

***University of Ibadan Journal of
Science and Logics in ICT
Research (UIJSLICTR)***
ISSN: 2714-3627

A Journal of the Faculty of Computing, University of Ibadan, Ibadan, Nigeria

Volume 17 No. 1, June 2026

journals.ui.edu.ng/uijslictr
http://uijslictr.org.ng/
uijslictr@gmail.com



Prepulse Inhibition Response Classification Model using Machine Learning Techniques

Olaniyan Abolade Shekinah¹, Adeniji Oluwashola David², Adeyemo Adebolajo A.³ and Oguntunde Toyin⁴

¹Department of Computer Science, Adeseun Ogundoyin Polytechnic, Oyo State, Nigeria

^{2,4}Department of ICT and Cybersecurity, University of Ibadan, Nigeria

³Institute of Child Health, College of Medicine, University of Ibadan, Ibadan, Nigeria

¹308mobeni@gmail.com, ²od.adeniji@mail.ui.edu.ng, ³aadeyemo@com.ui.edu.ng

⁴t.oguntunde@ui.edu.ng

Abstract

The use of acoustic has gained attention because of its application in signal processing to analyse sound in the field of speech recognition and computational acoustic. While Machine learning (ML) techniques for classification is rapidly developing with compelling results and significant future promise, little information is known on its application in classifying startle reflex waveforms. Hence, this study was aimed at developing an improved prepulse inhibition classification model using machine learning technique. A dataset of eight hundred (800) prepulse inhibition responses were collected from the prepulse inhibition experiment performed by introducing sound at pulse alone trial 45dB, prepulse + pulse trial 85dB and randomised trials using ASLAS device. For training, validation, and testing, the data was divided into ratios of 70:15:15. Decision trees (DT) (Entropy) and (Gini), logistic regression (LR), and random forests (RF) models were selected to categorize waveforms into startle and non-startle at 85dB stimulus intensity. Performance measures were accuracy, precision, recall, and F1-score. Improved acoustic device (ASLAS) utilised RF, LR, DT (Gini) and DT (Entropy) on its dataset to yield accuracy of 96.0, 97.0, 93.0 and 94.0%; precision of 74.0, 78.0, 92.0 and 75.0%; recall of 80.0, 82.0, 75.0 and 84.0%; and f1-score, 72.0, 82.0, 83.0 and 93.9%, respectively. LR gives best accuracy (97%), DT (Gini) has highest precision (92%), while DT (Entropy) has highest Recall (84%) and F1-Score (93%). This result implies that the model predictions are correct and a high F1 score means the model performs well on both metrics precision and recall. The importance of this is that it enhanced decision-making through the elimination of blind spots by forecasting future outcomes and also helps in proactive risk management by identifying potential vulnerabilities and anomalies which minimizes losses.

Keywords: Acoustic Startle Response, Prepulse Inhibition, Machine Learning, Wave form Classification

1. Introduction

Loss of hearing is one of the most prominent diseases in children and adults, and can negatively impact quality of life and social functioning if it is not diagnosed early properly (Lieu et al., 2020). The Acoustic Startle Reflex (ASR) serves as a warning system for survival and sensitive tool that can help elucidate the contribution of neuroactive steroids (Hantsoo et al., 2018). Acoustic startle response (ASR) modification procedures in mice are not as

sensitive as traditional operant techniques but may be used for assessing auditory behavior in certain situations (Lauer et al., 2017).

Evaluating The startle reaction in subjects is thought-provoking considering the exceptionally short duration which happens amid 5 and 12ms between several muscle units (e.g., neck, hind limb). The rodent startle reaction is usually evaluated in cramped, restricted spaces that limit animal mobility because of this difficulty (Madeline et al., 2020). PrePulse Inhibition (PPI) is the most used type of ASR modification. Prepulse inhibition of the acoustic startle reflex is a quantitative measure of central processing and may be used as a marker for short and long term plasticity in brainstem sensorimotor gating mechanisms (Gomez et al., 2020).

