MAPLE: A Novel Processing Technique for Adult Autism Prediction

Francis Julee Rajam
Department of Computer Science, St. Joseph's College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, Tamil Nadu, India
francisjulee@yahoo.com (corresponding author)

Britto Ramesh Kumar Swakkin
Department of Computer Science, St. Joseph's College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, Tamil Nadu, India
dr.sbritto@outlook.com

Received: 10 March 2025 | Revised: 5 April 2025 | Accepted: 9 April 2025

 DOI: https://doi.org/10.48084/etasr.10864

ABSTRACT

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder that affects individuals throughout their lives. Predicting autism in adults is a critical challenge with significant implications for early intervention and support. Current autism prediction systems in adults frequently encounter difficulties arising from incomplete or noisy data, which can affect the accuracy of prediction. This paper presents MAPLE (Missing data imputation and Anomaly removal for Preprocessing with Learning Enhancement), a novel algorithm designed to address these challenges and improve data quality in existing approaches. Existing systems for the prediction of autism in adults often struggle with incomplete data and noise, which can compromise the accuracy of the predictions. MAPLE addresses these issues by incorporating two key algorithms: MARVEL and SAFARI. MARVEL, the Multi-model Approach for Regression-based Value Estimation of Lost data, efficiently imputes missing values by leveraging diverse regression models. This technique ensures robust imputations even in incomplete data, contributing to more accurate predictions. SAFARI, the Statistical Anomaly Filter with Automated Range Identification, identifies and removes outliers from the imputed dataset, enhancing the model's robustness and generalization capabilities and, as a result, improving data quality. Comprehensive experiments were carried out using a real-world autism prediction dataset for adults to evaluate the performance of MAPLE, focusing on improving data quality. The results demonstrate that MAPLE outperforms existing systems, achieving significantly higher prediction accuracy while effectively addressing data quality challenges. This improvement is crucial for early diagnosis and intervention, ultimately enhancing the quality of life for individuals with autism.

Keywords-Autism Spectrum Disorder (ASD); anomaly detection; machine learning; predictive modelling; data preprocessing

I. INTRODUCTION

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition characterized by various symptoms and behaviors [1]. This disorder significantly affects individuals throughout their lives, from early childhood to adulthood. Although ASD is frequently diagnosed during childhood, its influence persists and evolves as individuals grow and transition to adulthood. The impact of autism is not limited to childhood. Instead, it extends its reach into adulthood, affecting various aspects of daily life, including education, employment, social interactions, and overall well-being. The nuanced nature of autism means that people can experience various challenges and strengths that persist and manifest differently as they age.

Historically, the focus of autism research and intervention has been mainly on children [2]. Recognizing the enduring impact of autism in adulthood has underscored the need for early intervention and tailored support services [3]. Existing systems, such as ADOS-2 (Autism Diagnostic Observation Schedule, Second Edition) and ADI-R (Autism Diagnostic Interview-Revised), have laid a solid foundation for the prediction of autism in adults [4]. The disadvantages associated with existing systems become particularly pronounced due to incomplete data and noise, affecting predictive accuracy. Collectively, these limitations hinder the ability of the systems to provide accurate predictions for adults on the autism spectrum [5].

In [6], the identification of individuals with ASD was explored using various machine learning approaches. In [7], the performance of six classification algorithms was evaluated in predicting ASD. In [8], the focus was on the early diagnosis of ASD by harnessing the capabilities of machine learning approaches. In [9], the focus was on ASD diagnosis, particularly in adolescents and adults. In [10], a comprehensive exploration of ASD analysis and detection was performed, harnessing the capabilities of various machine-learning techniques. In [11], a specialized autism classification algorithm was developed based on a logistic regression model. In [12], an innovative approach to ASD detection was introduced, focusing on structural Magnetic Resonance Imaging (sMRI) and outlier detection. In [13], the quality of the data was investigated, which is one of the fundamental challenges in the prediction of ASD. This approach involved ensemble techniques, including Bayesian boosting and gradient boosted trees. In [14], federated learning was employed for ASD detection. In [15], the detection of adults with autism was investigated, focusing on molecular correlations to better understand the neural systems that underlie predictive procedures in autism.

