# Machine Learning Technique to predict Autism Spectrum Disorder using Data Mining

Dina Ayman Abu Taleb
Computer Science Department
Faculty of Computer and Information Science,
Ain Shams University
Cairo, Egypt

Mohmed Mabrouk Morsey
Computer Science Department,
Faculty of Computer and Information Science,
Ain Shams University
Cairo, Egypt

Manal Omar
Medical Studies Department,
Faculty of Postgraduate Childhood Studies,
Ain Shams University
Cairo, Egypt

El-Sayed M. El-Horbaty
Computer Science Department,
Faculty of Computer and Information Science
Ain Shams University
Cairo, Egypt

## ABSTRACT
Autism Spectrum Disorder (ASD) is a developmental disability caused by differences in brain. Scientists believe there are multiple causes of ASD that act together to change the most common ways people develop. The diagnosis process for ASD must be early to provide the required clinical and mental care. In this paper, Machine Learning classifiers were applied using data mining techniques to develop a prediction model for ASD. The dataset utilized comprised 507 instances of children aged between 12 and 36 months. The experimental results show that all classifiers Logistic Regression (LR), Support Vector Machine (SVM), Decision tree, Naïve Bayes, and Artificial Neural Network (ANN) have an excellent accuracy between 95% to 100%. Furthermore, LR, SVM, and ANN have 100% accuracy.

## General Terms
Predictive Modeling, Classification algorithms, Data mining, Healthcare Analytics.

## Keywords
Machine Learning, Classification algorithms, Data mining, Autism Spectrum Disorder.

## 1. INTRODUCTION
The neurodevelopmental disorders appear on the development stage in which the brain process has some differences from a normal person. These differences would be very specific and do not have much influence on the patient's life or would have a wide effect on the patient's life including social and intellectual ability. Sometimes the neurodevelopmental disorders appear before the child enrolls school. There are many disorders under the neurodevelopmental disorders. One of them that highly spread recently is autism spectrum disorder, in which the patient has a deficit in social communication skills, stereotype behaviors, activities, interest and sameness in lifestyle without acceptance of change. ASD may associate with other mental and behavioral disorders such as anxiety [1].

According to the world health organization "WHO", 1 in 100 children has an autism worldwide [24]. The autism spectrum disorder has a level of severity in which each level has a group of symptoms. These levels influence the autistic's life.

The severity levels are:

Level 1: in which the autistic requires support but not a lot to help improve the skills and behaviors.

Level 2: in which the autistic requires substantial support.

Level 3:in which the autistic requires very substantial support [1].

Some symptoms that appear on the autistics: do not have feelings, such as sensitivity of pain, express love, etc., not doing any actions to sounds, not interacting with others and environment, staying lonely, no eye contact or it is difficult to make eye contact, and not laughing [2].

The parents should be aware of their child and notice any irregular behaviors to help in the early detection process and determine if the child has ASD or not. This early detection process will have an impact on the life of the child by providing suitable healthcare services, medicine and reducing the cost of late diagnosis [3]. The diagnosis process is done through many tools. Some of these tools are as follows: ADI-R (Autistic Diagnosis Interview Revised), in which the parents have a clinical interview to answer questions about communication, reciprocal social interactions and restricted behaviors. The second tool ADOS (Autism Diagnostic observation Schedule) is in which the therapist rate quality and response of each task. These tasks related to communication and social response and behaviors.

Additionally, CARS (Childhood Autism Rating Scale) is a diagnosis tool that has 15 evaluation items to assess if the child has ASD or not [4]. The previous tools are time consuming and require a specific skills and qualified people such as doctors and therapists to do it. As a result of that, many researchers did researches in this field trying to apply Machine Learning techniques to help in the prediction process of autism spectrum disorder as early as possible which will help the doctors and therapists in the early diagnosis of the patient and provide the suitable healthcare quickly to provide a good lifestyle for patients.

