SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ DERGİSİ

Sakarya University Journal of Science
SAUJS

e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |
http://www.saujs.sakarya.edu.tr/en/

Title: Vote-Based: Ensemble Approach

Authors: Abdul Ahad ABRO

Recieved: 2021-03-23 18:56:37

Accepted: 2021-05-31 09:24:51

Article Type: Research Article

Volume: 25
Issue: 3
Month: June
Year: 2021
Pages: 858-866

How to cite
Abdul Ahad ABRO; (2021), Vote-Based: Ensemble Approach. Sakarya University Journal of Science, 25(3), 858-866, DOI: https://doi.org/10.16984/saufenbilder.901960
Access link
http://www.saujs.sakarya.edu.tr/en/pub/issue/62736/901960

Sakarya University Journal of Science 25(3), 858-866, 2021

Vote-Based: Ensemble Approach

Abdul Ahad ABRO*[1]

Abstract

Vote-based is one of the ensembles learning methods in which the individual classifier is situated on numerous weighted categories of the training datasets. In designing a method, training, validation and test sets are applied in terms of an ensemble approach to developing an efficient and robust binary classification model. Similarly, ensemble learning is the most prominent and broad research area of Machine Learning (ML) and image recognition, which assists in enhancing the capability of performance. In most cases, the ensemble learning algorithm yields better performance than ML algorithms. In this regard, numerous approaches had been studied significantly and used to accomplish better yields from the existing literature; however, the outcomes of these methods are inadequate. Unlike existing methods, the proposed technique aggregates an ensemble classifier, known as vote-based, to employ and integrate the advantage of ML classifiers, which are Naive Bayes (NB), Artificial Neural Network (ANN) and Logistic Model Tree (LMT). This paper proposes an ensemble framework that aims to evaluate datasets from the UCI ML repository by adopting performance analysis. The experimental consequences reveal that the intended approach outperforms than the conventional approaches. Furthermore, the experimental outputs indicate that the suggested method provides more accurate results according to the base learner approaches in terms of accuracy rates, an area under the curve (AUC), recall, precision, and F-measure values. This method can be used for binary classification, image recognition and machine learning problems.

Keywords: Machine Learning, Artificial Neural Network (ANN), Ensemble learning, Data Mining, Classification.

1. INTRODUCTION

Machine Learning and Ensemble Learning multiple approaches intend to merge specific decisions by weighted or unweighted vote-based to classify new events as an active research area. These systems are mainly aimed towards achieving efficient results in classification rather than using a single model. Ensemble Learning is an approach of ML, which associates distinct base models to develop a single predictive model [1]. This is one of the sophisticated approaches of data assessment to pact with reflections having several datasets, as automated tools are being applied in it to locate patterns and relationships. Several methods are used in ensemble learning to progress

*Corresponding author: abdulahadabro1@gmail.com
[1]Department of Computer Engineering, Ege University, Turkey.
ORCID: https://orcid.org/0000-0002-3591-9231

prediction models [2]. It is much viable for resolving classification problems due to its robustness, highly concise prediction and measurements of variable significance. Ensemble learning is successfully deployed for its significant performance in numerous aspects, including medical, remote sensing, pattern recognition and sensors (IoT) for its magnificent outputs. NB, introduced by Chen et al. [3], proposed the particular relevant method, which chose some of the attributes to design the NB approach. The outcomes refer to the classification accuracy by maintaining efficiency, time and simplicity. ANN, suggested by Khwaja et al. [3], merged the bagging and boosting to train the model on sampling original training data in bagged-boosted ANNs. The results show the decreased variance compared to single ANN, boosted ANN, bagged ANN.