In [16], a systematic review on prediction-related impairments in ASD was presented, determining common trends and research gaps. In [17], the adult results of children with autism were investigated, developing a typology and prediction model based on childhood characteristics. In [18], the neural correlations of hierarchical predictive procedures in autistic adults were studied, adding to the comprehension of the brain mechanisms involved in prediction. In [19], it was found that low-level prediction-based sensory and motor procedures were not affected in autism, challenging the notion of generalized prediction impairments. In [20], a machine-learning framework was developed to categorize ASD in adults, prioritizing algorithmic precision. In [21], an ASD detection framework was proposed, which combined clinical and computational techniques. In [22], various algorithmic methods were examined to classify ASD, emphasizing the advantages and drawbacks of current methods.

In [23], eye-tracking data and machine learning were used to identify ASD, creating a novel tool for early detection. In [24], machine learning was used to perform psychological evaluations of autism in schools and communities and improve diagnostic accuracy. In [25], machine learning was used to assess developmental milestones in children with autism, providing insights for early intervention. In [26], machine learning was used to improve the Quantitative Checklist for Autism in Toddlers (Q-CHAT), resulting in more effective early autism detection. In [27], 2D video-based pose estimation was used to automatically predict ASD in young children, demonstrating the power of computer vision methods. In [28], feature selection methods in machine learning were investigated to improve ASD classification accuracy, model interpretability, and efficiency.

Predicting autism in adults presents a formidable challenge, mainly due to the nuanced presentation of the condition in this population [29]. Unlike childhood diagnosis, which often involves well-established assessment tools and protocols, the path to detecting autism in adults is less defined [30]. Existing research on ASD prediction and data preprocessing has made significant strides, but several key research gaps and challenges remain unaddressed [31, 32].

II. PROPOSED METHODOLOGY

The proposed approach, MAPLE (Missing data imputation and Anomaly removal for Preprocessing with Learning Enhancement), incorporates several key components designed to enhance data quality and improve predictive accuracy for ASD diagnosis in adults. Figure 1 shows the architecture of the proposed approach.

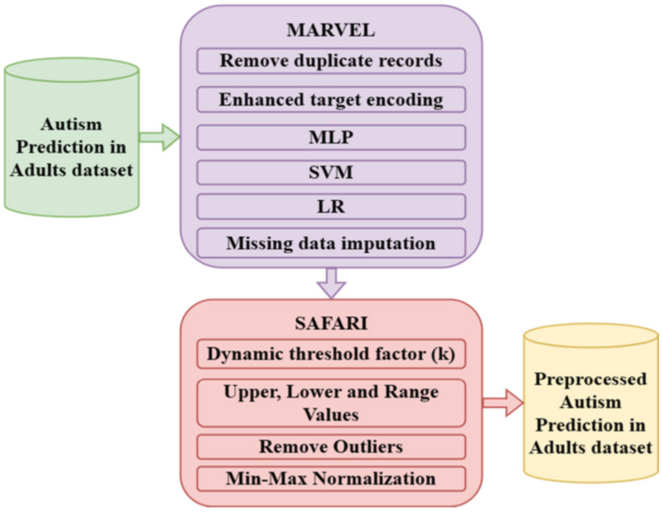

Fig. 1. Architecture of MAPLE.

A. Data Preprocessing

- Data imputation was performed using the MARVEL algorithm to efficiently impute missing values. The first step in the MAPLE algorithm involves handling missing data using the MARVEL technique. This method employs multiple regression models to estimate and fill in missing values.
- Anomaly Detection: After imputing missing values, the next step was to identify anomalies within the dataset. For this purpose, the SAFARI algorithm was utilized, which calculates the lower and upper limits based on the lower 30th and upper 70th percentiles of the data. Instances that fall outside these limits are flagged as anomalies. This step is crucial for ensuring that only normal data points are considered for further analysis.
- Setting Anomaly Threshold: The algorithm establishes a dynamic anomaly threshold, which is defined to classify data points as outliers.
- Creation of Refined Dataset: Following the identification and removal of outliers, MAPLE generates a new dataset called outlier-removed data. Figure 2 shows the pipeline diagram of the data preprocessing process in the MAPLE algorithm.