Chowdhury et al.,[5] published a paper about how to predict ASD using Machine Learning classifiers they applied Machine Learning classifiers: naïve Bayes, KNN, LR, gradient boosting, SVM, decision tree, and MLP classifier to the dataset. After applying data mining flow and using a Machine Learning classifier. The SVM gave the best result with an accuracy of 95%. Vakadkar et al., [6] the authors subtract the process used to determine if the child is susceptible to ASD in its nascent stages by using data mining and Machine Learning algorithms SVM, RFC, Naïve Bayes, LR, and KNN. At the end discussed

the result of the model and evaluation metrices measured in this work.

Jaffer et al.,[7] discussed how to employ data mining to combine behavioral and clinical characteristics with demographic data and provide a way to diagnose ASD in adults quickly, easily, and get an acceptable result using Machine Learning algorithms and data mining tools.

This paper surveys the researches that used Machine Learning techniques in the prediction process for autism spectrum disorder.

The rest of this paper is organized as follows: Section 2: elaborates background and related works of autism predicting methods. The methodology that followed to build a prediction model was presented in Section 3. Section 4: discusses and explains the results that obtained after the experiments. Conclusion and future work given in section 5.

# 2. BACKGROUND AND RELATED WORKS

## 2.1 BACKGROUND:

Machine Learning is used in many areas. It is becoming popular in healthcare sector and used in many processes such as diagnosing, detecting, and prediction diseases. Some researchers made researches in healthcare sector used ML techniques on different topics. One of these topics is autism spectrum disorder. Some of ML classifiers applied on this field were:

Logistic Regression (LR):

LR is an algorithm used to predict the binary outcome either something happens or does not, the outcome lies between 0,1. It represent the relation between dependent variable which is categorical always and independent variable which can be categorical or numeric [20].

Support vector machine (SVM):

SVM is an algorithm that transform the data into linear decision space by determining the hyperplane which separates the training data into different classes [21].

Decision trees:

A decision trees classifier starts with a root which can be an assumption or question or etc., then find a route, this process repeats until completed a route and received the output. Every node in the tree represents a choice and the leaves represent states of the output [21].

Naïve bayes:

Naïve bayes classifier depends on Bayes theorem and conditional probabilities. It calculates if a data belongs to a specific category or not [20, 22].

Artificial neural network (ANN):

ANN is a Machine Learning algorithm in which a set of data parameters is trained, and these parameters are then used to determine the output of the model [21].

Data Mining, is a field related to Machine Learning that become in the forefront of new technologies in the last few years in which a pattern of data discovered automatically or semiautomatically. These patterns must be meaningful and lead to make prediction and decisions. The patterns take two forms black box and transparent box. The main difference between them is whether or not the structure of pattern can be represented, examined, and used to make future helpful decisions [8].

The learning process in data mining is not a theoretical but it is a practical in which the input is a data in form of set and the output takes the form of prediction [8].

## 2.2 RELATED WORKS:

Some studies that utilized Machine Learning classifiers are discussed.:

Shaon et al., (2019) [9] have built a model to predict ASD on adult using Random Forest algorithm. The steps they followed were, data collection in which they used a dataset available online from UCI repository, data preprocessing by applied filters to clean the data, then used prediction algorithm in which they used Random Forest and got 0.96% accuracy.

Suman et al., (2020) [2] have done research on ASD using Machine Learning and deep learning techniques. The goal of this research was to predict and analyze ASD in children, adults, and adolescents. In this study, Naïve Bayes, SVM, logistic regression, KNN, neural networks, and convolutional neural networks were applied to three datasets. The results showed that the convolutional neural network has the highest accuracy compared to other algorithms used in the prediction process.

Vakadkar et al., (2021) [6] applied Machine Learning algorithms to a toddler's dataset to determine if the child has ASD or not. In this paper, the researchers split the dataset into two parts: a training set and a test set, then applied LR, SVM, KNN, Naïve Bayes, and random forest algorithms. The LR has the highest accuracy relative to other algorithms, which is 97.15%.

Sumaia et al., (2022) [7] have proposed a model to detect autism spectrum disorder in adults. In this work, the researchers used an ASD dataset for adults with 21 attributes from the UCI Machine Learning repository and applied the following algorithms: logistic regression, sequential minimum optimization (SMO), Naïve Bayes, and instance-based technique based on k-neighbors (IBK) with Weka tools to classify the data for the adults. The results obtained showed that SMO has the highest accuracy, 99.71%.

