Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(6), 928-939 | ISSN:2251-6727


# Foot Posture Gait Values Analysis Using Xgboost Algorithm for Cerebellar Ataxia

**Shanmuga Sundari M[1], Dr. Vijaya Chandra Jadala [2]**

[1]Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswarm. Andhra Pradesh, India

[2]Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswarm. Andhra Pradesh, India

***(Received: 07 October 2023 Revised: 12 November Accepted: 06 December)***

**KEYWORDS**

Cerebellar Ataxia,
Gait Values,
Machine Learning,
Motion effect,
Perdiction.

**ABSTRACT**:

Patients with Cerebellar Ataxia (CA) may benefit from the use of predicting systematic flow of motions with analysis of gait (AoG) prediction. The importance of this research is to identify the real model for evaluate the poor gait pattern in the human movements. The foot positions are captured and used for this research. The machine learning model will predict the neurology disease using the foot gait values. The accelerometer was attached to the patients to capture the walking positions. 12 different walking styles were carried out with the walking speed of 0.6 to 1.7 m/s. The values and motion movements are carried for gait analysis to find the disease prediction. Two AoG prediction models based on step-based abnormal gait patterns and conventional AoG generated from the signals sensing features were built using XGBoost to assess whether the use of the unsteady gait characteristics can more effectively identify AoG. Compare with the existing model, Ensemble method is providing better accuracy for foot gait value analysis.

## 1 Introduction

One important area of medical expertise is neurological specialization [1]. The brain instructs the body on how to respond to events. With this research, we can pinpoint the activity issue and determine the nervous system's capacity. A disruption in a person's activity rhythm may result in neurodegenerative disorders. Brain, spine, and nerve damage are the focus of neurosurgery. Our specialists use neurosurgery to treat neurological diseases. Finding activity patterns [2] in the medical world is difficult. We must observe the patient's motions in order to pinpoint the condition [3]. Frequently, we are unable to pinpoint the issue or obtain reliable results. Any gap in a person's medical care has the potential to be fatal. Neurologists will evaluate the complexity of disorders related to the nervous system. It helps to assess the disease's intensity, which influences treatment. For humans, the ability of the brain is more critical. The activity pattern shows the intensity of the CA condition. It will lead to lowering mortality and minimize the critical surgery. A multidisciplinary team of neuro specialists are working toward this aim.

Given the prevalence of oculomotor disorders in ataxia, quantitative measurements of saccades may be helpful for both screening for the disease and monitoring its progression. Unfortunately, eye tracking is not widely used since it requires specialised, expensive equipment as well as skill to use the devices. Due to these limitations, clinical studies of eye movement disorders in neurologic diseases often only involve a few subjects at one point in time. Consumer-grade electronic devices that use eye tracking offer the chance to improve accessibility and lighten the strain of quantitative oculomotor evaluations for both screening and intensity tracking reasons, and may enable monitoring in the home environment.

Artificial intelligence has significant advantages for the healthcare industry [4]. It assists in saving numerous lives by correctly predicting the disease. Several algorithms are run on the activity data, and the results are then analysed. The analysis's findings are reliable and accurate for forecasting. We use the Logistic Regression [5], Decision Tree (DT) [6], Random Forest (RF) [7] and K Nearest Neighbor (KNN) [8] algorithms to analyse the necessary disease prediction. We use the ensemble technique [9] to get more accurate predictions from the machine learning techniques.

To forecast the neuro-clinical data, we employ machine learning techniques. We used two longitudinal cohorts in order to train, cross-validate, and test the models' applicability to the other cohort. The objective was to predict the likelihood of ICDs at the upcoming appointment using the patient's overall data and genetic data.

The breakdown of the paper residue is as follows: In section II, which is related, we discussed the assessment of the literature and the flaws of the works. Part III describes the data set and experimental setup. A mathematical analysis of machine learning is presented in Part III. Part IV of the article describes the proposed system used in this research. The statement of the problem is described in section V with methodology. Feature extraction and reporting on the experimental evaluation is in Section VI. An description of the comparative study and result analysis is given in Parts VII and VIII. The findings and forthcoming projects are covered in Section IX.

## 2 Literature Survey

The importance of studying neuroscience cannot be emphasized with in some limit. Due to the complexity of the nervous system, a problem's diagnosis might be challenging. Additionally, the medical community observes novel virus varieties for which a cure is required. We need to provide patients with faster care in today's society. Many researchers used Bayesian quadratic discriminant analysis to test the kinematic and diagnostic aspects of CA patients. To anticipate the disease, the author calculated finger chase movement values. The study clarified how machine learning methods using density-based mining and root mean value probability work.

