model Packaged Hidden Naive Bayes (PHNB) [18], was proposed which resulted in a reduction in the test time on HD datasets and has higher accuracy on some particular datasets as compared to HNB. A new instance-weighting approach [19] was also introduced to improve the performance of naive Bayes text classifier in which every training instance in a subset is weighted according to the distance between it and the mean of the training subset and compared with other naive Bayes text classifiers.

C. Survey on k-nearest neighbor Algorithm

In pattern recognition, the k-nearest neighbor algorithm (k-NN) is a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space. The output depends on whether k-NN is used for classification or regression. For both classification and regression, there are a lot of applications can be considered like text mining, agriculture, finance and medicine. In [20], the KNN approach was used for predicting Economic Events. A KNN model-based approach was introduced in [21] to overcome the low efficiency and dependency on k then compared with C5.0 and standard kNN to be more efficient and data tuples for classification with a 90.41% reduction rate on average.

AdaNN [22] overcomes the limitation of the traditional *k*NN algorithm. It focuses on finding suitable k for each test

example. KNN (94.95%) and MKNN (99.51%) are compared in [23] using confusion matrix and resulted that MKNN is more accurate about 5-7% from KNN. In medical diagnosis and treatment by cutting out the ROI (region of interest), using image segmentation by kNN approach can be used [24]. The k-nearest neighbor decision rule is known to provide a useful nonparametric procedure for pattern classification [25]. GNNR [26] was proposed for the same which is based on the k nearest neighbour density estimation technique.

D. Survey on Support Vector Machine

SVMs can be used to solve various real-world problems. SVMs are helpful in text and hypertext categorization. It was used in [27] for the sentiment detection from a Punjabi News Article and found 90% accuracy. A new method of features selection based on ranking scores can be derived from SVMs [28]. Classification of images can also be performed using SVMs. This is also true of image segmentation systems, including those using a modified version SVM that uses the privileged approach as suggested by Vapnik [1]. In [29], classification of multisensory datasets, consisting of multitemporal SAR data and optimal imagery were addressed using an SVM-based fusion approach. Hand-written characters can be recognized using SVM. The SVM algorithm has been widely applied in the biological and other sciences. Smile detection method [32] and face recognition [34] based on SVM and HM-SVM respectively and in medical disease predictions also SVM is widely used like in Cancer detection [28,30,31] and brain tumor detection [34] using MRI images. They have been used to classify proteins with up to 90% of the compounds classified correctly. Permutation tests based on SVM weights have been suggested as a mechanism for interpretation of SVM models. Support vector machine weights have also been used to interpret SVM models in the past. Posthoc interpretation of support vector machine models in order to identify features used by the model to make predictions is a relatively new area of research with special significance in the biological sciences. The study in [35] provides a data mining approach for malicious software detection using a linear SVM algorithm. A research on DDoS attacks detection based on RDF-SVM [36] was addressed using optimal feature subset over KDD99 dataset and resulted to have a higher detection rate and recall rate compared to other methods.

III. DETAILED ANALYSIS OFALGORITHMS

Relevant details should be given including experimental design and the technique (s) used along with appropriate statistical methods used clearly along with the year of experimentation (field and laboratory).

A. Classification and Regression Tree

CART follows a tree-structured classification scheme where the nodes represent the input variables and the leaves correspond to decision outcomes. DTs are one of the earliest and most prominent ML methods that have been widely applied for classification purposes. Based on the architecture of the DTs, they are simple to interpret and "quick" to learn. When traversing the tree for the classification of a new sample we are able to conjecture about its class. The decisions resulted from their specific architecture allow for adequate reasoning which makes them an appealing technique. The optimization formula to calculate entropy is

$$E(S) = \sum_{i=1}^{c} - p_i \log_2 p_i \qquad (1)$$

B. Naïve Bayes

A Naive Bayes classifier is also described as an elementary probabilistic classifier which is constructed on the implementation of Bayes' theorem along a dedicated set of assumptions for conditional independence. To get optimization problem, let x and y be the random variables then according to Bayes Rule [37] it will be:

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)}. \qquad (2)$$

C. K-Nearest neighbor

The k-Nearest neighbor technique is a non-parametric method for classification. The neighbors are taken from a set of objects for which the class for classification is known. KNN simply stores the training data. It is a simple technique that is capable of handling extremely complex tasks such as identifying cancerous masses. The optimization formula for kNN, using Euclidean distance as the distance metric is

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}. \qquad (3)$$

Where p and q are different attributes. After calculating distances all k nearest neighbors, we predict the class using majority votes of k nearest points.