Figure 2 illustrates a typical preprocessing pipeline for an autism prediction dataset. It starts with duplicate removal, followed by converting categorical features (Gender, Ethnicity, Jaundice, Autism) into numerical representations. Missing values in the Autism column are then imputed using a method called MARVEL, which involves training machine learning models on complete data. Next, anomaly removal (using SAFARI) identifies and removes outliers, such as the individual with Age 40. Finally, the numerical features are normalized using min-max scaling to bring all values within a range of 0 to 1, preparing the data for the training of the machine learning model.

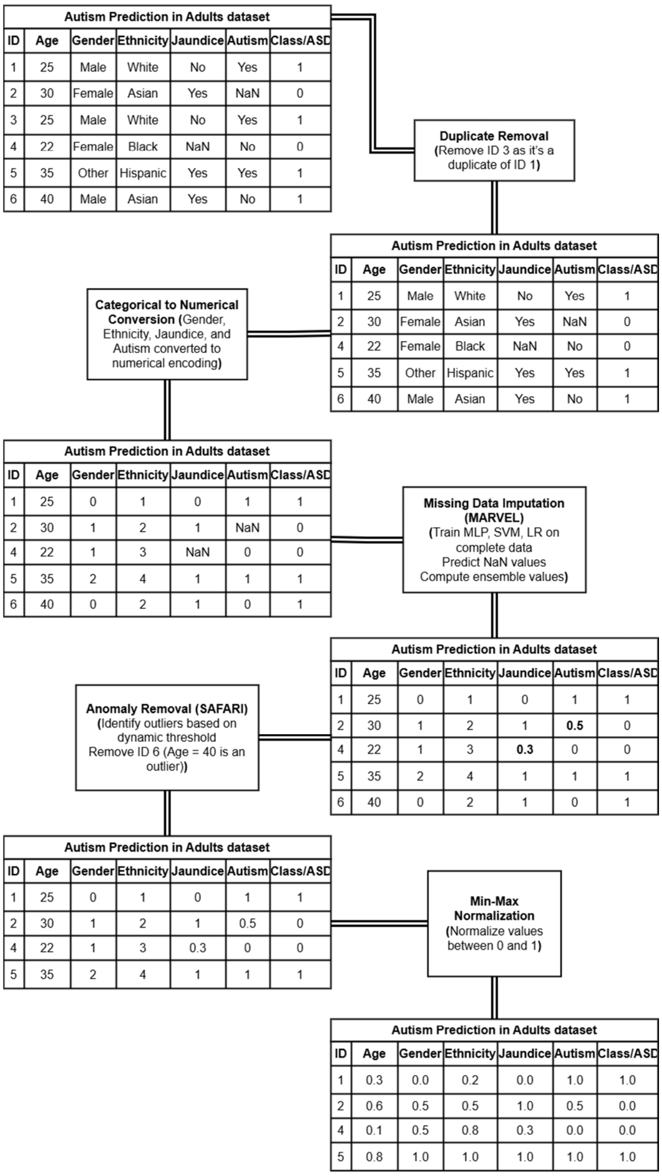

Fig. 2. Data preprocessing.

```
Algorithm 1: MAPLE
import org.apache.commons.lang3.StringUtils;
import org.apache.spark.ml.feature.MinMaxScaler;
import org.apache.spark.ml.feature.MinMaxScalerModel;
import org.apache.spark.sql.Dataset;
import org.apache.spark.sql.Row;
import org.apache.spark.sql.SparkSession;
import java.util.ArrayList;
import java.util.List;

public class AutismDatasetPreprocessor {
  public static void main(String[] args) {
    SparkSession spark =
    SparkSession.builder().appName("Autism Dataset
    Preprocessor").getOrCreate();
  // Load the Autism Prediction in Adults dataset
    Dataset<Row> autism Dataset =
      spark.read().format("csv")
        .option("header", "true")
        .option("inferSchema", "true")
        .load("autism_dataset.csv");
    // Remove duplicate records
    Dataset<Row> duplicatesRemovedDataset =
      removeDuplicates(autismDataset);
    // Perform categorical to numerical conversion
    // using enhanced target encoding
    Dataset<Row> categoricalToNumericalDataset =
      categoricalToNumericalConversion(
        duplicatesRemovedDataset);
    // Apply MARVEL algorithm for missing data
    // imputation
    Dataset<Row> imputedDataset =
      applyMARVEL(categoricalToNumericalDataset);
    // Use SAFARI algorithm for outlier removal
    Dataset<Row> outlierRemovedDataset =
    applySAFARI(imputedDataset);
    // Apply Min-Max normalization
    Dataset<Row> normalizedDataset =
     applyMinMaxNormalization(outlierRemovedDataset);
    // Return preprocessed dataset
    normalizedDataset.show();
  }

  private static Dataset<Row>
    removeDuplicates(Dataset<Row> dataset) {
    return dataset.dropDuplicates();
  }

  private static Dataset<Row>
    categoricalToNumericalConversion(Dataset<Row>
      dataset) {
    // Implement enhanced target encoding logic here
    // For demonstration purposes, assume a simple
    // one-hot encoding
    return dataset;
  }

  private static Dataset<Row>
    applyMARVEL(Dataset<Row> dataset) {
    // Implement MARVEL algorithm logic here
    // For demonstration purposes, assume a simple
    // mean imputation
    return dataset;
  }

  private static Dataset<Row>
    applySAFARI(Dataset<Row> dataset) {
    // Implement SAFARI algorithm logic here
    // For demonstration purposes, assume a simple z-
    // score-based outlier removal
    return dataset;
  }

  private static Dataset<Row>
    applyMinMaxNormalization(Dataset<Row> dataset) {
    MinMaxScaler scaler = new MinMaxScaler()
      .setInputCol("features")