ICDs and related disorders were found to be prevalent in cross-sectional studies in a range of 15-20% [10-12], with an estimated yearly incidence of over 10% [13-15] and a five-year cumulative incidence of over 50% [16]. These problems can also impact PD patients who have had their condition for more than five years.

Traditionally, ideas like the sway path are used to evaluate postural stability. This method is commonly used with force plate platforms to determine the variation of the centre of foot force [17]. There is interest in creating examinations of balance problems based on internal measurement units (IMUs), as these procedures are laborious and the equipment is pricey.

The research on neurodegenerative CA described in the references [18-20] is early and constrained by either features extraction methods or attempts to correlate innovative measuring systems with traditional medical scores. So, the need for a thorough framework for the qualitative evaluation of postural cerebellar impairments based on IMU sensors still warrants further study.

In this study, we built on earlier work [21] to present a thorough methodology for evaluating CA with wearable sensors. Our main focus in using contemporary machine learning techniques to determine the viability of diagnosing postural impairments in CA is the motion test.

The random forest algorithm's k mean value is used to calculate the human activity pattern. The author employed classifiers to identify the deviation after taking human activity into account as the pattern. To comprehend the pattern styles, they used various graphical data. Bitmaps were employed in utility pattern mining, and the output shown as visual graphs. The majority of instances of ALS are caused by a deadly neurological ailment that is sporadic, and this chronic disease's origin will evolve at different phases. It is quite challenging to pinpoint the mutation's genesis.

Machine learning aids in the early diagnosis of disease. Based on an action, the process mining [22] is evaluated. The technique for coordinating the order of activities is called activity mining. The movement is tracked and recorded to create a data collection. The author used Peripheral Measurements to analyze the disease. Algorithms for machine learning are utilized to forecast the results of human pattern.

The Naive Bayes algorithm [23] was employed by the author to forecast real developments. On the activity pattern data set, SVM actions are utilised to forecast the values. The measures were generated and assessed using this method. Every step of the process was calculated and evaluated using the process mining concept.

The topographical approach and values were used by the author to assess the road colocation pattern [24]. The author employed a neural network and colocation rules to employ this project. For humans, the capacity of the brain is more important. The body as a whole is activated using the brain as a model. The actions are well-synchronized and cooperate with the brain. The pattern displays negative numbers when the actions fail.

This research proposes a machine learning-based mechanism for CA prediction. Three algorithmic phases make up the presented model for the complete methodology: choosing the pre-processing features, discretizing the CA dataset, and deep learning-based

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(6), 928-939 | ISSN:2251-6727

classification. In a later section of this research, efficiency is examined and comparison analysis and evaluation of outcomes for CA datasets are reported.

This research proposes a machine learning-based mechanism for CA prediction. Three algorithmic phases make up the presented model for the complete methodology: choosing the pre-processing features, discretizing the CA dataset, and deep learning-based classification. In a later section of this research, efficiency is examined and comparison analysis and evaluation of outcomes for CA datasets are reported.

The approaches employed in the pertinent principles and drawbacks are explained in Table 1 below.

**Table 1.** Literature survey concepts

| Ref. | Approaches | Methods Used | Drawbacks |
|---|---|---|---|
| [1] | IOT devices | Microprocessor unit, matlab | Low accuracy using sensors. |
| [2-4] | Machine learning | SVM, KNN, Decision tree | It was only applicable to one subject and didn't apply to several ailments. |
| [5] | Machine learning | Classification algorithms | A single notion was used to identify the best classifier. |
| [6,7] | Machine learning, Ensemble method | AdaBoost | This study established the effectiveness of the Gaussian-RBF kernel. However, this is not used in any program. |
| [8] | Machine learning | SVM, and decision tree | Small database used. |
| [9-12] | Machine learning | Classification algorithms | They only employed one sample, and in real-time circumstances, classification did not flow dynamically. |
| [13,14,15] | Machine learning | Tree structured algorithm | Decision-making processes are lacking in some subjects that this research needs to include. |
| [18,23,25] | Machine learning, Ensemble method | SVM, encryption and fuzzy logic system | Lacking to make grouping and integrate the self-adaptive features. |
| [28,30,] | Activity Process mining | Alpha algorithm | Real time dataset is missing. |
| [32] | Machine learning | Naive Bayes | This study did not look into concerns with development and mobility. |

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(6), 928-939 | ISSN:2251-6727

## 3 Dataset

The National Center for Biotechnology Information (NCBI) [25] is where the dataset was gathered. The dataset consists of a number of different patient movement attributes. The dataset contains 19 occurrences that could result in activity movements. As the condition is uncommon, there are fewer cases. The table below shows the mobility values for neurological patients.