Olaniyan Abolade Shekinah, Adeniji Oluwashola David, Adeyemo Adebolajo A. and Oguntunde Toyin (2026) Prepulse Inhibition Response Classification Model using Machine Learning Techniques, *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR)*, Vol. 17 No. 1, pp. 230 – 237.

Prepulse inhibition research has advanced significantly since 1992 (Swerdlow et al., 2016) to accurately measure complex neurological systems. Prepulse inhibition is a quick and accurate method for assessing hearing in animals and thus offers a high throughput automated screening tool (Longenecker et al., 2016).

Assessing PPI of acoustic startle involves series of standardized protocol designed to measure sensory gating. This method is not only applicable to rodents but can also be applied to humans and other animals. Machine learning techniques offers skilled approaches to spontaneously classify responses as startle or non-startle. This is accomplished by using features to train classification models generated from experimental. The models are capable of identifying predicted trends and features of non-startle responses. The developed model classifies prepulse inhibition responses into startle and non- startle.

2. Literature Review

2.1 Related works

Virag *et al.* (2020) developed Acoustic Startle Platform (PASTA). The PASTA is based on digital kitchen scale which enables data acquisition store and examine at a fast sample rate. the data which is not possible with the ordinary kitchen scale. The data collected at high sampling rate was used to extract and analyse the respiratory patterns in the study technique during a freeze behavior.

Fawcett and Longenecker *et al.*, (2022) use a machine learning technique to automatically categorize startle reflex animal waveforms models leading to data reliability. They provided a comprehensive protocol written in the computer language R to help scientists in categorizing SR waveforms into non-startles and startles. They loaded labeled acoustic startle responses in order to train the machine learning model in the computational environment with a limited number of features to classify the responses into startle and non-startle. Results show that data consistency will be enhanced by the machine learning-based method and interoperability between labs using the startle reflex paradigm and interoperability between labs using the startle reflex paradigm

Emmanuel *et al* (2021) uses a combination of

Artificial Neural Network and Random Forest hybrid (RF-ANN) models to predict the occurrence of mental disorder in IT workers. The algorithms for Random Forest and Artificial Neural Networks were created independently, and their outcomes were noted and contrasted with those of the hybrid model. The hybrid model using “Bagging Ensemble, showed a substantial progress in its performance when compared to the individual results of the Artificial Neural Networks and Random Forest models. The weighted average performance for precision, recall, and accuracy in the Hybrid Random Forest and Artificial Neural Networks Model is 84.5%, while precision is 82.5%. The hybrid model's outcome accurately demonstrates a significant improvement in performance over the RF and ANN models' standalone performances. This implies that an enhanced dataset could be examined and compared with the outcomes by using the RF-ANN model.

Stevanovic *et al* (2025) presented a new automated platform for assessing acoustic startle response adult zebrafish. To evaluate the sonic startle in adult zebrafish, an automated kinematic analysis platform was created. It was discovered that females exhibit less habituation and more reactivity than men. Their findings increase interest in utilizing adult zebrafish to evaluate the possible impact of environmental contaminants on ASR plasticity.

Oguntunde *et al.*, (2025) employ machine learning methods to create a model for transaction alert system to distinguish between fake transaction alert and genuine one. To ascertain the genuity of the transaction alert. MATLAB programming is used to simulate a fake transaction alert model utilizing the Random Forest (RF) and Support Vector Machine (SVM) algorithms. At performance evaluation, RF algorithm method was more suited for the creation of the phony transaction detection of alerts due to its great exactness, precision, memory and F-measures value of 97.6, 97.48, 97.58, and 97.8%.

Thiago *et al* (2024) created a computer model to simulate the modulation and inhibition of prepulse by cortical and subcortical units. They created a computer model of a brain circuit used for modulation and ASR and PPI mediation. The model reproduces the consequences of modifications in inter stimulus interval, prepulse

strength, and ASR habituation as observed in experiments. They adjusted models parameter to replicate a 60dB virtual background. As the Inter Stimulus Intensity (ISI) rises, the %PPI rises as well and peaks between 70 and 100ms. The computational model that is being provided provides a useful tool to look at the disorder-related impairment in PPI.