Aishwarya et al., (2023) [10] have built a model using Naïve Bayes and decision tree classifier depending on data mining techniques to use it in building web application that can predict ASD in toddlers and kids. The accuracy of the model is high depending on the two classifiers used.

Khan et al., (2023) [11] have a review paper on researches that discusses how to apply Machine Learning techniques to detect or predict ASD. In this paper, the researchers highlighted some lessons that need to be considered. Some of these lessons are: SVM, RF, and LR are the outperforming algorithms in ASD researches, and precision and recall are the most used performance metrics.

Table 1 provides a summary of the related works, highlighting the algorithms utilized and the accuracy levels achieved in each study.

**Table 1 Summery of related works**

| Author | Year | Aim | Dataset | Algorithms | Results |
|---|---|---|---|---|---|
| Shaon et al., [9] | 2019 | To predict ASD in adults based on behavioral attributes. | ASD dataset for adults from UCI | Random forest | 96% |
| Suman et al., [2] | 2020 | To predict and analysis ASD problems in children, adult, and adolescents | ASD datasets for children, adults, and adolescents | Naïve Bayes, SVM, LR, KNN, Neural Network, CNN | CNN gives better accuracy on all datasets as following: 98.30%, 99.53%, 96.88% |
| Vakadkar et al., [6] | 2021 | To detect if the child is having ASD in nascent stages | ASD toddler's dataset from Kaggle | SVM, Random Forest, Naïve Bayes, LR, KNN | LR gives the best accuracy 97.15% |
| Sumaia et al., [7] | 2022 | To use data mining techniques to provide a quick and easy way to predict ASD | ASD dataset for adults from UCI | LR, SMO, Naïve bayes, IBK | SMO gives the best accuracy 99.71% |
| Aishwarya et al., [10] | 2023 | To predict early onset ASD on toddlers and kids. | ASD dataset for kids and toddlers from UCI | Naïve Bayes | 95% |
| Khan et al., [11] | 2023 | To review and displays the ML approaches used to predict ASD | UCI, ABIDE, Kaggle, and real time datasets | Different ML algorithms used in many studies | There is different range of accuracy that gained depending on different factors |

Table 1 illustrates most of the researches are focused on more than one age group and not discussed the age group we selected. Also, they used online datasets that not updated and some of them used one algorithm to build the prediction model. In this paper, we introduce a ML technique to predict ASD in children (single age group) between 12-36 months and build a prediction model that can be used to build a digital product.

## 3. METHODOLOGY

The proposed methodology is based on a Machine Learning architecture for classification, employing data mining techniques to develop a model for predicting ASD. The pseudo-code involved in building the prediction model are outlined as follows:

Step 1: Data Collection:

A dataset that containing relevant features and labels for ASD in children was sourced from the Kaggle website. This dataset focuses on toddlers aged between 12 and 36 months and comprises 17 attributes. The attributes include both behavioral features and individual characteristics. The data was collected through questionnaires filled out by parents or caregivers, which provided insights into the behavioral patterns and characteristics of the children. The dataset is specifically tailored for early diagnosis, which aligns with the study's objectives.

Step 2: Data Preprocessing:

The dataset was loaded on Weka tool and preprocess was performed to handle missing values and apply normalization filter to ensure uniformity and improve model convergence. As the dataset consisted of numeric data, there was no need to convert categorical variables.

Step 3: Feature Selection:
Feature selection was performed to identify the most informative attributes for predicting ASD and to eliminate irrelevant or redundant features. This step ensured that the model focused on critical variables, improving its efficiency and accuracy.

Step 4: Model Selection:

The algorithms selected for the prediction process include Five Machine Learning algorithms were chosen for the prediction process based on their effectiveness in similar classification tasks LR, SVM, Decision Tree, Naïve Bayes, and Networks ANN. These algorithms were selected due to their ability to handle classification problems and their proven success in healthcare analytics and predictive modeling.