On the other hand, SVM, proposed by Nieto et al. [4], comprised of statistical learning theory with a latest class model, generates the classification values which possess the well-known accuracy of the multivariate function. SVM is a paradigm that utilizes classification algorithms for two-group challenges. It is accuracy and predictive performance on the survival of traumatic brain injuries performed significantly improved than logistic regression. SVM is a valuable approach for resolving classification and regression challenges. While the LMT [5], the combination of decision tree and logistic regression models give accurate outcomes of these algorithms; whereas, the high computational cost makes it inadmissible. The suggested method LogitBoost, with 14 benchmark datasets, provides the training time decreed while accuracy remains constant in fast incremental learning of logistic model tree. The main idea of this research is summarized in the following manner:

I) Vote-based ensemble learning method improves the binary classification performance accuracy. II) Comparative analysis of three base learners and seven datasets from the UCI ML Repository based on five evaluation criteria; accuracy (Acc), AUC, precision, recall and F-measure.

This research is organized into several sections. In section 2, related works pertain to machine and ensemble learning are presented. In section 3, the methodology is discussed in detail. Section 4, provides experimental design, the definition of the datasets, performance evaluation and results. Finally, the conclusion and future work are suggested in section 5.

2. RELATED WORK

In literature review, each ensemble and ML model has its pros and cons. Generally, its behaviour majorly depends upon the features of various suggested areas. Therefore, the performance evaluation of ensemble and ML models for vote-based assessment is significantly desired, although many assessment assignments have been agreed out by researchers, such as [6,7,8]. There are a wide variety of methods to build ensembles. In this work, we mainly focus on the vote-based ensemble method, which comprises other supervised ML algorithms. In [9], ensemble classifiers have been well researched and utilized to enhance the accuracy in multiple tasks. Many ensemble approaches, including weighted majority voting (WMV), majority voting, max combiner, mean combiner and median combiner, were introduced.

In contrast, specific classifiers can be aggregated utilizing any of these approaches. WMV is most demanding among the other method due to its theoretical usage, sensitivity and efficient results. In [10], naïve Bayes attempts to weigh up the general knowledge of classification in a multi-domain e-commerce platform. This model is designed to the immense computational efficiency of the traditional naïve and has an improved capability of classification for dealing with datasets. It enhances the performance and adaptability of the method. In other studies [11,12], a single layer of neurons was introduced between the input and output layers. The network was trained using epochs and an Adam optimizer with a default learning rate, whereas ANN is one of the methods which perform best in terms of sensitivity, followed by the SVM, decision tree and Logistic regression methods. In [13], the Logistic model tree, Random Forest (RF) and classification and regression tree (CART) were constructed using training data. As per this detailed study, all three models show valuable performances; the RF model has the maximum analytical ability in comparison to the LMT and CART models. However, there are still few state-of-the-art models, for instance [14,15], CART which has been employed for assessment rarely, and therefore should be more examined and linked in a more advanced manner. The aim of classification is to correctly predict the target class for each case in the data. Whereas in the prototype build training process, a classification algorithm co-ordinate among the principles of the predictors and the standards of the target. Distinct classification algorithms accomplish distinct methods for finding associations. These associations are prototype, which can operate to a distinct dataset in which the class is unidentified.

3. METHODOLOGY

This section provides the analysis of the suggested method, preprocessing of data and classification of algorithms used in detail.

3.1. The Analysis of Proposed Method

This system consists of several stages: datasets, base learners, comparative analysis of results, conclusion and future work shown in Figure 1. In addition, 10-fold cross-validation used for all learners and datasets to obtain generalization performance of the system is shown.

3.2. Preprocessing of Data

The ranges of the values in data preprocessing may be high. In this scenario, classification algorithms could be affected significantly or negatively by some features. Therefore, data values are normalized to [0,1] range using min-max normalization technique [16], in Equation (1)

$$\hat{x}_i = \frac{x_i - \min_{x_i}}{\max_{x_i} - \min_{x_i}} . (\max x_{new} - \min x_{new}) + \min x_{new} \tag{1}$$

For mapping value of a feature xi from the range [min(xi), max(xi)] to a new range [minxnew, maxxnew], the normalized feature x̂i is computed.