**Table 1.** Neurological patient's Daily activities

| S.No | Movements | Direction |
|---|---|---|
| i | Stance | Capturing the stand and straight position |
| ii | Swing | Arm or hand rotation |
| iii | Step length | Walking length |
| iv | Speed | Walking strength and speed |
| v | Hip | Finding the hip rotation |
| vi | Knee | Measuring the knee movements |
| vii | Ankle | Measure the ankle movement degree |
| viii | Cadence | Continous Movement |

Table 1 explains the different movements [26] that the patients make while they attempt to perform an activity. The order in which activities occur across time is determined by the activity pattern. The flow of activity that showed the disease's severity but was neglected throughout their treatment. The movement values utilized in this study's dataset are shown in Table 2.

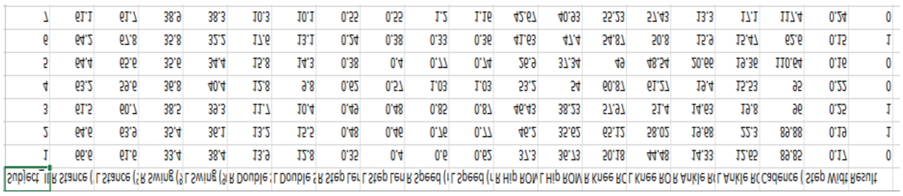

**Fig. 1** Cerebellar Ataxia Patient Data set

Fig.1 displays a sample of the data set's values. 19 attributes make up the data collection. Table 1 provides explanations of the qualities. These values are gathered from the testing apparatus. Next, as shown in Fig. 2, the data value are captured through the signals. The final characteristic indicates whether or not the patient has Cerebellar Ataxia. The outcome of this binary decision is displayed in the table.

The experimental set-up for gathering patient signal is shown in Fig. 2. The actions of the patients are tracked via wireless connections. It converts the electromagnetic signals into data.

Journal of Chemical Health Risks
www.jchr.org
JCHR (2023) 13(6), 928-939 | ISSN:2251-6727

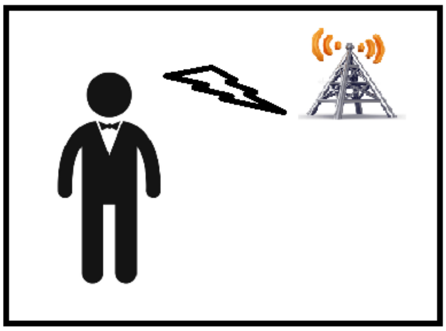

**Fig.2** Experimental setup

## 4 Proposed system

Data processing and the extraction of interesting variations that complied with inclusion criteria were done using the suggested system. The multiple text-like documents were analyzed using the Python libraries pandas [27] and NumPy. The method used to impute incomplete data was called "forward-filling," and it called for using the attribute and object in question's most recent non-missing values. Baseline missing values were imputed using the baseline values from the training dataset. We chose to utilise this easy approach because it can be applied at any size without any training and only a small percentage of the data from a small number of variables was allotted.

We distributed 30% of the patients to the test set for the outer loop and 70% of the patients to the training set at random. On the training dataset, we used a 5-fold subject-level cross-validation [28] approach in the inner loop to optimize the model hyper-parameters. These hyper-parameters regulate how closely the algorithms follow the training set of data.

Data from training and testing are split 70:30. For improving the parameters, we employed a 5-fold cross [29] validation approach. A model is developed using the inner loop to control the algorithms. The hyper parameters are regarded as the models at the training level.

### 4.1 Architecture

Fig 3. Explains the complete strucutre of proposed system. The process flow and the list of algorithms used for this research. The machine learning approaches are giveing the best result for predictions.

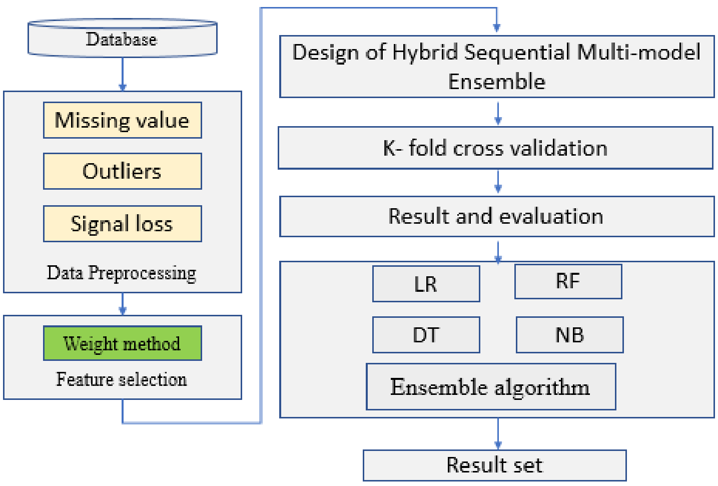


**Fig.3** Experimental setup

## 5 Methodology

### 5.1 Logistic regression

The first step of the workflow entails a number of health metrics for which datasets are available for the 14 attributes. Then, as pre-processing is necessary since it aids in data cleaning [30], we must process this dataset using machine learning methods. The datasets are then divided into two halves. They are testing and training data, respectively. We employ an 80-20 split, in which 80% of the data is used for training and 20% for testing. Here, we use the training data and labels found in the files x train and y train. It also includes x test and y test, which are used to test the data and

Journal of Chemical Health Risks
www.jchr.org
JCHR (2023) 13(6), 928-939 | ISSN:2251-6727

labels using the Sklearn library.