Step 5: Model Training:
The dataset was split into two subsets:
Training set: 80% of the dataset, used to train the model and enable it to learn patterns and relationships between input features and ASD diagnosis.
Testing set: 20% of the dataset, reserved for evaluating the model's performance.
Each algorithm was trained on the training data to learn the patterns and relationships between input features and ASD diagnosis. The performance of each algorithm was evaluated on the testing set using metrics such as accuracy, precision, recall, and F-measure.

Step 6: Model Evaluation and Optimization:
The performance of each trained model was evaluated using the testing set. The evaluation metrics included: accuracy, Precision, Recall and F-measure. Based on these metrics, the models were compared, and the best-performing algorithm was identified. Logistic Regression, SVM, and ANN demonstrated the highest accuracy and were considered for further optimization.

Patient data is essential for conducting the prediction process or other medical procedures for disorders., these data will help in doing the prediction process. The patients' data that was used has multiple forms. One of these forms is text data, which is a set of questions and answers for defining the symptoms of the patients. It was collected by conducting an interview or a questionnaire that was answered by the patient or caregiver. The most popular and employable source of text data for autism spectrum disorder is the UC Irvine Machine Learning Repository (UCI), which is an online website created by David Aha containing a dataset, domain theories, and data generators used for analytical purposes [12].
The available dataset on UCI related to ASD is divided into 3 categories: autism Spectrum Disorder Screening Data for Children, Autism Screening Adult, and autism Spectrum Disorder Screening Data for Adolescent. These datasets, created by Dr. Fadi Thabtah, contained 20 attributes: 10 related to individuals characteristics and 10 related to behavioral features. The children dataset contained 292 instances, 704 instances in the adult dataset, and 104 instances in the adolescent dataset [17-19].
Suman et al., [2] applied Machine Learning algorithms to the three types of datasets from the UCI repository to detect ASD. In addition to the previous research, Akter et al., [14] also used the three datasets for children, adults, and adolescents to enhance the accuracy of a Machine Learning model that can detect autism spectrum disorder. Autism screening adults were employed in [5,7,13,15] to predict ASD using Machine Learning algorithms and data mining techniques.
By the same token, Kaggle is an online website that provides datasets in different disciplines. Autism screening data for toddlers is a dataset created by Dr. Fadi Thabtah contained 18 attributes: 10 attributes about behavioral features called them (Q-Chat-10) and the others about individual characteristics with 1054 instances [16]. Looking at Vakadkar et al., [6] they employed classification algorithms on the autism toddler's dataset by splitting the dataset into two parts, 80% of the data used for training and 20% for testing to detect ASD in children. Khan et al., [11] published a review paper outlined on it a group of researches that employed ML techniques to predict or detect ASD in early stages, and they found that
the most used dataset used in these researches were UCI, Kaggle, and ABIDE.
The dataset used for this paper has been collected from Kaggle. The type of dataset contains data for toddlers (from 12-36 months) from Saudi Arabia. The data collected from questionnaire that consists of 17 attributes [16].
The details of dataset are given in Table 2 (see Table 2).

Once the dataset, which serves as the input for the study, is defined, a process known as data preprocessing is applied.
Data Preprocessing is applied on the data before feeding it into the model to ensure that the data is clean, consistent, accurate, and complete as the dataset could have some problems such as there is a missing value, repeated data, discretization, data reduction. After the preprocessing is complete, in which we remove region attribute, the person completing the test attribute, and screening score attribute because these attributes do not help in the purpose of study, the data become ready to use in analysis, training, and building the model.
To build the model, the dataset is split into two parts: the first part is training part with 80% of the dataset, the second part is testing part with 20% of the dataset. The percentage 80:20 is the most commonly used to split the datasets.
The classification algorithms LR, SVM, Decision Tree, Naïve Bayes, and ANN were applied to the training set. The performance of each algorithm was analyzed to determine the most effective model for ASD prediction.
By employing the outlined methodology, the study aimed to develop a robust, accurate, and efficient prediction model capable of aiding early detection of ASD in toddlers.