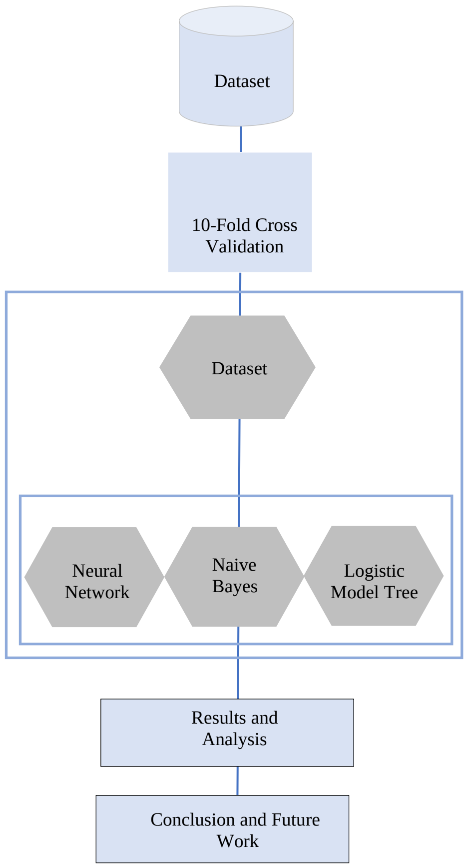

Figure 1 Overview of the proposed system

3.3. Classification of Algorithms

In this paper, a framework has been proposed for an ensemble learning method, including NB, ANN and LMT. Ensemble Learning, such as vote-based, enables one to diminish various influences such as classification error. Furthermore, combinations of many classifiers, particularly in the case of unstable classifiers, which may generate a more reliable classification than a single classifier.

The key concept of this analysis is to establish and provide data comprised of diverse attributes to present new methods related to binary classification. NB is a robust ML algorithm, which is used for predictive modelling [17]. It is an algorithm for classifying binary and multi-class problems. This approach is appropriate for binary or categorical input values. ANN is an information processing paradigm, which is considered as Universal Function Approximators. It is a modest and very influential process. It is considered as the class of feedforward artificial neural network and composed of a highly interconnected processing model to solve the classification and regression model [18]. Whereas [19], SVM is a robust supervised ML algorithm, which is used for classification and regression challenges. It is a versatile and high-dimensional space-effective algorithm. LMT is defined as a set of logistic regression and decision tree learning. It is creating a more accurate model than C4.5 and CARD in real-world datasets. It is also a well-known enhanced decision tree learner [20].

4. EXPERIMENTAL DESIGN

We describe and present the experimental process, evaluation measures and experimental results for this study in the subsections:

4.1. The Experimental Process

In the experimental process, seven datasets have been used ML Repository for classification schemes [21]. The number of instances, attributes, and classes for each dataset are presented. The specifications of these datasets are demonstrated in Table 1. The performance of our algorithm is being associated to several other state-of-the-art learning schemes on datasets and shows that it produces accurate outcomes.

All experiments are performed on a total of 3 ML classifiers by using WEKA (Waikato Environment for Knowledge Analysis) ML toolkit and JAVA programming language [22]. We utilized default parameter values for all classifiers. We carry out 10-fold cross-validation to all datasets to yield reliable results. This cross-validation is imposed on the actual dataset casually segregated into 10 similar sized sets, one of which is used as test validation, while the remaining sets are used for training operations. The process is repeated 10 times and considered the averages of the results.

Dataset characteristics are evaluated concerning the attributes and the number of instances. These datasets are typically used to solve ML-related concerns. Datasets are chosen according to their distinct parameters from the Repository. It is determined by investigating the appropriate data or datasets which are being utilized in the various research papers related to ML issues. The proposed vote-based ensemble learning technique has been introduced for this process. This method utilized the imbalanced classification problems of binary (two-class) where the positive case, such as (class 1), is taken as an unusual and negative case (class 0) is taken as normal. In this work, three different ML approaches have been carried out along with the ensemble learning method, which is considered appropriate for this mechanism. However, the performance metrics are calculated based on the datasets according to binary classification problems.