The model predicts the output as either 1 or 0, based on logistic regression, which employs binary classification [31]. The anticipated output of zero indicates that the subject is disease-free. In contrast, the output of 1 indicates that the person has cardiac disease. A target table is used here to illustrate it. Every dataset has some values; null values are never represented in a dataset. When sending the data for prediction, the pre-processing stage is crucial. The distribution of many parameters, such as age and sex, cholesterol and fasting blood, resting ECG and thalamus, exang and old peak, slope and ca, etc., is thoroughly examined. As a result of this study, the likelihood of developing heart disease can be accurately predicted.

The straight line equation:

$$y = b_0 + b_1x_1 + b_2x_2 + \ \dots . + b_nx_n \quad (1)$$

The value y can be in between 0 and 1, finally the equation 1 shows like below:

$\frac{y}{1-y}$ when y value is 0 then the output also 0; when y=1 then the output is ∞

But we need range between -[infinity] to +[infinity], then take logarithm of the equation 8 it will become:

$$log\left[\frac{y}{1-y}\right] = b_0 + b_1x_1 + b_2x_2 + \cdots + b_nx_n \quad (2)$$

### 5.2 Random forest

The random forest regression is also ensemble of trees only. Predicted value by the random forest at randomized forest [32]. This random variable can be used to describe the randomness induced by node splitting and the sampling procedure for tree construction. The trees are combined to form the finite forest estimate.

### 5.3 Naïve Bayes

The Nave Bayes algorithm is a one of the supervised learning method. It will help for classification issues. NB implies to categorize the text and finding the probability of the occurrence of the classes. NB classifiers are used to categorize the values dependently with the other attributes. So this algorithm is giving the accurate prediction in machine learning algorithms. Using the probability, it finds the likelihood of the object in the attribute set. This algorithm mostly used to text analysis.

$$P\left(\frac{A}{B}\right) = \frac{P\left(\frac{B}{A}\right)P(A)}{P(B)} \quad (3)$$

The probability value is calculated using formula 3.

### 5.4 Decision tree

The widely used classification that makes use of a number of logical operations is the decision tree (if-then rules). It facilitates decision-making [33]. Equations 4 and 5 explain the information entropy and gain values used in its definition, which is dependent on the nearest neighbor value that is computed.

$$Entropy(P) = -\sum_j (P_i(log_2(\pi)) \quad (4)$$

$$Gain(S,F) = Entropy(S) - \sum_f \in values(F) \frac{||s_f||}{||s||} Entropy(s_f) \quad (5)$$

## 6 Feature Extraction from Ensemble Test Data

Typical people can keep their equilibrium even when their eyes or blink; nevertheless, slight deviations may happen because of background noise and an object's respiration. A waveform that is mainly stable in the blinking stage and unsteady in the closing eyelids is produced by patients with sensory ataxia who can maintain posture and balance during the blinking stage and vigorously move their bodies during the closing period.

The following is the attribute's physical relevance: The signal's mean square value reflects the energy of the signal, its mean value describes its stability, the standard deviation reflects the degree of dispersion, the signal's kurtosis reflects the impact characteristics, and the signal's skewness reflects the asymmetry of the signal.

### 6.1 XGBoost Algorithm

The XGBoost algorithm can be used for foot gait value analysis by leveraging its powerful capabilities in handling structured data and predicting target variables. Here's how you can connect foot gait value analysis with the XGBoost algorithm which is explained in equation (6-10):

Step 1: Dataset Preparation: Collect a dataset of foot gait values, including various features such as step length, step duration, foot angle, pressure distribution, etc. Ensure that you have a sufficient amount of labeled data, where the target variable represents a specific aspect of foot gait analysis (e.g., identifying abnormal gait patterns).

Step 2: Data Preprocessing: Preprocess the foot gait dataset by handling missing values, outliers, and encoding categorical variables if necessary. Perform feature scaling

Journal of Chemical Health Risks
www.jchr.org
JCHR (2023) 13(6), 928-939 | ISSN:2251-6727

or normalization to ensure that all features have a similar scale, which helps XGBoost perform better.