**Table 2: List of Attributes in the dataset**

| Feature | Description |
|---|---|
| A1: Question 1 | Does your child look at you when you call his/her name? |
| A2: Question 2 | How easy is it for you to get eye contact with your child? |
| A3: Question 3 | Does your child point to indicate that she/he wants something? (e.g. toy that is out of reach) |
| A4: Question 4 | Does your child point to share interest with you? (e.g. pointing an interesting sight) |
| A5: Question 5 | Does your child pretend? (e.g. care for dolls, talks on a toy phone) |
| A6: Question 6 | Does your child follow where you are looking? |
| A7: Question 7 | If you or someone else in the family is visibly upset, does your child show signs of wanting to comfort them? (e.g. stroking hair, hugging them) |
| A8: Question 8 | Would you describe your child's first word as |

| | |
|---|---|
| A9: Question 9 | Does your child use simple gestures? (e.g. wave goodbye) |
| A10: Question 10 | Does your child stare at nothing with no apparent purpose? |
| Region | List of Saudi regions |
| Age | Toddlers (months) |
| Gender | Male or Female |
| Screening Score | 1-10 (the final score obtained based on the scoring algorithm of the screening method used. Less than or equal 3 no ASD traits; > 3 ASD traits) |
| Family member with ASD history | Whether any immediate family member has an ASD |
| Who is completing the test | Parent, other |
| Class | 0 or 1 (0 : no ASD traits, 1: ASD traits) |

## 4. EXPERIMENTAL RESULTS

The experiment took place using Weka 3.9 which is used for Machine Learning and data mining purposes. The prediction model was developed, and the accuracy of various algorithms was measured to determine the most effective one.

**Table 3 The Performance of LR, SVM, Decision trees, Naive Bayes, and ANN**

| Algorithms | Accuracy | Precision | Recall | F-measure |
|---|---|---|---|---|
| LR | 100 | 1.00 | 1.00 | 1.00 |
| SVM | 100 | 1.00 | 1.00 | 1.00 |
| Decision Tree | 95.049 | 0.982 | 0.933 | 0.957 |
| Naïve Bayes | 99.009 | 0.984 | 1.00 | 0.992 |
| ANN | 100 | 1.00 | 1.00 | 1.00 |

Table 3 presents the evaluation of the performance for the algorithms used in terms of the accuracy, precision, recall, and f-measure and compared to the performance of the other algorithms used.

As seen from Table 3 and the chart Figure 1 the Logistic Regression, SVM, and ANN gave the best results with best performance over all other algorithms.