Oyeniran *et al.*, (2020) developed an anomaly detection system for Zero-day attack in Network Traffic. They propose an ensemble machine language comprising XGBoost, LightGBM, Random Forest and CatBoost to detect attack effectively using a voting ensemble classifier. For training and testing, the dataset was divided into 75% and 25% portions respectively using stratified sampling technique. Individual models were trained and optimized with predictions. The results showed that ensemble voting classifier outperformed individual models achieving 99.95% accuracy followed by Random forest (99.43%), XGBoost (99.41%) and CatBoost (99.41%). The ensemble approach overcome the limitations of the traditional methods in handling encrypted traffic and multiclass classification.

Moein *et al.*, (2010) developed a decision-support system to assess hearing loss symptoms. The Support Vector Machine (SVM) and Multi-Layer Perceptron Neural Network (MLP) were utilized to classify hearing loss indicators in 150 patients from an otolaryngology clinic.

3. Methodology

3.1 Construction of Test-bed

This involves configuration for the measuring of hearing performance in rodent. It consists of:

- (1) a locally built kitchen scale with simple, low-capacity single load cell, which is a more cost-effective and simpler in design connected to a computers system via controller circuit board to read load cell resistance and a communication between controller board and the computer and
- (2) weighted mouse cage used as hearing test platform. The acoustic device for hearing test performance is built using a cost effective, robust and readily available material. The components of the acoustic device is given below:
 - a. A locally built Kitchen scale used for registering the test animal's activity. The core of kitchen scale is the load cell

connected to a readily available circuit board.

- b. A spotless mouse cage mounted atop the kitchen scale to restrict movement
- c. A readily available controller circuit board used to read load cells resistance.
- d. A Uno Arduino communication board acts as a bridge between the load cells controller board and the computer
- e. USB cable for communication to the computer system
- f. Computer acts as a control station which stores the scale data, automatically plays the startle sound at the proper timestamps after logging the received data.through the loudspeaker.
- g. A normal Audio cable for connection between the speaker and the computer system.
- h. A standard loudspeaker to present the shocking stimulus.

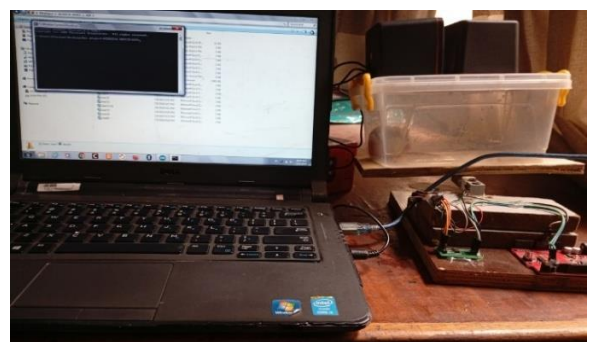


Figure 1: A schematic display of Testbed for the Adaptive Single Load Acoustic Startle (ASLAS)

3.2 Data Collection

This involves the collection of startle and non-startle signals for feature extraction and classification. Both the startle and non-startle responses will be collected together during the testbed implementation. To extract data from the platform, a custom Python script written for general purpose experimentation was used for hear testing. When the recording is completed, the script automatically saves data in the designated directory in a comma-separated value (csv) format, using the ".csv" extension to make file loading in the downstream analysis easier.. To make data capturing and visualization easier, the function `pd.read_csv()` was executed to load the csv data and `pygame.mixer.init()` to load sound. Moreover, the functions `plt.plot()` and `plt.show()` were utilized to plot many graphs, although the function `writer.writerow()` yields an open csv

file for writing of the data. The 85 dB noise burst caused noticeable startle, but the 45 and 55 dB noise bursts caused little to no startle, according to the i/o function.

