World Academy of Science, Engineering and Technology
International Journal of Medical and Health Sciences
Vol:17, No:12, 2023

Artificial Intelligence-Based Detection of Individuals Suffering from Vestibular Disorder

D. Hişam, S. İkizoğlu

***Abstract*—**Identifying the problem behind balance disorder is one of the most interesting topics in medical literature. This study has considerably enhanced the development of artificial intelligence (AI) algorithms applying multiple machine learning (ML) models to sensory data on gait collected from humans to classify between normal people and those suffering from Vestibular System (VS) problems. Although AI is widely utilized as a diagnostic tool in medicine, AI models have not been used to perform feature extraction and identify VS disorders through training on raw data. In this study, three ML models, the Random Forest Classifier (RF), Extreme Gradient Boosting (XGB), and K-Nearest Neighbor (KNN), have been trained to detect VS disorder, and the performance comparison of the algorithms has been made using accuracy, recall, precision, and f1-score. With an accuracy of 95.28 %, Random Forest (RF) Classifier was the most accurate model.

***Keywords*—**Vestibular disorder, machine learning, random forest classifier, k-nearest neighbor, extreme gradient boosting.

Open Science Index, Medical and Health Sciences Vol:17, No:12, 2023 publications.waset.org/10013416.pdf

I. Introduction

THE inner ear's VS, which is composed of microscopic components, detects head motion and body gravity. The body can maintain balance, maintain proper spatial orientation when moving, and correctly process visual images while moving thanks to the processing of this information by the vestibular regions of the brain [1]. To find anomalies in human balance systems, doctors primarily employ portable balance devices and record body sways nowadays [2]. This system has concerns with time consumption and usability. However, computerized disease classification can save a lot of work and result in a rapid and precise diagnosis of VS issues. Therefore, a study was conducted to examine the accuracy of three ML algorithms applied to the sensory data taken from patients with VS diseases and healthy individuals to diagnose VS from 51 healthy and VS samples using data collected from four pressure sensors on each foot [3]. Another study was conducted as a follow-up to the first one to develop a ML algorithm that could be used to analyze sensory data on gait acquired from subjects to detect VS disorders [4].

AI approaches are used to analyze data and images and recognize patterns by developing algorithms that help doctors quickly and accurately diagnose a specific disease or disorder. Considering the countless AI applications that have been achieved recently, which made significant developments in many different scientific and industrial sectors, it is evident that artificial intelligence continues to revolutionize various fields with its transformative capabilities. Additionally, these algorithms are always being learned, which improves the accuracy of the diagnoses. However, there are times when medical professionals must simultaneously consider the patient's symptoms, treatment alternatives, potential side effects, an illness with comparable symptoms, past medical history, and a few other considerations. By evaluating a vast amount of data and assuring a thorough understanding of patient health records, AI technology provides a way for doctors to assist patients [5]. Therefore, academics and clinicians alike are interested in the use of Wearable Sensors to measure normal gait metrics. A system for activity recognition, including gait, was created by Zebin et al. [6] and consisted of five IMU sensors worn on the lower back, thighs, and shanks. A Convolutional Neural Networks (CNN)-based model is used to automatically extract the features from time-series raw data, and this method is more accurate than manually writing the features or using shallow learning. In a different study, Ordóez and Roggen [7] used 12 accelerometers close to the limb joints in addition to 7 IMUs on the chest, arms, and legs.

Pataky and colleagues showed great subject recognition accuracy (99.6%) using dynamic plantar pressure data, picture processing, and feature extraction. They found that foot pressure-based identification might be widely used in the security and healthcare industries [8]. Their study also highlighted the distinction between inter-subject pressure patterns. Quian's team combined a 1D pressure profile with 2D position trajectories of the centers of pressure of both feet to produce a 3D centroid of the ground reaction force vector (COP) trajectory over a footstep. When the COP data are paired with additional variables such the mean pressure and stride length, the average identification rate is 92.3% [9]. Others have employed pressure data in conjunction with gait factors including stride length, stride cadence, and the proportion of time spent on the toes to the heels to achieve high subject recognition rates [10], [11]. A technique for estimating age in an old population was also developed using changes in foot pressure distribution [12]. A few sensorized insoles have also been developed for use in sports, gait measurement, the prevention of diabetic ulcers, and other purposes [13], [14].