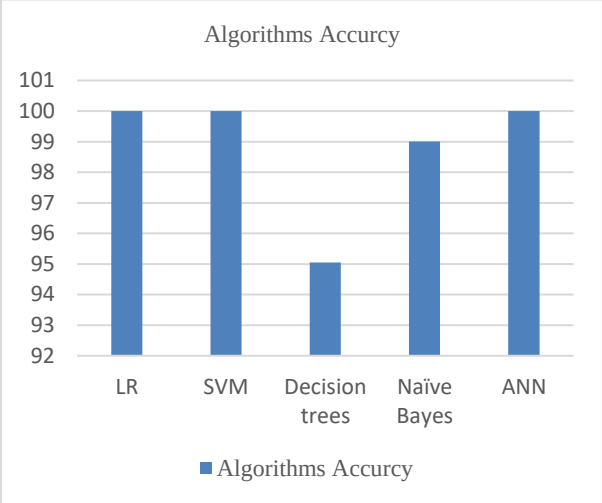


**Figure 1 Comparison between the accuracy of LR, SVM, Decision trees, Naive Bayes, and ANN**

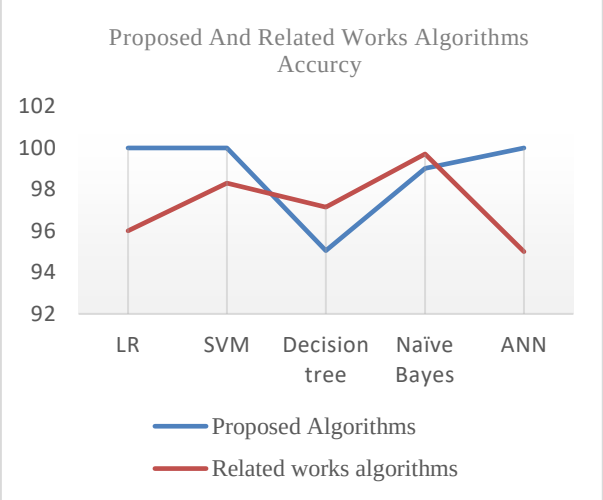


**Figure 2 Comparison of Proposed and Related Works Algorithms Accurcy**

To compare the results with algorithms from related works, Table 4 presents the algorithms applied by researchers to children's datasets for building prediction models, along with their respective accuracy levels. The Table 4 also includes the proposed algorithms and the accuracy achieved in this paper.

**Table 3 Comparative study between our proposed algorithms and related works algorithms**

| Our proposed Algorithms | Accuracy | Related works Algorithms | Accuracy |
|---|---|---|---|
| LR | 100 | Random forest | 96 |
| SVM | 100 | CNN | 98.30 |
| Decision Tree | 95.049 | LR | 97.15 |
| Naïve Bayes | 99.009 | SMO | 99.71 |
| ANN | 100 | Naïve Bayes | 95 |

# 5. CONCLUSION AND FUTURE WORK

A model was developed to predict autism spectrum disorder (ASD) using Weka, a Machine Learning tool, by applying various Machine Learning techniques to a dataset of autistic children. Performance metrics were evaluated to analyze the performance of the models and to identify the model with the highest accuracy for implementation. The findings indicate that LR, SVM, and ANN achieved the highest accuracy of 100% compared to other algorithms. This outcome may be attributed to potential biases within the dataset used**.**

Future research can focus on improving the robustness of the prediction model by addressing dataset limitations and biases. It is recommended to use updated datasets or collect new data specific to age groups to ensure the model's applicability to a broader population. Additionally, integrating these algorithms into digital healthcare tools can facilitate early diagnosis and support by providing reliable predictions for mothers, family doctors, nursery teachers, and pediatricians. Developing an accessible and user-friendly digital platform could enhance the utility of the model, enabling real-time assessments and aiding in better management of ASD through early detection and intervention strategies. Furthermore, exploring ensemble methods and hybrid algorithms could improve the accuracy and reliability of predictions in varied scenarios.

# 6. REFERENCES

[1] American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders, American Psychiatric Association, 2013. (doi: 10.1176/appi.books.9780890425596.)

[2] S. Raj and S. Masood, “Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques,” in Procedia Computer Science, Elsevier B.V., 2020, pp. 994–1004. (doi: 10.1016/j.procs.2020.03.399).

[3] M. D. Hossain, M. A. Kabir, A. Anwar, and M. Z. Islam, “Detecting Autism Spectrum Disorder using Machine Learning,” Sep. 2020, (doi: 10.1007/s13755-021-00145-9).

[4] E. Navarro-Pardo, M. F. López-Ramón, Y. Alonso-Esteban, and F. Alcantud-Marín, “Diagnostic Tools for Autism Spectrum Disorders by Gender: Analysis of Current Status and Future Lines,” Children.

[5] K. Chowdhury and M. A. Iraj, “Predicting Autism Spectrum Disorder Using Machine Learning Classifiers,” in Proceedings - 5th IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology, RTEICT 2020, Institute of Electrical and Electronics Engineers Inc., Nov. 2020, pp. 324–327.(doi: 10.1109/RTEICT49044.2020.9315717).

[6] K. Vakadkar, D. Purkayastha, and D. Krishnan, “Detection of Autism Spectrum Disorder in Children Using Machine Learning Techniques,” SN Computer Science, vol. 2, no. 5, Sep. 2021,(doi: 10.1007/s42979-021-00776-5).

[7] N. Al-Qazzaz, S. Jaffer, I. Abdulazez, and T. Yousif, “Data Mining for Autism Spectrum Disorder detection among Adults,” Al-Nahrain Journal for Engineering Sciences, vol. 25, no. 4, pp. 142–151, Dec. 2022,( doi: 10.29194/njes.25040142).