Consequently, 45 or 55 dB white noise bursts (96-97ms duration) and 85 dB white noise bursts (96-98ms duration) were employed as pulse stimuli were used as prepulse stimuli. PPI testing started with a two-minute baseline, which was followed by a phase of habituation and PPI. 45 pulses spaced 96 milliseconds apart are presented during the habituation period. 800 trials total—120 pulse-only trials and 680 prepulse/pulse trials spaced 96 ms apart—make up the PPI phase. The prepulse was given at an interstimulus interval (ISI) of either 96 or 97 ms before the pulse in the prepulse/pulse trials. There was pseudo-randomization in the 800 trials.

3.3 Data Preprocessing

The first step in data preparation is a careful analysis of the dataset to identify and fix any errors, inconsistencies, or missing values.

3.3.1 Data Cleaning

Outliers and unnecessary information were eliminated from the prepulse inhibition response data. To improve data quality and dependability, data entry errors, formatting inconsistencies, and incorrect values were addressed. MATLAB functions including summary, missing, and outlier were used to find missing values, outliers, duplicates, and inconsistencies. Additionally, missing values were added by imputation or removal using functions like fillmissing and ismissing.

3.3.2 Feature Extraction

Relevant features such as startle response amplitude, latency, and habituation patterns, are extracted from the data. The features that indicate startle and non-responses are also extracted. Some features have poor predictive ability because of variations in sensor responses, startle response latency, and magnitude. Both filtering and normalizing are necessary to exclude high frequency information unrelated to the startle reflex response in order to extract highly predictive features from SR waveforms. Additionally, it guarantees easy comparison of waveforms across a variety of amplitudes. To ensure that only highly predictive characteristics free of redundant information are used to train the machine learning models, the retrieved

features are first pre-processed. This is accomplished by using data cleansing to exclude attributes with almost zero volatility. In order to prevent the amplitude and variability of any one feature from controlling the machine learning model, the remaining characteristics are then scaled and centered before the model is trained. The Random Forests, Decision Trees, and Linear Regression techniques are then used to train individual machine learning models.

3.4 Model Classification

At this phase, the features being extracted were preprocessed for training and classification. Classifiers such as Decision Tree, Random Forests and Logistics were used to categorize based on their PPI patterns. Figure 2 shows the three phases of the model.

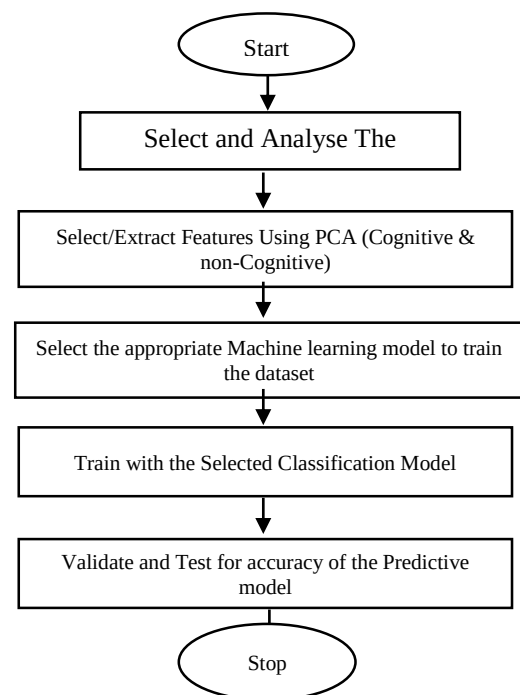


Figure 2: Phases of Model Implementation

3.5 Model Selection

It is recommended to utilize Python classifiers with the scikit-learn module. Scikit learn is a Python module that simplifies machine learning for non-specialist, focusing on performance, ease of use, documentation and API consistency. (Pedregosa et al., 2011). Random Forest is a popular machine learning method for classification and other tasks because of its computational efficiency, insensitivity to tuning parameters, and built-in cross validation and interpretation tools (Erwan Scornet et al., 2026).