D. Support Vector Machine

Support Vector Machines (SVMs) are the newest supervised machine learning technique (Vapnik, 1995) [1]. SVMs revolve around the notion of a "margin"—either side of a hyperplane that separates two data classes. Maximizing the margin and thereby creating the largest possible distance between the separating hyperplane and the instances on either side of it has been proven to reduce an upper bound on the expected generalization error. The most well-known SVM classifier [1] can be expressed as:

$$\min_{\mathbf{w},b} \frac{1}{n}\sum_{i=1}^{n}\left[1 - y_i\left(b + \mathbf{w}\cdot\mathbf{x}_i\right)\right]_+ + \lambda\left\|\mathbf{w}\right\|_2^2. \quad (4)$$

Where **w** is the vector coefficient, b is the intercept, **x** is sample and y is the class label.

IV. EXPERIMENTAL SETUP

A. The Data Set

In this paper, a breast cancer dataset is used, which is available from the UCI machine learning repository. This is a relatively small scale dataset, which is composed of 569 data samples and each data sample has 31 different features.
The given dataset is divided into 90% training and 10% testing sets based on the 10-fold cross-validation strategy. In order to evaluate the size, shape and texture of each cell nuclei, 10 characteristics were derived namely the radius, perimeter, area, compactness, smoothness, concavity, concave points, symmetry, fractal dimension and texture.
Next, we calculate the confusion matrix for each algorithm respectively using simple coding python tool. An example of the confusion matrix has been given below in Table 1. The confusion matrix contains True Positive (TP), False Positive (FP), False Negative (FN), and True Positive (TP) the explanation for which have been given in section *B.*

Table 1. Confusion Matrix terminology

n=165	**Predicted: NO**	**Predicted: YES**	
Actual: NO	TN = 50	FP = 10	60
Actual: YES	FN = 5	TP = 100	105
	55	110	

B. Methodology

Table 2. Terminology used for comparison.

Parameters	Definition	Formula
Accuracy	Ratio of correctly predicted observation	(TN+TP)/number of instances
Misclassification Rate	Shows how often the classifier is wrong. Also called Error Rate.	1-Accuracy
True Positive Rate	It shows if it is actually yes, then how often classifier predicted yes.	TP/FN+TP
False Positive Rate	It shows if it is actually no, then how often classifier predicted yes.	FP/TN+FP
Specificity	It shows if it is actually no, then how often classifier predicted no.	1-False Positive Rate
Precision	It shows when it predicts yes, then how often it is correct.	TP/FP+TP

V. COMPARISION BASED RESULT AND CONCLUSION

The results of each algorithm are recorded in Table 3. It consists of all the six above mentioned terminologies for each algorithm respectively. The data is represented graphically in Figure. 1.
From Figure 1 we conclude that the SVM algorithm is best suited for the breast cancer prediction with its specificity being highest among all (94.7%) and misclassification rate being lowest among them (19.5%).

Table 3. Performance comparisons of algorithms

PERFORMANCE MATRIX	CART	NAÏVE BAYES	KNN	SVM
Accuracy	0.756	0.769	0.787	0.801
Misclassification Rate	0.243	0.230	0.212	0.195
True Positive Rate	0.392	0.620	0.502	0.569
False Positive Rate	0.052	0.152	0.068	0.072
Specificity	0.927	0.847	0.931	0.947
Precision	0.794	0.680	0.784	0.803

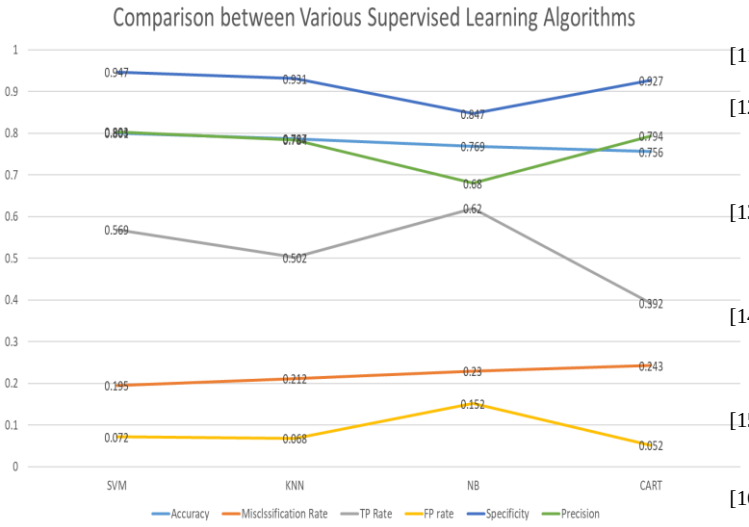

Figure 1. Graphical representation of performance comparison

REFERENCES

[1] Cortes C, Vapnik V. "*Support-vector Networks"*. Machine Learning, 20(3):pp. 273-297. 1995.