Step 3: Dataset Split: Split the preprocessed dataset into a training set and a test set. The training set will be used to train the XGBoost model, while the test set will evaluate its performance on unseen data. Typically, an 80:20 or 70:30 split is used, but you can adjust this ratio based on the dataset size and specific requirements.

XGBoost Model Training: Train an XGBoost model on the prepared training set. XGBoost is a gradient boosting algorithm that constructs an ensemble of decision trees to make predictions. Specify the XGBoost hyperparameters such as the learning rate, maximum depth, number of trees (boosting rounds), regularization parameters, etc.

Step 4: Model Evaluation: Evaluate the trained XGBoost model using the test set. Calculate relevant evaluation metrics such as accuracy, precision, recall, F1 score, or any other suitable metric based on the specific foot gait analysis task. This step helps assess the model's performance and identify any potential issues, such as overfitting or underfitting.

Step 5: Hyperparameter Tuning: Perform hyperparameter tuning to optimize the XGBoost model's performance. This can be achieved through techniques like grid search, random search, or more advanced methods like Bayesian optimization. Iterate over different combinations of hyperparameters and evaluate the model's performance until satisfactory results are obtained.

Step 6: Feature Importance Analysis: Analyze the importance of features in the XGBoost model to gain insights into foot gait analysis. XGBoost provides a feature importance score based on how much each feature contributes to the model's predictions. This analysis helps identify the most influential gait parameters and understand their relationship with the target variable.

Step 7: Predictions: Once the XGBoost model is trained and fine-tuned, you can use it to make predictions on new, unseen foot gait data. Preprocess the new data using the same preprocessing steps applied to the training set. Feed the preprocessed data into the trained XGBoost model to obtain predictions, which can be used for various applications like identifying abnormal gait patterns or providing personalized recommendations.

By connecting foot gait value analysis with the XGBoost algorithm, you can leverage the algorithm's strengths in handling structured data and building accurate predictive models.

*Algorithm 1: XGBoost Pseudo code*

Input : Gait values (Attributes)
Output: Disease prediction
Step 1: Calculation of weight value using attribute
Step 2: Error function used to calculate error that stated in equation 6

$$\in_t = \sum_i N(w_n)(t) I(y_n \neq (h_t(x_n)) \quad (6)$$

Step 3: Define the weight value using below equation 7

$$\alpha_t = \log(\frac{1-\alpha_t)}{\alpha_t} \quad (7)$$

Step 4: Weight updating using below equation 8

$$(w_n)^{(t+1)} = W_n^{(t)} \exp(\alpha_t \frac{I(y_n \neq (h_t(x_n)))}{z_t} \quad (8)$$

Weight normalization using the above equation 9
Step 5: Output is calculated by below equation 10

$$f(x) = sign \sum_{t=1}^{T} \alpha(t) h_t(x) \quad (10)$$

# 7 Experimental Result

## 7.1 Conventional AoG Detection Features Extraction and Selection

Pre-stops' gait patterns were dissimilar from those of pre-AoGs, as was the case in the past (Fig. 4). The VT variation pattern for pre-stops differed slightly from that of pre-AoG, and the cadence tended to be steadier. Similar trends were seen for other metrics in pre-AoG times as well, but they tended to be milder and to get better or worse one or two steps later. The FP rate could be lowered and pre-AoG and pre-stops could be distinguished using these various trends.

## 7.2 Results for AoG values with coordinates

We decide which sensor data to display with both of our feet. The graph's x-axis in Figure 4 displays two distinct hues. The variants are depicted using the colour yellow. Patients with CA have radically different foot positions than people with other conditions.

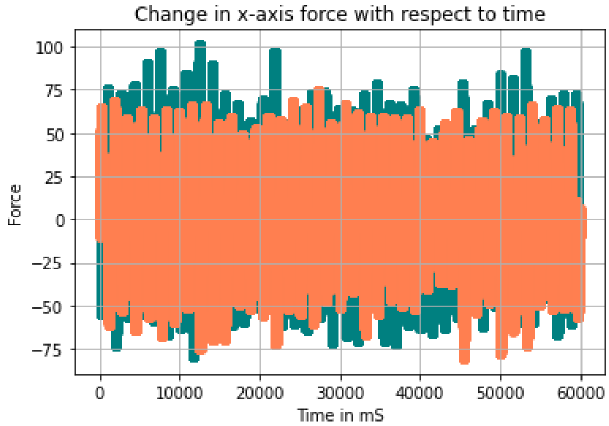


**Fig. 4** The Gait value analysis with x-axis

### 7.3 y axis of LRFM1 sensor data to see different variations of sensor data

Fig. 5 shows the foot y axis positions; the left leg is shown in blue, and the right leg is shown in red. The varied numbers displayed in fig. 5 are the result of the left and right legs moving in unison.