      .setOutputCol("scaledFeatures");
```

```
    MinMaxScalerModel scalerModel =
      scaler.fit(dataset);
    return scalerModel.transform(dataset);
  }
}
```

```
Algorithm 2 MARVEL
Input: Dataset with missing values (data)
Output: Imputed dataset
1. Separate the dataset into complete (train Data)
   and incomplete (test Data) data.
2. Identify attributes with missing values (missing
   Attributes).
3. For each missing attribute in missing attributes:
   a. Train regression models (MLP, SVM, LR) on
      training data.
   b. Predict missing values in test data using
      trained regression models (MLP, SVM, LR).
   c. Calculate the ensemble imputed value as the
      average of predicted values from all models.
4. Replace missing values in test Data with ensemble
   imputed values.
5. Combine the training data and imputed test data to
   create the imputed dataset (imputed Data).
6. Return imputed data.
```

```
Algorithm 3 SAFARI
Input: Imputed dataset (data)
Output: Dataset with outliers removed
1. Load the imputed dataset as instances (data
   Instances).
2. Define a dynamic threshold factor (k) based on the
   dataset size.
3. Initialize an array for anomaly counts (anomalies)
   with the same size as data instances.
4. For each numeric attribute in data instances:
   a. Extract the attribute values into an array
     (attribute Values).
   b. Sort the attribute values in ascending order.
   c. Calculate the lower (Lower) and upper (Upper)
      values based on the sorted data using
      percentiles
      (e.g., lower 30th and upper 70th percentiles)
   d. Calculate the attribute’s range (Range =
        Upper - Lower).
   e. Calculate lower and upper limits (Lower Limit,
      Upper Limit) using k and range.
   f. Identify instances with values between the
      Lower and Upper limits as normal; otherwise,
      consider them anomalies and update the anomaly
      counts.
5. Set an anomaly threshold (anomaly Threshold) to
   determine outliers (e.g., two or more anomalies).
6. Create a new dataset without outliers (outlier
   Removed Data) by filtering instances with
   anomalies below the threshold.
7. Return the outlier-removed data.
```

III. RESULTS AND DISCUSSION

A comprehensive evaluation of the performance of the MAPLE algorithm in ASD detection for adults was carried out, comparing it with existing algorithms [29-31]. The MAPLE algorithm was run using the Java Development Kit (JDK) version 1.8, and the experimental code was implemented and executed using the Apache NetBeans IDE 15, resulting in a seamless creation and evaluation workflow. The dataset used for the MAPLE algorithm is a real-world autism prediction in adults dataset [33]. This dataset often contains incomplete and noisy data, which can significantly affect the accuracy of predictions. This dataset comprises various features, including patient ID, scores based on the Autism spectrum Quotient (AQ) 10-item screening tool (A1_Score to A10_Score), age, gender, ethnicity, history of jaundice at birth, family history of autism diagnosis, country of residence, previous screening test history, AQ1-10 screening test results, patient age description, relation of the patient who completed the test, and the target column labeled Class/ASD (0 for No and 1 for Yes).