Table 1
Datasets Specifications

Datasets	Attributes	Instances	Classes
Audiology	69	226	24
Balance Scale	4	625	3
Credit Approval	15	690	2
Heart (Statlog)	13	270	2
Ionosphere	34	351	2
Sonar	60	208	2
Zoo	17	101	7

4.2. Analysis of Algorithm

The hybrid nature of algorithm produces dynamically efficient outcomes with respect to different ensemble classifiers and target class for each case in the data. The result of vote-based ensemble approach is progressive. This approach provides more beneficial and efficient outputs by using the advantages of these algorithms.

5. MEASURES OF EVALUATION

This section describes the five performance evaluation measures of the suggested method, consisting of accuracy, AUC, precision, recall and F-measure.

Accuracy represents how near a measurement is to an identified or accepted figure. It is further defined in Equation (2).

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

In equation 2, TN, FN, FP and TP show the number of True Negatives, False Negatives, False Positives and True Positives. AUC is the area under the ROC curve for classifier performance. Its value will always be between 0.0 and 1.0. ROC graphs are two-dimensional graphs. In this curve, the TP rate is plotted on the Y-axis and FP rate is plotted on the X-axis [23]. If AUC value is close to 1, the classifier is more reliable and better than a random classifier.

Precision is a positive analytical value [8]. Precision defines how reliable measurements are, although they are farther from the accepted value. The precision is shown in Equation (3).

$$\text{Precision}=\frac{\text{TP}}{\text{TP}+\text{FP}} \tag{3}$$

The recall is the hit rate [24]. The recall is the reverse of precision; it calculates false negatives against true positives. The recall is illustrated in Equation (4).

$$\text{Recall}=\frac{\text{TP}}{\text{TP}+\text{FN}} \tag{4}$$

F-measure can be defined as the weighted average [8] of precision and recall. This rating considers both false positives and false negatives. The F-measure is presented in Equation (5).

$$\text{F}-\text{measure}=\frac{2}{1/\text{precision}+1/\text{recall}} \tag{5}$$

In the weighting operation, these criteria are adjusted by the reference class prevalence proportionally in the data. Tables 2-8 present accuracy, AUC, precision, recall and F-measure individual weighted values for all datasets [25,26].

Table 2
Weighted values for audiology dataset

Audiology (Standardized)					
Methods	Acc(%)	AUC	Precision	Recall	F-Measure
NB	73.4513	0.943	0.750	0.943	0.500
ANN	83.1858	0.978	1.000	0.832	0.857
SVM	81.8584	0.941	1.000	0.819	0.857
LMT	**84.0708**	0.957	1.000	0.841	0.667

Table 3
Weighted values for balance scale dataset

Balance Scale					
Methods	Acc(%)	AUC	Precision	Recall	F-Measure
NB	90.4000	0.971	0.901	0.904	0.938
ANN	**90.7200**	0.977	0.916	0.907	0.911
SVM	87.6800	0.879	0.868	0.877	0.909
LMT	89.7600	0.981	0.859	0.898	0.873

Table 4
Weighted values for credit approval

Credit Approval					
Methods	Acc(%)	AUC	Precision	Recall	F-Measure
NB	77.6812	0.896	0.793	0.777	0.769
ANN	83.6232	0.895	0.836	0.836	0.836
SVM	**84.9275**	0.856	0.861	0.849	0.850
LMT	84.7826	0.920	0.852	0.848	0.848

Table 5
Weighted values for heart (statlog) dataset

Heart (Statlog)					
Methods	Acc(%)	AUC	Precision	Recall	F-Measure
NB	83.7037	0.898	0.837	0.837	0.837
ANN	78.1481	0.839	0.784	0.781	0.782
SVM	**84.0741**	0.837	0.841	0.841	0.840
LMT	83.3333	0.897	0.833	0.833	0.833