**Fig. 5** The Gait value analysis with y-axis

### 7.4 z axis change of LRFM2 sensor

The z axis readings for the left and right legs were obtained using the LRFM sensor and are displayed in Fig. 6. The right leg is shown in red, whereas the left leg is shown in blue. Figure 6 displays the values that were noted at the relevant period.

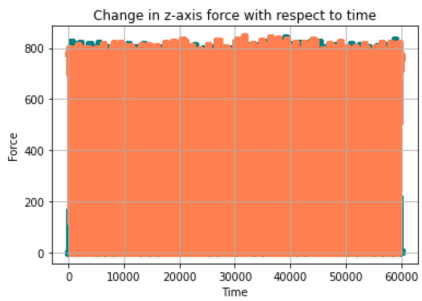


**Fig. 6** The Gait value analysis with z-axis

### 7.5 Visualize all axis in 3D graph

Journal of Chemical Health Risks
www.jchr.org
JCHR (2023) 13(6), 928-939 | ISSN:2251-6727

In Fig. 7, the right leg is depicted with blue colour, while the left foot is depicted with red colour. This 3D visualisation was created using the Matplotlib tool. The leg posture of CA patients varies and requires the longest amount of time to be maintained in the usual position. The right and left feet of a human are seen in the above illustration moving. (with 0,6 m/s, GP1).

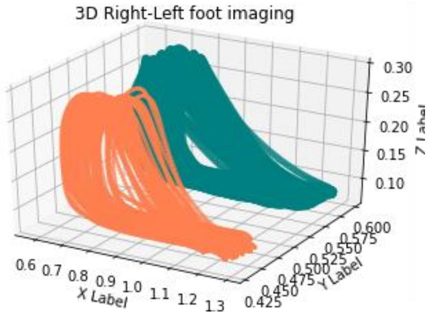


**Fig. 7** Correlation matrix for Gait analysis

The relationship between left and right position for average people is seen in Fig. 7. Between the two legs, there should be adequate alignment of the coordination. We must recognise and resolve the issue behind any misalignment of the legs' axes. In Fig. 6, the x, y, and z axes are used to show the alignment of both legs. The graph clearly shows the alignment.

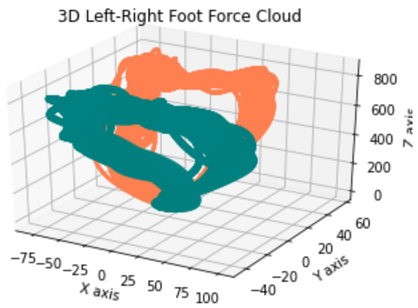


**Fig. 8** Correlation matrix for Gait analysis

Figure 8 depicts the change in leg position. It will provide details regarding the individual and their condition. Only that patient will experience this type of limb misalignment when the neurological condition strikes at that moment.

## 8 Comparative Analysis

The CA properties are explained in Fig 9 along with the result value. It explains how the outcome and accompanying traits relate to one another. So, we can comprehend how the severity of the disease relates to the level of the attribute. The ankle attribute gives the major problem in this case. So that the attribute reflects more in the above fig 9.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(6), 928-939 | ISSN:2251-6727

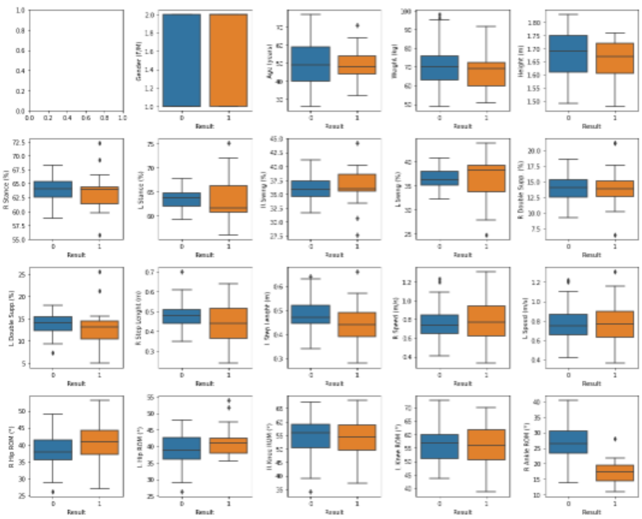

**Fig.9** Attribute relation with CA disease

The fig 9 shows the prediction of CA. The attribute values are plotted in the dotted graph. This interprets with the CA disease status. It explains the possibility of the disease occurrence with the attribute.

### 8.1 Result Analysis

The final result analysis is given in the below figure. The accuracy, F1-score, recall, precision and R2-Score values are calculated based on the confusion matrix. The various machine learning algorithms are applied in the CA dataset and derived the result as in Fig. 6. The result shows that the boosting algorithm gives the better result compared to the remaining all machine learning algorithms.