A. *Performance Metrics*

Table I shows several performance metrics that were employed to assess the effectiveness of the MAPLE algorithm [34]. These metrics include accuracy, precision, recall, F1-score, sensitivity, and specificity [36]

TABLE I. PERFORMANCE METRICS

Metrics	Formula
Accuracy [34]	$\frac{TP+TN}{TP+TN+FP+FN}$
Precision [34]	$\frac{TP}{TP+FP}$
Recall [34]	$\frac{TP}{TP+FN}$
F1-measure [34]	$\frac{(2*Precision*Recall)}{(Precision+Recall)}$
Specificity [35]	$\frac{TN}{TN+FP}$

Table II presents a comparison of the performance metrics of the MAPLE algorithm against three other models [29-31]. The metrics evaluated include accuracy, precision, F1-score, sensitivity (recall), and specificity. Using these metrics, the findings and performance of the proposed model is examined below.

1) *Accuracy*

This metric indicates the overall correctness of the predictions made by the models. MAPLE achieved an accuracy of 91.6312%, which is higher than that of the other models. This increased accuracy is mainly due to MAPLE's sophisticated preprocessing methods, particularly the data imputation and anomaly elimination steps.

TABLE II. PERFORMANCE METRICS COMPARISON

Metrics	[29]	[30]	[31]	MAPLE
Accuracy (%)	81	76.8	90.6	91.6312
Precision (%)	85.26	78.29	90.58	91.76
Recall (%)	83.08	75.29	90.60	91.69
F1-measure (%)	81	72.5	90.62	91.63
Specificity (%)	86	79.9	90.58	91.76

2) *Precision*

MAPLE scored 91.76%, again outperforming the other models. The MAPLE method, when integrated with SVM classification, achieved the highest precision (91.76%) among the algorithms tested. This high precision is due to MAPLE's resilient preprocessing, which ensures that the input data is clean, noise-free, and enhanced for classification. By efficiently converting features, MAPLE allows the SVM classifier to differentiate between relevant and irrelevant data, resulting in a significant reduction in false positives. This preprocessing benefit ensures that the SVM model consistently detects true positives with high accuracy.

3) F1-Score

The comparison of F1-score demonstrates the better efficiency of the MAPLE algorithm, with the highest F1-score (91.69%) among all techniques. This is due to its distinctive preprocessing methods, such as MARVEL missing data imputation and SAFARI outlier removal. By imputing missing data, MAPLE ensures that the dataset is complete and representative, avoiding information loss that could jeopardize the classifier's effectiveness. Additionally, eliminating outliers improves data excellence and mitigates the influence of extreme values, allowing the SVM classifier to focus on significant trends.

4) Sensitivity (Recall)

This metric reflects the model's ability to identify all relevant cases (true positives). MAPLE achieved a sensitivity of 91.63%, which is significantly higher than that of [30]. Sensitivity assesses the model's capacity to accurately detect individuals with ASD.

5) Specificity

This measures the model's ability to correctly identify negative cases (non-ASD individuals). MAPLE's specificity is 91.76%, indicating it is also effective at classifying non-ASD individuals accurately. Specificity assesses the model's capacity to accurately identify individuals without ASD.

The experimental findings show that the MAPLE algorithm surpassed existing models in predicting autism in adults, as evidenced by higher accuracy, sensitivity, and specificity values. MAPLE's success can be attributed to its data preprocessing technique, which addresses typical issues such as missing data and outliers. MAPLE ensures high-quality data by utilizing the MARVEL algorithm for missing data imputation and the SAFARI algorithm for anomaly elimination, resulting in better model efficiency. These advances show that MAPLE is a potential technique for autism prediction, with the ability to boost early detection and intervention for individuals with autism.