In [15], a classification of sensor data for gait analysis yields an overall accuracy of 80% when the RF model is used, as opposed to other ML methods like KNN and ANN. To evaluate the role of gait analysis in Parkinson's disease detection, nine classification models with different metrics were used, including Naive Bayes (NB), KNN with k = 3, Support Vector Machine (SVM) with linear kernel, Decision Tree (C4.5),

D. H. is with Istanbul Technical University Mechatronics Engineering Dept. Turkey (phone: +905385629753; e-mail: elsisy20@ itu.edu.tr).

S. İ. is with Istanbul Technical University Control and Automation Engineering Dept. Turkey (e-mail: ikizoglus@itu.edu.tr).

Open Science Index, Medical and Health Sciences Vol:17, No:12, 2023 publications.waset.org/10013416.pdf

Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Adaboost (ADA), Subspace Technique (SUB), and RF with 50 trees. For data collected from a smartphone-based gait analyzer, Probabilistic Neural Network (PNN) was able to achieve the best accuracy of 91.13% when compared with eight ML models in [16].

In this study, three ML models were created to recognize VS from sample data in [4], [5]. Each of the models presented in the literature review has been used independently on relevant applications with positive outcomes. So, all these models have been compared in this paper. Implementing ML models such as the RF Classifier, KNN and XGB Classifier to detect VS diseased subjects is the main purpose of this paper.

II. Methodology

A. Data

Sensory data collection was made using information from 51 people, 20 of whom were healthy and 31 of whom had VS balance disorder. The age range for the healthy group was 35–62, with a mean of 41.6 and a standard deviation of 6.8, and the VS group's range was 48–64, with a mean of 54.8 and a standard deviation of 5.7. The healthy sample's weight ranges from 50.4 to 78, with a mean of 61.3 and a standard deviation of 8.8, and from 68 to 90.7 for the VS samples, with a range of 75.6 to 8.8. Four pressure sensors were used for each foot, as seen in Fig. 1.

B. Data Preprocessing

On the data, two normalization procedures were applied. First, all sensor values were normalized to the same scale to ensure that the subject's weight, age, and shoe size had no bearing on the results. We may compare and analyze the pressure sensory data from several participants in this way. The steps listed below were taken to obtain this:

1. Find the sum of 4 pressure sensor data for each foot at every time instance (sum of A0-A3 for the right foot and A4-A7 for the left foot at each time point).
2. Find the max value for each sum vector.
3. Normalize each sensor using the max value.

For example, for A0 normalization:

1. Find the sum of A0 A1 A2, and A3 for each time instance, then generate a 1 x length(A0) sum vector.
2. Find max (sum vector).
3. Normalized A0 = A0/max (sum vector).

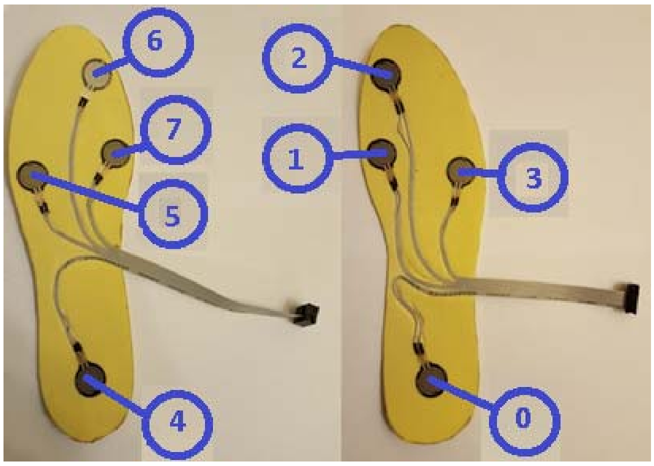

Fig. 1 Pressure sensor position in the both right and left foot

Curves drawn using a set of sample data from sensor A0 are shown in Fig. 2.