[2] P. Manoj Kumar, R. Sugumaran and D. Zerr, *"A rule-based classifier using Classification and Regression Tree (CART) approach for urban landscape dynamics,"* IEEE International Geoscience and Remote Sensing Symposium, Toronto, Ontario, Canada, Vol.2, pp. 1193-1194, 2002

[3] A. M. Amiri and G. Armano, *"Early diagnosis of heart disease using classification and regression trees,"* The 2013 International Joint Conference on Neural Networks (IJCNN), Dallas, TX, 2013, pp.1-4. doi: 10.1109/IJCNN.2013.6707080

[4] A. F. A. Pinem and E. B. Setiawan, *"Implementation of classification and regression Tree (CART) and fuzzy logic algorithm for intrusion detection system,"* 2015 3rd International Conference on Information and Communication Technology (ICoICT), Nusa Dua, pp. 266-271, 2015.

[5] H. R. Bittencourt and R. T. Clarke, "*Use of classification and regression trees (CART) to classify remotely-sensed digital images,"* IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477), Toulouse, Vol.6, pp. 3751-3753. 2006 doi: 10.1109/IGARSS.2003.1295258

[6] N. Bhargava, R. Purohit, S. Sharma and A. Kumar, "Prediction of arthritis using classification and regression tree algorithm," *2017 2nd International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore,pp. 606-610. 2017.

[7] Xuegong, Zhang. "*Introduction to statistical learning theory and support vector machines".* Acta Automatica Sinica 26.1 (2000): pp.32-42.

[8] Weinberger, Kilian Q., John Blitzer, and Lawrence Saul. "*Distance metric learning for large margin nearest neighbor classification".* Advances in neural information processing systems 18(2006): pp.1473

[9] Nigel Williams and Sebastian Zander and Grenville J. Armitage. "*A preliminary performance comparison of five machine learning algorithms for practical IP traffic flow classification".* Computer Communication Review (2006).volume=36:pp.5-16.

[10] S. R. Safavian and D. Landgrebe, "*A survey of decision tree classifier methodology,"* in IEEE Transactions on Systems, Man, and Cybernetics, vol. 21, no. 3, pp. 660-674, May-June 1991. doi: 10.1109/21.97458

[11] Watson, Thomas J. . *"An empirical study of the naive Bayes classifier."* (2001).

[12] Y. Huang and L. Li, *"Naive Bayes classification algorithm based on small sample set,"* 2011 IEEE International Conference on Cloud Computing and Intelligence Systems, Beijing, pp. 34-39. 2011 . doi: 10.1109/CCIS.2011.6045027

[13] Sanjay S Bhadoria and Rajendra Kumar Patel, "*Web Text Content Extraction and Classification using Naïve Bayes Classifier Algorithm*", International Journal of Scientific Research in Computer Science and Engineering, Vol.2, Issue.5, pp.1-4, 2014

[14] Sona Taheri and Musa Mammadov. 2013. "*Learning the naive Bayes classifier with optimization models".* Int. J. Appl. Math. Comput. Sci. *23*, Vol.4 pp.787-795. (December 2013), DOI: https://doi.org/10.2478/amcs-2013-0059

[15] Zheng F., Webb G.I. (2011*) "Tree Augmented Naive Bayes".* In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. 2011.

[16] (2011) C4.5. In: Sammut C., Webb G.I. (eds). *"Encyclopedia of Machine Learning".* Springer, Boston, MA

[17] Jiang, Liangxiao et al. *"A Novel Bayes Model: Hidden Naive Bayes."* IEEE Transactions on Knowledge and Data Engineering *21* ,pp.1361-1371. 2009

[18] Y. Ji, S. Yu and Y. Zhang, *"A novel Naive Bayes model: Packaged Hidden Naive Bayes,"* 6th IEEE Joint International Information Technology and Artificial Intelligence Conference, Chongqing, pp. 484-487. 2011. doi: 10.1109/ITAIC.2011.6030379

[19] Y. Wu, *"A New Instance-weighting Naive Bayes Text Classifiers,"* 2018 IEEE International Conference of Intelligent Robotic and Control Engineering (IRCE), Lanzhou, pp. 198-202. 2018. doi: 10.1109/IRCE.2018.8492960

[20] Imandoust, S.B. & Bolandraftar, Mohammad. (2013). *"Application of K-nearest neighbor (KNN) approach for predicting economic events theoretical background".* Int J Eng Res Appl. 3.pp. 605-610. 2013.

[21] Guo G., Wang H., Bell D., Bi Y., Greer K. (2003) *"KNN Model-Based Approach in Classification".* In: Meersman R., Tari Z., Schmidt D.C. (eds) On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE. OTM. Lecture Notes in Computer Science, vol. 2888. Springer, Berlin, Heidelberg. 2003.