IV. CONCLUSION

The MAPLE algorithm emerges as a groundbreaking solution to address critical challenges in the prediction of adult ASD. One of the most notable innovations of MAPLE is its dual approach to improving data quality. The MARVEL algorithm employs multiple regression models to impute missing values, ensuring that the imputed data are not only accurate but also reflective of the underlying patterns in the dataset. The SAFARI component of MAPLE introduces a sophisticated method for identifying and removing outliers. By setting dynamic thresholds based on statistical percentiles, SAFARI ensures that only valid data points are retained for analysis. This is particularly important in the context of autism prediction, where outliers can skew results and lead to misdiagnosis. The meticulous identification of anomalies enhances the integrity of the dataset, making MAPLE a more reliable tool compared to existing algorithms. The experimental results demonstrated MAPLE's remarkable superiority over existing techniques, showcasing its potential to significantly improve the accuracy of ASD diagnosis in adult populations. Future work could further refine and validate the algorithm on diverse datasets, along with exploring potential enhancements and extensions, unlocking even greater advancement in autism prediction and healthcare.

REFERENCES

[1] A. Genovese and M. G. Butler, "Clinical Assessment, Genetics, and Treatment Approaches in Autism Spectrum Disorder (ASD)," *International Journal of Molecular Sciences*, vol. 21, no. 13, Jul. 2020, Art. no. 4726, https://doi.org/10.3390/ijms21134726.

[2] K. Khowaja *et al.*, "Augmented Reality for Learning of Children and Adolescents With Autism Spectrum Disorder (ASD): A Systematic Review," *IEEE Access*, vol. 8, pp. 78779–78807, 2020, https://doi.org/10.1109/ACCESS.2020.2986608.

[3] P. M. Dietz, C. E. Rose, D. McArthur, and M. Maenner, "National and State Estimates of Adults with Autism Spectrum Disorder," *Journal of Autism and Developmental Disorders*, vol. 50, no. 12, pp. 4258–4266, Dec. 2020, https://doi.org/10.1007/s10803-020-04494-4.

[4] V. Allgar, B. Wright, A. Taylor, A. L. Couter, and H. Phillips, "Diagnosing Autism Spectrum Disorders in Deaf Children Using Two Standardised Assessment Instruments: The ADIR-Deaf Adaptation and the ADOS-2 Deaf Adaptation," *Journal of Clinical Medicine*, vol. 10, no. 19, Sep. 2021, Art. no. 4374, https://doi.org/10.3390/jcm10194374.

[5] M. E. Alqaysi, A. S. Albahri, and R. A. Hamid, "Diagnosis-Based Hybridization of Multimedical Tests and Sociodemographic Characteristics of Autism Spectrum Disorder Using Artificial Intelligence and Machine Learning Techniques: A Systematic Review," *International Journal of Telemedicine and Applications*, vol. 2022, pp. 1–26, Jul. 2022, https://doi.org/10.1155/2022/3551528.

[6] A. Novianto and M. D. Anasanti, "Autism Spectrum Disorder (ASD) Identification Using Feature-Based Machine Learning Classification Model," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 17, no. 3, Jul. 2023, Art. no. 259, https://doi.org/10.22146/ijccs.83585.

[7] S. F. Mohd Radzi, M. S. Hassan, and M. A. H. Mohd Radzi, "Comparison of classification algorithms for predicting autistic spectrum disorder using WEKA modeler," *BMC Medical Informatics and Decision Making*, vol. 22, no. 1, Nov. 2022, Art. no. 306, https://doi.org/10.1186/s12911-022-02050-x.

[8] R. Francese and X. Yang, "Supporting autism spectrum disorder screening and intervention with machine learning and wearables: a systematic literature review," *Complex & Intelligent Systems*, vol. 8, no. 5, pp. 3659–3674, Oct. 2022, https://doi.org/10.1007/s40747-021-00447-1.