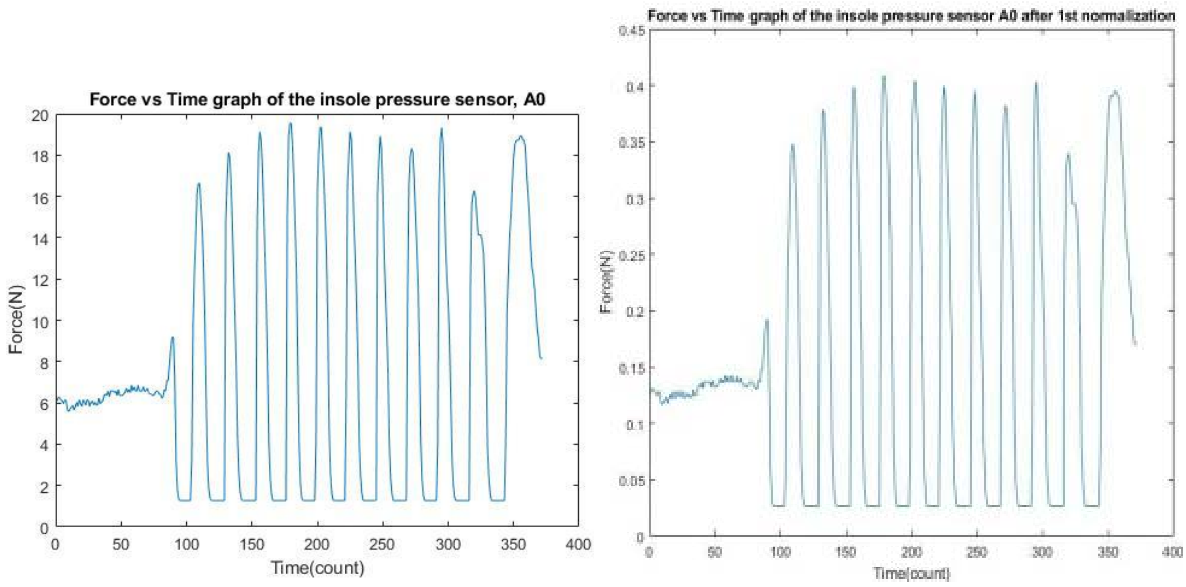

Fig. 2 Sample sensor A0 data before and after the first Normalization

World Academy of Science, Engineering and Technology
International Journal of Medical and Health Sciences
Vol:17, No:12, 2023

Open Science Index, Medical and Health Sciences Vol:17, No:12, 2023 publications.waset.org/10013416.pdf