[8]E. F. Ian H. Witten, “Data Mining Practical Machine Learning Tools and Techniques”, 2005.

[9] S. B. Shuvo, J. Ghosh, and A. S. Oyshi, “A Data Mining Based Approach to Predict Autism Spectrum Disorder Considering Behavioral Attributes,” 2019 10th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2019, Jul. 2019, (doi: 10.1109/ICCCNT45670.2019.8944905).

[10] D. Aishwarya, R. P. Charulatha, M. Niranjan, and K. Malathi, “Effective Prediction of Autism Disease using Data Mining Techniques,” in 6th International Conference on Inventive Computation Technologies, ICICT 2023 - Proceedings, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 761–766. (doi: 10.1109/ICICT57646.2023.10134517).

[11] K. Khan and R. Katarya, “Machine Learning Techniques for Autism Spectrum Disorder: Current trends and future directions,” in 2023 International Conference on Innovative Trends in Information Technology, ICITIIT 2023, Institute of Electrical and Electronics Engineers Inc., 2023.(doi: 10.1109/ICITIIT57246.2023.10068658).

[12] https://archive.ics.uci.edu/about

[13] J. Alwidian*, A. Elhassan, and R. Ghnemat, “Predicting Autism Spectrum Disorder using Machine Learning Technique,” International Journal of Recent Technology and Engineering (IJRTE), vol. 8, no. 5, pp. 4139–4143, Jan. 2020, (doi: 10.35940/ijrte.E6016.018520).

[14] T. Akter, M. I. Khan, M. H. Ali, M. S. Satu, M. J. Uddin, and M. A. Moni, “Improved Machine Learning based Classification Model for Early Autism Detection,” in International Conference on Robotics, Electrical and Signal Processing Techniques, 2021, pp. 742–747. (doi: 10.1109/ICREST51555.2021.9331013).

[15] Brian McNamara, Camila Lora, Donyoung Yang, Fabiana Flores, and Paul Daly, “Machine Learning Classification of Adults with Autism Spectrum Disorder.”

[16] “Autism screening data for toddlers.” Accessed***: Feb. 25, 2024***. [Online]. Available: https://www.kaggle.com/datasets/fabdelja/autism-screening-for-toddlers?resource=download&select=Toddler+data+description.docx

[17] “Autistic Spectrum Disorder Screening Data for Children - UCI Machine Learning Repository.” Accessed: ***Feb. 25, 2024***. [Online]. Available: https://archive.ics.uci.edu/dataset/419/autistic+spectrum+disorder+screening+data+for+children

[18] "Autism Screening Adult - UCI Machine Learning Repository." Accessed: ***Feb. 25, 2024***. [Online]. Available: https://archive.ics.uci.edu/dataset/426/autism+screening+adult

[19] "Autistic Spectrum Disorder Screening Data for Adolescent - UCI Machine Learning Repository." Accessed: ***Feb. 25, 2024.*** [Online]. Available: https://archive.ics.uci.edu/dataset/420/autistic+spectrum+disorder+screening+data+for+adolescent

[20] "5 Types of Classification Algorithms in Machine Learning." Accessed: ***Apr. 20, 2024***. [Online]. Available: https://monkeylearn.com/blog/classification-algorithms/

[21] "8 Types of Machine Learning Classification Algorithms." Accessed: ***Apr. 20, 2024***. [Online]. Available: https://www.akkio.com/post/5-types-of-machine-learning-classification-algorithms

[22] "Classification algorithms: Definition and main models." Accessed: ***Apr. 20, 2024.*** [Online]. Available: https://datascientest.com/en/classification-algorithms-definition-and-main-models

[23] https://ml.cms.waikato.ac.nz//weka

[24] "Autism." Accessed: Feb. 04, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders?gad_source=1&gclid=Cj0KCQjw9vqyBhCKARIsAIIcLMFz4kdZSN_XYJDzkUQiTvgqqBszN_TXCijdJem96yPnt9BsrdoTnkwaAi82EALw_wcB