[9] C. Küpper *et al.*, "Identifying predictive features of autism spectrum disorders in a clinical sample of adolescents and adults using machine learning," *Scientific Reports*, vol. 10, no. 1, Mar. 2020, Art. no. 4805, https://doi.org/10.1038/s41598-020-61607-w.

[10] S. Raj and S. Masood, "Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques," *Procedia Computer Science*, vol. 167, pp. 994–1004, 2020, https://doi.org/10.1016/j.procs.2020.03.399.

[11] H. Sewani and R. Kashef, "An Autoencoder-Based Deep Learning Classifier for Efficient Diagnosis of Autism," *Children*, vol. 7, no. 10, Oct. 2020, Art. no. 182, https://doi.org/10.3390/children7100182.

[12] K. Devika, D. Mahapatra, R. Subramanian, and V. R. M. Oruganti, "Outlier-Based Autism Detection Using Longitudinal Structural MRI," *IEEE Access*, vol. 10, pp. 27794–27808, 2022, https://doi.org/10.1109/ACCESS.2022.3157613.

[13] R. Asgarnezhad, K. A. M. Alhameedawi, and H. Akram Mahfoud, "An Effective Model Of Autism Spectrum Disorder Using Machine Learning," *Indonesian Journal of Electrical Engineering and*

Informatics (IJEEI), vol. 11, no. 2, pp. 389–401, Apr. 2023, https://doi.org/10.52549/ijeei.v11i2.4060.

[14] M. S. Farooq, R. Tehseen, M. Sabir, and Z. Atal, "Detection of autism spectrum disorder (ASD) in children and adults using machine learning," *Scientific Reports*, vol. 13, no. 1, Jun. 2023, Art. no. 9605, https://doi.org/10.1038/s41598-023-35910-1.

[15] L. A. Sapey-Triomphe, J. Temmerman, N. A. J. Puts, and J. Wagemans, "Prediction learning in adults with autism and its molecular correlates," *Molecular Autism*, vol. 12, no. 1, Oct. 2021, Art. no. 64, https://doi.org/10.1186/s13229-021-00470-6.

[16] J. Cannon, A. M. O'Brien, L. Bungert, and P. Sinha, "Prediction in Autism Spectrum Disorder: A Systematic Review of Empirical Evidence," *Autism Research*, vol. 14, no. 4, pp. 604–630, Apr. 2021, https://doi.org/10.1002/aur.2482.

[17] A. Pickles, J. B. McCauley, L. A. Pepa, M. Huerta, and C. Lord, "The adult outcome of children referred for autism: typology and prediction from childhood," *Journal of Child Psychology and Psychiatry*, vol. 61, no. 7, pp. 760–767, Jul. 2020, https://doi.org/10.1111/jcpp.13180.

[18] L. A. Sapey-Triomphe, L. Pattyn, V. Weilnhammer, P. Sterzer, and J. Wagemans, "Neural correlates of hierarchical predictive processes in autistic adults," *Nature Communications*, vol. 14, no. 1, Jun. 2023, Art. no. 3640, https://doi.org/10.1038/s41467-023-38580-9.

[19] J. J. S. Finnemann, K. Plaisted-Grant, J. Moore, C. Teufel, and P. C. Fletcher, "Low-level, prediction-based sensory and motor processes are unimpaired in Autism," *Neuropsychologia*, vol. 156, Jun. 2021, Art. no. 107835, https://doi.org/10.1016/j.neuropsychologia.2021.107835.

[20] N. A. Mashudi, N. Ahmad, and N. M. Noor, "Classification of adult autistic spectrum disorder using machine learning approach," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 3, Sep. 2021, Art. no. 743, https://doi.org/10.11591/ijai.v10.i3.pp743-751.

[21] Md. J. Uddin *et al.*, "An Integrated Statistical and Clinically Applicable Machine Learning Framework for the Detection of Autism Spectrum Disorder," *Computers*, vol. 12, no. 5, Apr. 2023, Art. no. 92, https://doi.org/10.3390/computers12050092.

[22] S. G. Jacob, M. M. Bait Ali Sulaiman, and B. Bennet, "Algorithmic Approaches to Classify Autism Spectrum Disorders: A Research Perspective," *Procedia Computer Science*, vol. 201, pp. 470–477, 2022, https://doi.org/10.1016/j.procs.2022.03.061.