Table 6
Weighted values for ionosphere dataset

Ionosphere					
Methods	Acc(%)	AUC	Precision	Recall	F-Measure
NB	82.6211	0.935	0.842	0.826	0.829
ANN	91.1681	0.915	0.918	0.912	0.909
SVM	88.604	0.853	0.891	0.886	0.883
LMT	**93.1624**	0.922	0.934	0.932	0.930

Table 7
Weighted values for sonar dataset

Sonar					
Methods	Acc(%)	AUC	Precision	Recall	F-Measure
NB	67.7885	0.800	0.704	0.678	0.673
ANN	**82.2115**	0.878	0.822	0.822	0.822
SVM	75.9615	0.758	0.759	0.760	0.759
LMT	77.8846	0.846	0.779	0.779	0.779

Table 8
Weighted values for zoo dataset

Zoo					
Methods	Acc(%)	AUC	Precision	Recall	F-Measure
NB	95.0495	1.000	0.963	0.950	0.947
ANN	**96.0396**	0.993	0.960	0.960	0.958
SVM	96.0396	0.984	0.960	0.960	0.958
LMT	94.0594	0.997	0.941	0.941	0.939

Table 9
Our proposed vote-based approach

Proposed Vote-Based Classifier ANN, LMT, NB						
Classifier	Acc (%)	Impr. (%)	AUC	Precision	Recall	F-Measure
Audiology	**85.3982**	**1.3274**	**0.979**	0.914	0.854	**0.922**
Balance Scale	*90.7200	**0.0000**	**0.989**	0.891	*0.907	0.880
Credit Approval	**85.6522**	**0.7247**	0.918	0.857	**0.857**	**0.856**
Heart (Statlog)	**84.8148**	**0.7407**	**0.900**	**0.848**	**0.848**	**0.848**
Ionosphere	**94.5869**	**1.4245**	**0.952**	**0.947**	**0.946**	**0.945**
sonar	**82.6923**	**0.4808**	**0.893**	**0.827**	**0.827**	**0.827**
Zoo	**97.0297**	**0.9901**	0.999	**0.969**	**0.970**	**0.968**

- * Indicates the similar performance results concerning base learner.
- High Acc, AUC, Precision, Recall and F-measure are shown in Bold, while the greyed shows insufficient results.
- Impr. represents improvement according to the best results of Tables 2-8.

6. EXPERIMENTAL RESULTS

To sum up, Tables 2-8, have been designed according to the diverse datasets concerning the numerous approaches of ML in terms of different specifications. In Table 2, LMT has better outcomes, which provides 84.0708% Acc in comparison to others. Likely, in Table 3, ANN indicates 90.7200% Acc adequate consequences. Similarly, in Table 4, the SVM presents 84.9275% Acc effective results. Likewise, in Table 5, the SVM illustrates the 84.0741% Acc productive outcomes. In Table 6, LMT has shown the 93.1624% Acc result. Furthermore, in Table 7, ANN represents the 82.2115% Acc output. However, in the end, ANN shows a 96.0396% Acc result in Table 8. In general, ANN has more successive consequences than SVM; whereas the SVM provides more effective outputs than NB in most of the datasets. On the other hand, LMT has also provided satisfactory results to some extent, which is illustrated in Tables 2 and 8.

In Table 9, the vote-based ensemble learning method has been applied, in which the model is trained with the combined prediction preceding model. The vote-based has been set as a classifier and experienced the diverse datasets with numerous methods like ANN, LMT and NB in the given order. The Audiology, Credit Approval, Heart (Statlog), Ionosphere, Sonar and Zoo datasets have significant outputs regarding the accuracy, recall, precision, AUC and F-measure parameters in Table 9; however, the Balance Scale dataset show a similar outcome for Table 3.