Decision trees are popular in machine learning due to their simplicity and interpretability with applications in medical diagnosis and fraud detection (Ibomoiye et al., 2024) and also effective for data classification in various fields, (Bahzad et al., 2021)

3.5.1 Logistic Regression

Classification problems are carried out using Logistic Regression (LR) techniques. A probability classifier based on linear regression models is called logistic regression. One or more independent variables are used in linear regression to forecast an independent variable's value. A variant of linear regression that restricts the output value to any value between 0 and 1 is called logistic regression. LR is a powerful method for categorizing prepulse responses, and binary classification using logistic regression can be achieved using classification trees. (Predictive analysis 2020). Common applications of LR model is fraud detection and biomedical predictions. A model was proposed in this study using Scikit-learn's Logistic Regression package.

3.5.2 Decision Tree

By learning basic decision rules deduced from the data attributes, Decision Trees (DT) forecast the value of a target variable. Decision trees such as CART, can effectively model classification and regression problems with ensemble methods like random forest thus improving performing and accuracy (Njoku 2019). Decision trees is well known in data science for classification and regression task but with less stable predictions when compared with Random Forest due to multiple trees trained together (Talekar 2020)

3.5.3 Random Forest

One type of ensemble machine learning technique is called random Forest. It does not rely on a single decision tree to output results, but instead collects outputs derived separately from multiple independent decision trees. It is typically used for classification and regression, improving model accuracy and over fitting control. The resulting “forest” improves prediction accuracy in classification and regression tasks, but faces challenges like high computational costs and model interpretability (Ignatious et al., 2024). Similar to decision trees, random forests is capable of handling problems including both regression and classification. Algorithms for random forests

produce several decision trees for every assignment, combining the forecast of every tree, then select the most popular response as the outcome.

4.0 Result and Discussion

In the pulse alone trial experiment, startle reflex responses PPI test was conducted against the weight as shown in figure 3. It was observed that as the stimulus is constant, the weight changes. At initialization of sound, at time of less than 2s, the weight rises to 0.04g, the weight changes a bit to -0.03. There was a drastic increase in the weight at a given time of 4000ms and there was continuous stability in the weight as stimulus remains constant at a given interval of 4 seconds to 6 seconds.

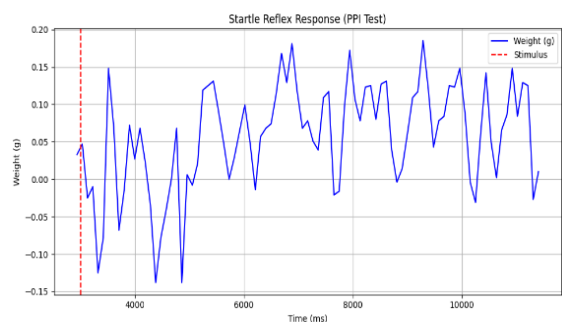


Figure 3: Results of Pulse Alone Trial simulation graph

In the case of Prepulse-pulse experiment, startle reflex One method for ensemble machine learning is random forest. It does not rely on a single decision tree to output results, but instead collects outputs derived separately from multiple independent responses PPI test was also conducted against the weight as shown in figure 4. It was observed that the stimulus is constant as the weight changes. At initialization of the sound, the weight is 0.1g and it remains stable, but at a given time of 8000ms, there was drastic change in the weight and when that occur, it was observed in the experiment that the weight was stable but dropped in between 8000ms and 10000ms and later increase in weight to 0.15 at the interval of 8seconds to 10 seconds.



Figure 4: Result of Prepulse – Pulse Trial simulation graph

The last graph figure 5 is Randomization and inter-trial graph, the trials were conducted in the random manner with changing inter trial intervals of 10-20s. The startle reflex responses PPI test was conducted against the weight as shown in figure 5. It was observed that as the stimulus is constant, the weight changes and it keep changing as the inter trial time interval increases. This proves our experiment of Randomization of time intervals right.