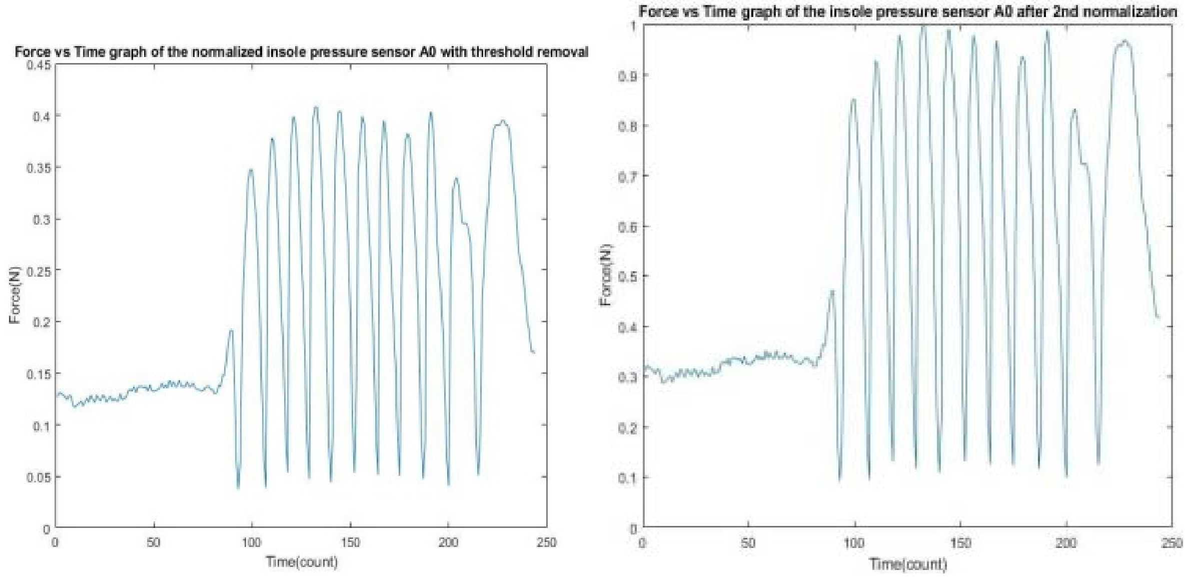

Fig. 3 Sample sensor A0 data after removing threshold and 2nd Normalization

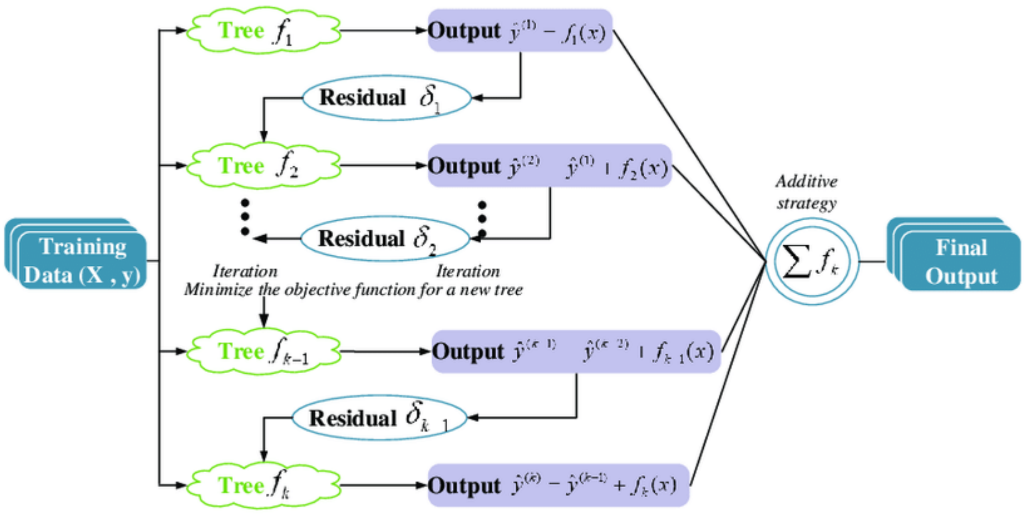

Fig. 4 The structure of extreme gradient boosting

Data were first normalized to eliminate the impact of weights, and then a second normalization was performed to scale the numbers from 0 to 1. The data relating to the duration of the foot in the air were eliminated before the second normalization phase. Data over a threshold value determined in a method to remove misleading data were taken into consideration for this. The second normalization was then carried out by dividing each value by the maximum of the whole data and utilizing the maximum of the relevant sensor data. For instance, the second step's normalization of the sensor A0 data for a gait cycle was A0 2ndnorm = A0/max (A0). The fully normalized data based on the raw data are displayed in Fig. 3.

C. Algorithms

Starting with various ML models that were effective in the same classification of medical data as seen in the introduction, we trained our data on them. Using the same data, all models have been trained to distinguish between healthy and VS individuals.

1) Extreme Gradient Boosting (XGB Classifier)

The gradient boosting concept is applied by XGBoost, one of the boosted tree algorithms. Compared to earlier gradient boosting techniques, XGBoost uses a more regularized model formalization to prevent data over-fitting, enhancing its performance [24]. The learning rate hyperparameter, which determines how rapidly the model learns, was employed in this study and was set to 0.1.

The number of estimators, or the number of trees you wish to build before combining predictions or forecasts, is 280 in this case. The higher the number, the slower the algorithm runs but the better it works.

2) Random Forest Classifier

The RF algorithm is a popular tree-based ensemble learning

World Academy of Science, Engineering and Technology
International Journal of Medical and Health Sciences
Vol:17, No:12, 2023

Open Science Index, Medical and Health Sciences Vol:17, No:12, 2023 publications.waset.org/10013416.pdf

model and the bagging-type ensemble it employs has demonstrated remarkable effectiveness in addressing complex classification and regression tasks across diverse domains [17]. RF splits each node using the best among a group of randomly selected predictors at that node, in contrast to other conventional trees [18]. This additional layer of randomization is what gives RF a better resilience to over-fitting [19] Each DT in the RF separates a class prediction, as seen in Fig. 5, and the class with the greatest votes is selected as the model prediction. Any of the component models will function more effectively as a committee of numerous uncorrelated trees than as an individual tree. Any tree in the RF has the potential to offer a misleading forecast, even though the other trees may produce the correct ones. The RF constructs several DTs, and they are ultimately merged to form an absolute and stable value that is mostly used when making predictions about the class and training it [20]-[23]. Thus, this group of trees provides a better outlook. The model utilized in this study was trained using just two classes, and its hyperparameters are 100 as the number of estimators.