[23] I. A. Ahmed *et al.*, "Eye Tracking-Based Diagnosis and Early Detection of Autism Spectrum Disorder Using Machine Learning and Deep Learning Techniques," *Electronics*, vol. 11, no. 4, Feb. 2022, Art. no. 530, https://doi.org/10.3390/electronics11040530.

[24] S. Ahmed, Md. F. Hossain, S. B. Nur, M. Shamim Kaiser, and M. Mahmud, "Toward Machine Learning-Based Psychological Assessment of Autism Spectrum Disorders in School and Community," in *Proceedings of Trends in Electronics and Health Informatics*, vol. 376, M. S. Kaiser, A. Bandyopadhyay, K. Ray, R. Singh, and V. Nagar, Eds. Springer Nature Singapore, 2022, pp. 139–149.

[25] M. M. Haque *et al.*, "Informing Developmental Milestone Achievement for Children With Autism: Machine Learning Approach," *JMIR Medical Informatics*, vol. 9, no. 6, Jun. 2021, Art. no. e29242, https://doi.org/10.2196/29242.

[26] G. Tartarisco *et al.*, "Use of Machine Learning to Investigate the Quantitative Checklist for Autism in Toddlers (Q-CHAT) towards Early Autism Screening," *Diagnostics*, vol. 11, no. 3, Mar. 2021, Art. no. 574, https://doi.org/10.3390/diagnostics11030574.

[27] N. Kojovic, S. Natraj, S. P. Mohanty, T. Maillart, and M. Schaer, "Using 2D video-based pose estimation for automated prediction of autism spectrum disorders in young children," *Scientific Reports*, vol. 11, no. 1, Jul. 2021, Art. no. 15069, https://doi.org/10.1038/s41598-021-94378-z.

[28] B. Elshoky, O. Ibrahim, and A. Ali, "Machine Learning Techniques Based on Feature Selection for Improving Autism Disease Classification," *International Journal of Intelligent Computing and Information Sciences*, vol. 21, no. 2, pp. 65–81, Jul. 2021, https://doi.org/10.21608/ijicis.2021.61582.1058.

[29] M. S. Ahammed, S. Niu, M. R. Ahmed, J. Dong, X. Gao, and Y. Chen, "Bag-of-Features Model for ASD fMRI Classification using SVM," in *2021 Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS)*, Shenyang, China, Jan. 2021, pp. 52–57, https://doi.org/10.1109/ACCTCS52002.2021.00019.

[30] W. Liu, M. Liu, D. Yang, M. Wang, and T. Tao, "Automatic diagnosis of autism based on functional magnetic resonance imaging and elastic net," in *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, Chongqing, China, Jun. 2020, pp. 104–108, https://doi.org/10.1109/ITOEC49072.2020.9141766.

[31] Y. Wang, J. Wang, F. X. Wu, R. Hayrat, and J. Liu, "AIMAFE: Autism spectrum disorder identification with multi-atlas deep feature representation and ensemble learning," *Journal of Neuroscience Methods*, vol. 343, Sep. 2020, Art. no. 108840, https://doi.org/10.1016/j.jneumeth.2020.108840.

[32] R. Srinivasan, R. Korah, and M. Ravichandran, "Revolutionizing Diagnostic Insights: Exploring Advanced Image Processing Techniques and Neural Networks in Traditional Indian Medicine," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19214–19220, Feb. 2025, https://doi.org/10.48084/etasr.8975.

[33] "Autism Prediction." Kaggle, [Online]. Available: https://kaggle.com/autismdiagnosis.

[34] F. Thabtah, "Autism Spectrum Disorder Screening: Machine Learning Adaptation and DSM-5 Fulfillment," in *Proceedings of the 1st International Conference on Medical and Health Informatics 2017*, Feb. 2017, pp. 1–6, https://doi.org/10.1145/3107514.3107515.

[35] F. Thabtah, "Machine learning in autistic spectrum disorder behavioral research: A review and ways forward," *Informatics for Health and Social Care*, vol. 44, no. 3, pp. 278–297, Jul. 2019, https://doi.org/10.1080/17538157.2017.1399132.