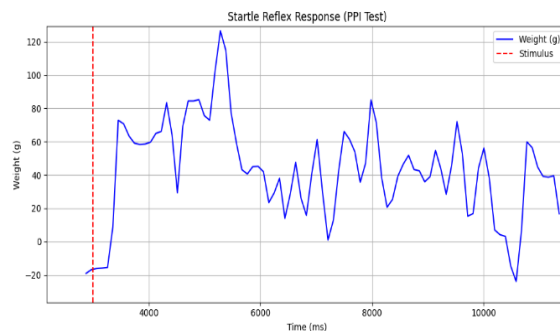


Figure 5: Result of Randomization and Inter-trials simulation

4.1 Training and Performance of Machine Learning Models

All features extracted from the SR waveforms were divided into 70% training samples, 15% validate samples, and 15% test samples following feature engineering. Some models of machine learning from many families including Random Forests, Decision Tree and Logistics Regression, is used to train the models using the training sets. A robust startle classifier is then produced by combining the several machine learning models using a generalized linear model. Using the exact same properties as the trained ensemble model, additional SR waveforms were categorized as startles or non-startles.

4.2 Model Evaluation

Table 1 shows the test accuracy as the quantity of training samples grows with every animal. Generally, the test accuracy grows as the quantity of training datasets for the animal and measurement methodologies increases. 800 SR waveforms were manually categorized for both the measuring method and the animal, using 560 to teach the Random Forest, Logistics Regression and Decision Tree models. The prediction accuracies attained taught by machine learning models distinctly on every data set for animal/startle modality establish the strong and over-all nature of the machine learning startle classification technique presented in this study.

The performance of the models is presented in Table 1.

Table 1: Result of Experiment at 85dB

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	96.0	74.0	80.0	72.0
Logistic Regression	97.0	78.0	82.0	82.0
Decision Tree (Gini)	93.0	92.0	75.0	83.0
Decision Tree (Entropy)	94.0	75.0	84.0	93.9

4.3 Comparison of result with the Existing Method

The Machine learning classifier was compared with Automated Startle classified which is one of the commonly used methods for classification. The machine learning classifier performed better than the Automated Startle classifier. The automated method closely follow the Machine learning classification at lower stimulus intensity of 55dB and 85db which gives 72% and 73.2% respectively.

However, Machine learning classifier performs well than the automated at 85dB intensity. At performance evaluation, the existing classifier (which utilized PASTA dataset), RF, LR, DT (Gini) and DT (Entropy), (all utilizing ASLAS data), had accuracy of 90.0%, 96.0%, 97.0% , 93.0% and 94.0%; precision of 76.0%, 74.0%, 78.0%, 92.0% and 75.0%; recall of 75.0%, 80.0%, 82.0%, 75.0% and 84.0%; and f1-score of 72.0%, 72.0%, 82.0%, 83.0% and 93.9%, respectively. The Machine learning classifier performed better when compared with the existing method regarding F1-score, recall, accuracy, and precision.

The Logistics Regression was found to be the algorithm with most accuracy, Decision Tree (Gini) has best precision while the Decision Tree (entropy) has highest Recall and the best f1-score. The study demonstrates that precision (accuracy) and the amount of time required to construct a model are important factors to be considered. Therefore, to have a successful predictive learning model, algorithms need to be precise, accurate, and minimally error-prone. The result implies that the model predictions are correct and a high F1-score means that the

model performs well on both precision and recall.

5.0 Conclusion and Recommendation

In this study, we introduce a low-cost device in form of controlled typical load cells for the kitchen, which can be utilized to excerpt data in a behavioral experiment. Moreover, we develop a entire code required to redesign the PASTA software includes a Python code for the control computer, a C++ and Arduino code for the communication board, and a R package for data processing, analysis, and visualization. The complete system consists of a locally constructed kitchen scale that is linked to a computer using a mouse cage as a test bed.