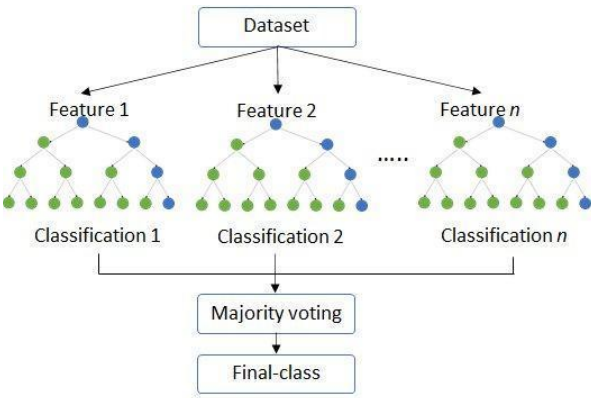

Fig. 5 Model structure of RF classifier

3) K-Nearest Neighbor

This instance-based learning technique computes each point using the majority vote of the nearest neighbors to store the instances of the training data and assigns the target class with the most representative nearest neighbors. It does not create a model to make predictions. It is sometimes referred to as a lazy classifier because predictions instead rely on most k-nearest points [25].

D. Models' Performance

Recall, Precision, and F1 score were used to evaluate the accuracy of the model performance. Accuracy is the ratio of correctly predicted events to all input samples. The model's performance in counting true positives among all accurately predicted positive outcomes is measured using a precision score. A model's performance in reliably counting true positives among all real positive values is measured using the recall score. When choosing between the precision and recall scores can impair the model's capacity to detect false positives and false negatives with high frequency, respectively, the F1-score, which is the harmonic mean of the precision and recall scores, is used as a metric. Additionally, a confusion matrix representing true positives, false positives, true negatives, and false negatives was evaluated for each model.

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \tag{1}$$

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

$$F1\ Score = \frac{Precision*Recall}{Precision+Recall} \tag{4}$$

as TP, TN, FP, FN are True Positive, True Negative, False Positive and False Negative.

III. Results and Discussion

A. Models' Evaluation

With all the hyperparameters mentioned in the algorithm section, which have been determined using trial and error several times to approach the most efficient parameters, ML models results are discussed in this section. The best performance was achieved by RF with an accuracy of 95.28% after training only on raw data without any manual feature extraction as shown in Fig. 6, in which all the model accuracy is represented.

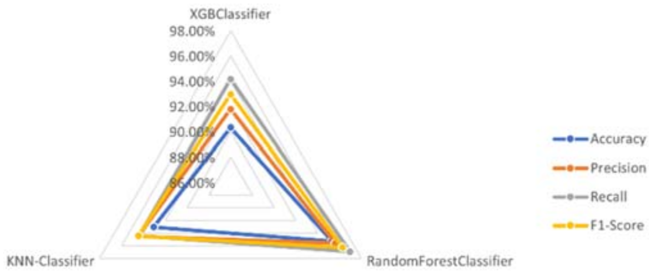

Fig. 6 Models Performance

Table I shows all the accuracies, Precision, F1-scores and Recall for all the models.

TABLE I
Models Performance

Model	Accuracy	Recall	Precision	F1 Score
XGB	90.33%	94.16%	91.78%	92.95%
RF	**95.28%**	**96.94%**	**95.59%**	**96.26%**
KNN	93.05%	94.39%	94.50%	94.45%

Results given in both Table I and Fig. 6 show that the best model with the best accuracy, precision, recall and F1-score is RF.

IV. Conclusion

This study has made important advancements in applying AI algorithms such as ML models to sensory data on gait analysis from humans to diagnose diseases of the VS. In this study, VS disorder was detected using 3 different ML models (RF

Classifier, XGB and KNN), with RF Classifier as the most accurate model. The following stage of work is applying feature extraction codes prior to training models, using simple DL models to train data on it, then training compound CNN-LSTM model using the same data.

References

[1] Khan, Sarah, and Richard Chang. "Anatomy of the vestibular system: a review." NeuroRehabilitation 32.3 (2013): 437-443.