In Table 9 demonstrates the comparison of all dataset results concerning our proposed vote-based meta-ensemble learning method. As it is clearly shown in Table 9, a Meta-ensemble classifier, vote with three base learners (namely, ANN, LMT and NB), provides highly accurate outcomes in comparison to others. Moreover, it is analyzed that when the vote-based ensemble method combines with ANN, it provides more accurate outcomes than ANN; whereas ANN does not provide better outcomes when applied individually.

7. CONCLUSION AND FUTURE WORK

Based on the numerical and experimental outcomes, the core findings of this research effort can be summarized as follows:

In this paper, the ensemble learning classifier is being widely used due to its effectiveness and high performance in various fields such as ML and pattern recognition. In this study, the vote-based was correlated with the standard implementations of NB, ANN, SVM and LMT. The experimental results with 07 datasets indicate the outperformance of the model among all the four methods by a large margin. It ensured that vote-based has a similar diversity-accuracy pattern to Neural Network but is more accurate and diverse than it. A reasonable performance has been achieved by utilizing our applied ensemble techniques when compared with similar studies in the literature. A marginal improvement has been fetched statistically and shows significant differences in favor of the implemented method. Many machine learning algorithms, on the other hand, are incompetent to deliver good results since they are dependent relative on datasets. The sensitivity of a algorithm can be considerably influenced by the size of the training and test sets. In the future, other hybridization of ensemble learning methods will be utilized for performance improvement.

Acknowledgements

No acknowledgements has been declared by the author.

Funding

The author received no financial support for the research, authorship or publication of this study.

The Declaration Of Conflict Of Interest/ Common Interest

No conflict of interest or common interest has been declared by the author.

The Declaration of Ethics Committee Approval

This work does not require ethics committee permission or any special permission.

The Declaration of Research And Publicatıon Ethics

The author of the paper declares that he complies with the scientific, ethical and quotation rules of SAUJS in all processes of the paper and that he does not make any falsification on the data collected. In addition, she declares that Sakarya University Journal of Science and its editorial board have no responsibility for any ethical violations that may be encountered, and that this study has not been evaluated in any academic publication.

REFERENCES

[1] M. A. Shehab and N. Kahraman, “A weighted voting ensemble of efficient regularized extreme learning machine,” *Comput. Electr. Eng.*, vol. 85, 2020.

[2] J. Cao, S. Kwong, R. Wang, X. Li, K. Li, and X. Kong, “Class-specific soft voting based multiple extreme learning machines ensemble,” *Neurocomputing*, vol. 149, no. Part A, pp. 275–284, 2015.

[3] A. S. Khwaja, A. Anpalagan, M. Naeem, and B. Venkatesh, “Joint bagged-boosted artificial neural networks: Using ensemble machine learning to improve short-term electricity load forecasting,” *Electr. Power Syst. Res.*, vol. 179, no. October 2019, p. 106080, 2020.

[4] P. J. G. Nieto, E. García-gonzalo, and J. C. Á. Antón, “Journal of Computational and Applied A comparison of several machine learning techniques for the centerline

segregation prediction in continuous cast steel slabs and evaluation of its performance," *J. Comput. Appl. Math.*, vol. 330, pp. 877–895, 2018.

[5] S. Lee and C. H. Jun, "Fast incremental learning of logistic model tree using least angle regression," *Expert Syst. Appl.*, vol. 97, pp. 137–145, 2018.

[6] H. Liu and L. Zhang, "Advancing Ensemble Learning Performance through data transformation and classifiers fusion in granular computing context," *Expert Syst. Appl.*, vol. 131, pp. 20–29, 2019.

[7] S. Shen, M. Sadoughi, M. Li, Z. Wang, and C. Hu, "Deep convolutional neural networks with ensemble learning and transfer learning for capacity estimation of lithium-ion batteries," *Appl. Energy*, vol. 260, no. December 2019, p. 114296, 2020.