### 8.2 Comparation table

The below table 2 gives the precise format of the evaluation analysis among the various machine learning algorithms. The existing machine learning algorithm's results are compared with XGBoost in the table 2.

**Table 2.** Comparative study for prediction of CA disease

| S.No | Algorithm | Prediction Accuracy (%) |
|---|---|---|
| 1 | DT | 97.5 |
| 2 | RF | 98.1 |
| 3 | Naïve Bayes | 93.8 |
| 4 | Logistic Regression | 98.6 |
| 5 | XGBoost | 99.6 |

The accuracy is measured based on the confusion matrix. So that it will find the accuracy which is true positive values. The comparative table explained the existing algorithms (RF, NB, LR and DT) result with XGBoost. Based on the metrics, XGBoost is giving a better effect.

## 9 Conclusion

One of the important problems in the medical field is the prediction of neurological diseases. There are several ways that neurological disorders might make it difficult to carry out regular tasks. It may occasionally result in life mortality. We explored with the neurological data set in this paper. The severity of the condition is reflected in the patient's activity problem. When a human's activity differs from the norm, we can assume they have CA illness. The heel-knee-movement analysis is used to collect the patient's activity disorder data in the data set.

The essential features are first discovered, the data is then preprocessed by removing outliers and using wavelet transform filtering, and finally the model is trained using Neural Network, DT, LR and RF machine learning techniques. The experimental results show that most of the algorithms can achieve about 90-98% predictive performance, accurately differentiating between sensory ataxia and cerebellar ataxia, proving the efficacy of the technical strategy described in this paper. XGBoost, however, has a 99.6% success rate in predicting neurological diseases. When compared to other machine learning algorithms, it is producing the best results.

## References

1. T. Ngo et al., "Balance deficits due to cerebellar ataxia: A machine learning and cloud-based approach," IEEE

Trans. Biomed. Eng., vol. 68,no. 5, pp. 1507–1517, May 2021.

2. A. Esteva et al., “A guide to deep learning in healthcare,” Nature Med., vol. 25, no. 1, pp. 24–29, Jan. 2019.
3. R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, “Deep learning for healthcare: Review, opportunities and challenges,” Briefings Bioinf., vol. 19, no. 6, pp. 1236–1246, Nov. 2018.
4. M. Adnan, S. Kalra, J. C. Cresswell, G. W. Taylor, and H. R. Tizhoosh, “Federated learning and differential privacy for medical image analysis,” Sci. Rep., vol. 12, no. 1, pp. 1–10, Dec. 2022.
5. M. Hao, H. Li, X. Luo, G. Xu, H. Yang, and S. Liu, “Efficient and privacy-enhanced federated learning for industrial artificial intelligence,” IEEE Trans. Ind. Informat., vol. 16, no. 10, pp. 6532–6542, Oct. 2020.
6. Y. Liu, J. J. Q. Yu, J. Kang, D. Niyato, and S. Zhang, “Privacy-preserving traffic flow prediction: A federated learning approach,” IEEE Internet Things J., vol. 7, no. 8, pp. 7751–7763, Aug. 2020.
7. Q. Wu, K. He, and X. Chen, “Personalized federated learning for intelligent IoT applications: A cloud-edge based framework,” IEEE Open J. Comput. Soc., vol. 1, pp. 35–44, 2020.
8. Y. Zhao, H. Liu, H. Li, P. Barnaghi, and H. Haddadi, “Semi-supervised federated learning for activity recognition,” 2020, arXiv:2011.00851.
9. Shanmuga Sundari, M., Sudha Rani, M., & Ram, K. B. (2023). Acute Leukemia Classification and Prediction in Blood Cells Using Convolution Neural Network. In International Conference on Innovative Computing and Communications (pp. 129-137). Springer, Singapore.
10. D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, “Federated learning for Internet of Things: A comprehensive survey,” IEEE Commun. Surveys Tuts., vol. 23, no. 3, pp. 1622–1658, 3rd Quart., 2021.
11. Q. Wu, X. Chen, Z. Zhou, and J. Zhang, “FedHome: Cloud-edge based personalized federated learning for in-home health monitoring,” IEEE Trans. Mobile Comput., early access, Dec. 16, 2020, doi:10.1109/TMC.2020.3045266.
12. N. Rieke et al., “The future of digital health with federated learning,” npj Digit. Med., vol. 3, no. 1, pp. 1–7, Dec. 2020.
13. R. Kumar et al., “Blockchain-federated-learning and deep learning models for COVID-19 detection using CT imaging,” IEEE Sensors J., vol. 21, no. 14, pp. 16301–16314, Jul. 2021.
14. Shanmuga Sundari, M., & Jadala, V. C. (2022). Improved Performance Analysis for Cerebellar Ataxia disease Classification using AdaBoost. NeuroQuantology, 9488-9497.
15. J. Lee, Y. Kagamihara, and S. Kakei, “A new method for functional evaluation of motor commands in patients with cerebellar ataxia,” PLoS ONE, vol. 10, no. 7, Jul. 2015, Art. no. e0132983.
16. C. Hohenfeld et al., “Application of quantitative motor assessments in friedreich ataxia and evaluation of their relation to clinical measures,” Cerebellum, vol. 18, no. 5, pp. 896–909, Oct. 2019.
17. Chen Y, Ghannam R, Heidari H. Air Quality Monitoring using Portable Multi-Sensory Module for Neurological Disease Prevention. In2019 UK/China Emerging Technologies (UCET) 2019 Aug 21 (pp. 1-4). IEEE.
18. Rukhsar, S. (2018). Discrimination of multi-class EEG signal in phase space of variability for epileptic seizure detection using error-correcting output code (ECOC). International Journal of Information Technology, 1-13.
19. Nasser, I. M., Al-Shawwa, M., \& Abu-Naser, S. S. (2019). Artificial neural network for diagnosing autism spectrum disorder. International Journal of Academic Information Systems Research (IJAISR), 3(2).
20. Shanmuga Sundari, M., Sudha Rani, M., & Kranthi, A. (2023). Detect Traffic Lane Image Using Geospatial LiDAR Data Point Clouds with Machine Learning Analysis. In Intelligent System Design (pp. 217-225). Springer, Singapore.
21. X. Yang and S. Member, "Activity Pattern Mining for Healthcare," IEEE Access, vol. 8, pp. 56730–56738, 2020, doi: 10.1109/ACCESS.2020.2981670.
22. J. V. Chandra, G. Ranjith, A. Shanthisri, R. Rakshitha, and K. Mahesh, "A Framework for Implementing Machine Learning algorithms using Data sets," no. October, 2019, doi: 10.35940/ijitee.K1263.0981119.
23. Ker, J. I., Wang, Y., Hajli, M. N., Song, J., \& Ker, C. W. (2014). Deploying lean in healthcare: Evaluating information technology effectiveness in US hospital pharmacies. International Journal of Information Management, 34(4), 556-560.
24. H. T. Lin and L. Li, "Support vector machinery for infinite ensemble learning," J. Mach. Learn. Res., vol.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(6), 928-939 | ISSN:2251-6727