[2] D. Basta, M. Rossi-Izquierdo, A. Soto-Varela, A. Ernst, Mobile posturography: posturographic analysis of daily-life mobility, Otol. Neurotol. 34 (2) (2013) 288–297.

[3] Heydarov, Saddam, et al. "Performance comparison of ML methods applied to motion sensory information for identification of vestibular system disorders." 2017 10th International Conference on Electrical and Electronics Engineering (ELECO). IEEE, 2017.

[4] İkizoğlu Serhat, and Saddam Heydarov. "Accuracy comparison of dimensionality reduction techniques to determine significant features from IMU sensor-based data to diagnose vestibular system disorders." Biomedical Signal Processing and Control 61 (2020): 101963.

[5] Khera, Preeti, and Neelesh Kumar. "Role of machine learning in gait analysis: a review." Journal of Medical Engineering & Technology 44.8 (2020): 441-467.

[6] T. Zebin, P. J. Scully, and K. B. Ozanyan, "Human activity recognition with inertial sensors using a deep learning approach," in Proc. IEEE Sensors, Oct./Nov. 2016, pp. 1–3. doi: 10.1109/ ICSENS.2016.7808590.

[7] F. J. Ordóñez and D. Roggen, "Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition," Sensors, vol. 16, no. 1, p. 115, 2016. doi: 10.1109/ICSENS.2016. 7808590.

[8] Pataky, Todd C., et al. "Gait recognition: highly unique dynamic plantar pressure patterns among 104 individuals." Journal of The Royal Society Interface 9.69 (2012): 790-800.

[9] Qian, Gang, Jiqing Zhang, and Assegid Kidané. "People identification using gait via floor pressure sensing and analysis." European Conference on Smart Sensing and Context. Springer, Berlin, Heidelberg, 2008.

[10] Yun, Jaeseok. "User identification using gait patterns on UbiFloorII." Sensors 11.3 (2011): 2611-2639.

[11] Takeda, Takahiro, et al. "Footage estimation by gait sole pressure changes." 2010 IEEE International Conference on Systems, Man and Cybernetics. IEEE, 2010.

[12] Crea, Simona, et al. "A wireless flexible sensorized insole for gait analysis." Sensors 14.1 (2014): 1073-1093.

[13] Ostadabbas, Sarah, et al. "Sensor architectural tradeoff for diabetic foot ulcer monitoring." 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2012.

[14] Naito, Yutaro, et al. "Quantification of gait using insole type foot pressure monitor: clinical application for chronic hemiplegia." Journal of UOEH 36.1 (2014): 41-48.

[15] McCalmont, Graham, et al. "eZiGait: toward an AI gait analysis and assistant system." 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2018.

[16] Altilio, Rosa, et al. "A comparison of machine learning classifiers for smartphone-based gait analysis." Medical & Biological Engineering & Computing 59.3 (2021): 535-546.

[17] R. Punnoose and P. Ajit, Prediction of Employee Turnover in Organizations using Machine Learning Algorithms. International Journal of Advanced Research in Artificial Intelligence (IJARAI), Vol. 5, No. 9, 1-5, 2016.

[18] A. Liaw and M. Wiener, Classification and Regression by Random Forest, R news, 2(3), 18-22, 2002.

[19] L. Breiman, Random forests. Machine Learning, 45(1), 5–32, 2001.

[20] G. Biau and E. Scornet, ''A random forest guided tour,' TEST, vol. 25, no. 2, pp. 197–227, Jun. 2016.

[21] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd ed. Canada: O'Reilly Media, 2019.

[22] R. Genuer, Random Forests with R (Use R!), 1st ed. Paris, France: Springer, 2020.

[23] A. Liaw and M. Wiener, ''Classification and regression by randomForest,'' R Project, vol. 3, no. 3, pp. 18–22, 2003.

[24] N. Alexy and K. Alois, Gradient Boosting Machines: A Tutorial. Frontiers in Neurorobotics. Vol 7(21) pp 3, 2013.

[25] Z hang H. The optimality of Naive Bayes. Proc Seventeenth Int Florida Artif Intell Res Soc Conf FLAIRS 2004. 2004;2: 562–567.View Article Google Scholar**.**

Open Science Index, Medical and Health Sciences Vol:17, No:12, 2023 publications.waset.org/10013416.pdf