Aside the cost, device offers numerous benefits compared to commercially available alternatives since the platform offers raw data that is compatible with various data analysis techniques and enables the researcher to contest for possible confusing factors such as trial speed. Furthermore, the device is very customizable due to its elastic hardware and design. Both the device and the load cell angle can be readily modified with only little changes. The result shows that Machine learning classifier performs better than the Automated classified.

There are several limitations with setup that must be taken into consideration such as the quality and calibration of the equipment, the base for the testbed and the position of the load cell which could have an impact on the precision and repeatability of the results. Additionally, in animals with elevated elastic action, motor responses to auditory stimuli may be covered by spontaneous movements. For this kind of control, synchronous video capture of motor movement can be quite helpful.

Future studies should focus on addressing the setup's limitations and improve its competences. Adding other stimuli in addition to white noise and allowing for more flexibility in the number of consecutive presentations are two possible improvements. Mechanical limitations must be used to restrict the animal's movement without creating undue tension in order to reduce the effect of impulsive actions and guaranty reliable data collection. The setup offers a lot of promise for recordings in multiple modalities and complicated designs for experiments, despite certain constraints. Future studies ought to focus

on improving the suggested setup's capabilities and correcting its shortcomings in order to make the approach more user-friendly and effective for researchers

References

1. Lieu, J., Kenna, M., Anne, S., & Davidson, L. (2020). Hearing Loss in Children: A Review.. JAMA, 324 21, 2195-2205 <https://doi.org/10.1001/jama.2020.17647>
2. Hantsoo, L., Golden, C. E. M., Kornfield, S. L., Grillon, C., & Epperson, C. (2018). Startling Differences: Using the Acoustic Startle Response to Study Sex Differences and Neurosteroids in Affective Disorders. Current psychiatry reports, 20, 40 - 40. <https://doi.org/10.1007/s11920-018-0906-y>
3. Lauer, A.M., Behrens, D., & Klump, G. (2017): Acoustic startle modification as a tool for evaluating auditory function of the mouse: Progress, pitfalls, and potential. *Neuroscience and biobehavioral reviews*. 77, 194–208. [doi:10.1016/j.neubiorev.2017.03.009](https://doi.org/10.1016/j.neubiorev.2017.03.009).
4. Madeline M. Pantoni, Gerald M. Herrera, Kaitlin R. Van Alstyne¹ and Stephan G. Anagnostaras (2020). Quantifying the Acoustic Startle Response in Mice Using Standard Digital Video. *Front Behav Neurosci*. 2020 Jun 3;14:83. doi: 10.3389/fnbeh.2020.00083. PMID: 32595460; PMCID: PMC7289120
5. Gómez-Nieto, R., Hormigo, S., & López, D. E. (2020). Prepulse Inhibition of the Auditory Startle Reflex Assessment as a Hallmark of Brainstem Sensorimotor Gating Mechanisms. *Brain Sciences*, 10. <https://doi.org/10.3390/brainsci10090639>
6. Swerdlow, N., Braff, D., & Geyer, M. (2016). Sensorimotor gating of the startle reflex: what we said 25 years ago, what has happened since then, and what comes next. *Journal of psychopharmacology (Oxford, England)*, 30, 1072 - 1081. <https://doi.org/10.1177/02698811>
7. Longenecker, R. J., Alghamdi, F., Rosen, M. J., & Galazyuk, A. (2016). Prepulse inhibition of the acoustic startle reflex vs. auditory brainstem response for hearing assessment. *Hearing research*, 339, 80-93. <https://doi.org/10.1016/j.heares.2016.06.006>
8. Virag D., Homolak J., Kodvanj I., Babic Perhoc A., Knezovic A., Osmanovic Barilar J., & Salkovic-Petrisic M. (2020). Repurposing a digital kitchen scale for neuroscience research: a complete hardware and software cookbook for PASTA. *Sci Rep*. 11(1): 2963. doi: 10.1038/s41598-021-82710-6. PMID: 33536528; PMCID: PMC7858584.