[8] A. A. ABRO, E. TAŞCI, and A. UGUR, "A Stacking-based Ensemble Learning Method for Outlier Detection," *Balk. J. Electr. Comput. Eng.*, vol. 8, no. 2, pp. 181–185, 2020.

[9] A. A. Aburomman, M. Bin, and I. Reaz, "A novel SVM-kNN-PSO ensemble method for intrusion detection system," vol. 38, pp. 360–372, 2016.

[10] F. Xu, Z. Pan, and R. Xia, "E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework," *Inf. Process. Manag.*, no. February, p. 102221, 2020.

[11] S. S. Panesar, R. N. D. Souza, F. Yeh, and J. C. Fernandez-miranda, "Machine Learning Versus Logistic Regression Methods for 2-Year Mortality Prognostication in a Small, Heterogeneous Glioma Database," *World Neurosurg. X*, vol. 2, p. 100012, 2019.

[12] A. A. Abro, M. Alci, and F. Hassan, "Theoretical Approach of Predictive Analytics on Big Data with Scope of Machine Learning."

[13] W. Chen *et al.*, "A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility," *Catena*, vol. 151, pp. 147–160, 2017.

[14] A. Kumar and A. Halder, "Ensemble-based active learning using fuzzy-rough approach for cancer sample classification," *Eng. Appl. Artif. Intell.*, vol. 91, no. December 2019, p. 103591, 2020.

[15] X. Zheng, W. Chen, Y. You, Y. Jiang, M. Li, and T. Zhang, "Ensemble deep learning for automated visual classification using EEG signals," *Pattern Recognit.*, vol. 102, p. 107147, 2020.

[16] T. Classification and B. K. Singh, "Investigations on Impact of Feature Normalization Techniques on Investigations on Impact of Feature Normalization Techniques on Classifier ' s Performance in Breast Tumor Classification," no. April 2015, pp. 10–15, 2017.

[17] L. Fan, K. L. Poh, and P. Zhou, "A sequential feature extraction approach for naïve bayes classification of microarray data," *Expert Syst. Appl.*, vol. 36, no. 6, pp. 9919–9923, 2009.

[18] E. Lella and G. Vessio, "Ensembling complex network 'perspectives' for mild cognitive impairment detection with artificial neural networks," *Pattern Recognit. Lett.*, vol. 136, pp. 168–174, 2020.

[19] R. Moraes, J. F. Valiati, and W. P. Gavião Neto, "Document-level sentiment classification: An empirical comparison between SVM and ANN," *Expert Syst. Appl.*, vol. 40, no. 2, pp. 621–633, 2013.

[20] N. Landwehr, M. Hall, and E. Frank, "Logistic model trees," *Mach. Learn.*, vol. 59, no. 1–2, pp. 161–205, 2005.

[21] UCI Machine Learning Repository, 2018, https://archive.ics.uci.edu/ml/index.php

[22] E. Frank, M. A. Hall, I. H. Witten, and T. Weka, "Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for 'Data Mining: Practical Machine Learning Tools and Techniques', Morgan Kaufmann, Fourth Edition, 2016.," p. 2016, 2016.

[23] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006.

[24] L. A. Bull, K. Worden, R. Fuentes, G. Manson, E. J. Cross, and N. Dervilis, "Outlier ensembles: A robust method for damage detection and unsupervised feature extraction from high-dimensional data," *J. Sound Vib.*, vol. 453, pp. 126–150, 2019.

[25] T. Fawcett, "ROC graphs: Notes and practical considerations for researchers," *Mach. Learn.*, vol. 31, no. 1, pp. 1–38, 2004.

[26] A. A. Abro, M. A. Yimer, and Z. Bhatti, "Identifying the Machine Learning Techniques for Classification of Target Datasets," *Sukkur IBA J. Comput. Math. Sci.*, vol. 4, no. 1, 2020.