9, pp. 285–312, 2008.

25. Shanmuga Sundari, M., Samyuktha, P., Kranthi, A., & Das, S. (2022). Evaluating Performance on Covid-19 Tweet Sentiment Analysis Outbreak Using Support Vector Machine. In Smart Intelligent Computing and Applications, Volume 1 (pp. 151-159). Springer, Singapore.
26. Shanmuga Sundari, M., & Nayak, R. K. (2023). Deviation and Cluster Analysis Using Inductive Alpha Miner in Process Mining. In Communication, Software and Networks (pp. 451-458). Springer, Singapore
27. Eldardiry, H., Neville, J., \& Rossi, R. A. (2020). Ensemble Learning for Relational Data. J. Mach. Learn. Res., 21, 49-1.
28. L. I. Junhuai, T. Ling, W. Huaijun, and A. N. Yang, "Segmentation and Recognition of Basic and Transitional Activities for Continuous Physical Human Activity," 2019, doi: 10.1109/ACCESS.2019.2905575.
29. Gopi, A. P., Jyothi, R., Narayana, V. L., Sandeep, K. S. (2020). Classification of tweets data based on polarity using improved RBF kernel of SVM. International Journal of Information Technology, 1-16.
30. S. Al-janabi and A. H. Salman, "Sensitive Integration of Multilevel Optimization Model in Human Activity Recognition for Smartphone and Smartwatch Applications," vol. 4, no. 2, pp. 124–138, 2021, doi: 10.26599/BDMA.2020.9020022.
31. M. S. Sundari and R. K. Nayak, "Efficient Tracing and Detection of Activity Deviation in Event Log Using ProM in Health Care Industry," in 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics, and Cloud)(I-SMAC), 2021, pp. 1238–1245.
32. M. Zhou, T. Ai, G. Zhou, and W. Hu, "A Visualization Method for Mining Colocation Patterns Constrained by a Road Network," IEEE Access, vol. 8, pp. 51933–51944, 2020, doi: 10.1109/ACCESS.2020.2980168.
33. Raju, C. S. K., Pranitha, K., Samyuktha, P., & Madhumathi, J. (2022, March). Prediction of COVID 19-Chest Image Classification and Detection using RELM Classifier in Machine Learning. In 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 1184-1188). IEEE.