[22] S. Sun and R. Huang, *"An adaptive k-nearest neighbor algorithm,"* 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery, Yantai, pp. 91-94. 2010. doi: 10.1109/FSKD.2010.5569740

[23] Okfalisa, I. Gazalba, Mustakim and N. G. I. Reza, *"Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification,"* 2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, pp. 294-298. 2017. doi: 10.1109/ICITISEE.2017.8285514

[24] P. Surlakar, S. Araujo and K. M. Sundaram, *"Comparative Analysis of K-Means and K-Nearest Neighbor Image Segmentation Techniques,"* 2016 IEEE 6th International Conference on Advanced Computing (IACC), Bhimavaram, 2016, pp. 96-100.

[25] K. K. Paliwal and P. V. S. Rao, *"Application of k-Nearest-Neighbor Decision Rule in Vowel Recognition," in* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-5, no. 2, pp. 229-231, March 1983. doi: 10.1109/TPAMI.1983.4767378

[26] B. Wang, Yong Zeng and Yupu Yang, *"Generalized nearest neighbor rule for pattern classification,"* 2008 7th World Congress

on Intelligent Control and Automation, Chongqing, pp. 8465-8470. 2008.doi: 10.1109/WCICA.2008.4594258

[27] Gangadeep Kaur, Kamaldeep Kaur, "*Sentiment Detection from Panjabi Text using Support Vector Machine*", International Journal of Scientific Research in Computer Science and Engineering, Vol.2, Issue.6, pp.39-46, 2017 .

[28] Ghattas, Badih & Ishak, Anis. (2019). *"An Efficient Method for Variables Selection Using SVM-Based Criteria"*. Journal of Machine Learning Research - JMLR. 2019.

[29] B. Waske and J. A. Benediktsson, *"Fusion of Support Vector Machines for Classification of Multisensor Data,"* in IEEE Transactions on Geoscience and Remote Sensing, vol. 45, no. 12, pp. 3858-3866, Dec. 2007. doi: 10.1109/TGRS.2007.898446

[30] G. Zhang, *"A Modified SVM Classifier Based on RS in Medical Disease Prediction,"* 2009 Second International Symposium on Computational Intelligence and Design, Changsha, pp. 144-147. 2009. doi: 10.1109/ISCID.2009.43

[31] .M. Sewak, P. Vaidya, C. Chan and Zhong-Hui Duan, *"SVM Approach to Breast Cancer Classification,"* Second International Multi-Symposiums on Computer and Computational Sciences (IMSCCS 2007), Iowa City, IA, pp.32-37. 2017 .

[32] M. Ma, Z. Gao, J. Wu, Y. Chen and X. Zheng, *"A Smile Detection Method Based on Improved LeNet-5 and Support Vector Machine,"* 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Guangzhou,pp.446-451.2018.

[33] M. H. Selamat and H. M. Rais, *"Enhancement on image face recognition using Hybrid Multiclass SVM (HM-SVM),"* 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, pp. 424-429. 2016. doi: 10.1109/ICCOINS.2016.7783253

[34] Parveen and A. Singh, *"Detection of brain tumor in MRI images, using combination of fuzzy c-means and SVM,"* 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN), Noida, pp. 98-102. 2015 doi: 10.1109/SPIN.2015.7095308

[35] B. Sanjaa and E. Chuluun, *"Malware detection using linear SVM,"* Ifost, Ulaanbaatar, pp. 136-138. 2013 doi: 10.1109/IFOST.2013.6616872

[36] C. Wang, J. Zheng and X. Li, *"Research on DDoS Attacks Detection Based on RDF-SVM,"* 2017 10th International Conference on Intelligent Computation Technology and Automation (ICICTA), Changsha, pp. 161-165. 2017 .doi: 10.1109/ICICTA.2017.43

Authors Profile

Pragati Prakash pursuing her Bachelor of Technology (B-Tech) in Computer Science and Engineering Department, Institute of technology, Guru Ghasidas Vishvavidyalaya,A Central University, Bilaspur, C.G., India. Her research interests are Network Security, Algorithms etc.

Nidhi Ekka pursuing her Bachelor of Technology (B-Tech) in Computer Science and Engineering Department, Institute of technology, Guru Ghasidas Vishvavidyalaya, A Central University, Bilaspur, C.G., India. Her research interests are Network Security, Algorithms etc.

Manjit Jaiswal is currently working as a Assistant professor in the Computer Science and engineering Department, Institute of technology, Guru Ghasidas Vishvavidyalaya, A Central University Bilaspur, Chhattisgarh, India. He received Master of Technology (M-Tech) degree in 2012 from MANIT, Bhopal, M.P., India. His research interest fields are Parallel computing, Firewall, Network Security, Distributed Sysem, Algorithms,Routing Algorithm etc.