9. Fawcett, T. J., Longenecker R. J., Chad S. C., and Joseph P. W. (2022). Machine learning, waveform preprocessing and feature extraction methods for classification of acoustic startle waveforms. *MethodsX*.8: 101166. doi: 10.1016/j.mex.2020.101166. PMID: 33354518; PMCID: PMC7744771.
10. Emmanuel O.A., Adeniji O.D., Awoyelu T.M., and Fasae O.D. (2021). Hybridization of Machine Learning Techniques in Predicting Mental disorder. <https://journals.researchparks.org/index.php/IJHCS> e-ISSN: 2615-8159 p-ISSN: 2615-1898 Volume 03 Issue: 6 Aug 2021
11. Stevanovic M et al (2025). Zebra_K, a kinematic analysis automated platform for assessing sensitivity, habituation and prepulse inhibition of the acoustic startle response in adult zebrafish. *Science of The Total Environment*, Volume 958, 178028, ISSN 0048-9697, <https://doi.org/10.1016/j.scitoten.2024.178028>.
12. Oguntunde T. and Abioye C.O. (2025). Development of Comparative Fake Transaction Alert Detection Model Using Machine Learning. *UIJSLICTR Vol .15 No. 1*. ISSN: 2714 3627.
13. Bezerra, T.O., Roque, A.C., & Salum C. (2024). A Computational Model for the Simulation of Prepulse Inhibition and Its Modulation by Cortical and Subcortical Units. *Brain Sci.* 14(5):502. doi: 10.3390/brainsci14050502. PMID: 38790479; PMCID: PMC11118907.
14. Oyeniran, B.G. & Oguntunde, T. & Akinola, S.O. (2020). Development of an Anomaly Detection System for Zero-Day Attacks in Network Traffic Using Ensemble Machine Learning Algorithms. *Advances in Multidisciplinary & Scientific Research Journal Publication*. 11. 97-104. 10.22624/AIMS/CISDI/V11N1P10.
15. Moein, M., Davarpanah, M., Ali, M., & Ataei, M. (2010). Classifying ear disorders using support vector machines. *2010 Second International Conference on Computational Intelligence and Natural Computing*, 1, 321-324. <https://doi.org/10.1109/cinc.2010.5643830>
16. Fabian Pedregosa, G. Varoquaux, Alexandre Gramfort, V. Michel, B. Thirion, O. Grisel, Mathieu Blondel, Gilles Louppe, (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 2825-2830
17. Erwan, S. and Giles Hooker. (2026). Theory of Random Forests. *Annual Review of Statistics and Its Application*. 13:99-121 <https://doi.org/10.1146/annurev-statistics-112723-034707>
18. Ibomoiye, D. M. and Jere, N. (2024). A Survey of Decision Trees: Concepts, Algorithms, and Applications. in *IEEE Access*, vol. 12, pp. 86716-86727, 2024, doi: 10.1109/ACCESS.2024.3416838.
19. Bahzad T. J. and Adnan M. A. (2021) Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of applied science and technology trends* 2(1). 20 – 28. ISSN: 2708-0757
20. Logistic regression and classification (2020). *Predictive Analytics*. <https://doi.org/10.1002/9781119464761.ch7>
21. Njoku, O. (2019). Decision Trees and Their Application for Classification and Regression Problems.
22. Talekar, B. (2020). A Detailed Review on Decision Tree and Random Forest. *Bioscience Biotechnology Research Communications*. <https://doi.org/10.21786/bbrc/13.14/57>
23. Ignatious K. Pious, A. Rajalakshmi, P. K. R, V. C M, M. Nalini and S. S. R, (2024). Enhancing Prediction Accuracy Through Random Forest in Classification and Regression *International Conference on Smart Technologies for Sustainable Development Goals (ICSTSDG)*, Chennai - 600077, Tamil Nadu, India, 2024, pp. 1-6, doi: 10.1109/ICSTSDG61998.2024.11026358.