Improving Early Autism Detection with Chi-Square Feature Selection, Machine Learning, and Explainable AI

Aymen Abu-Errub
Department of Computer Science, Faculty of Information Technology, Jadara University, Irbid, Jordan
A.Abu-Errub@jadara.edu.jo (corresponding author)

Received: 27 June 2025 | Revised: 25 July 2025 | Accepted: 14 August 2025

 DOI: https://doi.org/10.48084/etasr.12999

ABSTRACT

This study presented a framework that utilized Chi-square feature selection and Machine Learning (ML) classifiers to improve the early detection of Autism Spectrum Disorder (ASD) for children 12 to 36 months old. Six classifiers -Light Gradient Boosting Machine (LGBM), Extra Trees (ET), Decision Tree (DT), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Multi-Layer Perceptron (MLP) -were tested. The findings revealed that the integration of Chi-square feature selection with SVM achieved perfect accuracy, precision, recall, and F1-score, while the other models demonstrated notable gains (up to 90%). Additionally, a SHapley Additive exPlanation (SHAP) analysis was conducted to interpret the model predictions and highlight the key behavioral features, while a literature comparison with recent research showed that the current method outperformed the latter. This study demonstrated that integrating robust feature selection with explainable ML models can significantly advance the reliability of early ASD screening tools.

Keywords- Autism Spectrum Disorder (ASD); chi-square feature selection; machine learning; explainable AI

I. INTRODUCTION

The ASD is a complex neuro-developmental abnormality involving restrictions on the social interaction, communication, and repetitive behaviors. The early identification of ASD is significant since it helps a child engage in timely intervention during early years, which substantially improves the future outcomes in cognitive, social, and adaptive functioning [1]. However, conventional diagnosis that relies on behavioral assessments and clinical observation, has often led to serious delays. In [2], it was noted that it took three years for a child to diagnose autism, which prevented the access to early intervention services. Such waits highlight the need for more efficient and accurate diagnostic tools that will identify ASD at an early stage, preferably on toddlers aged 12-36 months.

The advancements in ML show promise in addressing the limitations of traditional diagnostic approaches, particularly in promoting early ASD detection [3-5]. The effectiveness of these models, however, is highly dependent on the quality and relevance of the input parameters. Many datasets include features that are redundant or irrelevant, which can confuse the model and lower its accuracy. For these reasons, robust feature selection is appropriate in order to enhance the predictive accuracy. For instance, chi-square feature selection has emerged as an effective method for identifying the most relevant parameters by ranking the input attributes according to their statistical significance and their dependency on the class label [6].

Beyond the predictive accuracy, AI-driven tools should also be trusted and interpretable. This study integrates Explainable Artificial Intelligence (XAI), specifically SHAP, to provide insights into the model decision-making. XAI techniques not only strengthen the clinical trust by offering clear justifications for predictions, but also help identify the key behavioral markers that drive the diagnostic outcomes [7]. In this work, a comprehensive framework is proposed for early ASD detection in toddlers, combining Chi-square feature selection with a suite of advanced ML classifiers, including Light Gradient Boosting Machine (LGBM), ET, DT, KNN, SVM, and MLP.

II. RELATED WORK

Research has demonstrated the potential of ML in ASD detection, while also revealing persistent methodological gaps. Authors in [8] evaluated several ML models in a child-focused study, such as SVM, Random Forest (RF), and Logistic Regression (LR), with their results demonstrating that LR was the most accurate model. However, many limitations were detected due to the lack of sufficiently large, open-source ASD datasets. Similarly, in [9] the importance of early detection was highlighted, particularly in the first two years of life. Their neural network model delivered superior prediction capacity, sensitivity, and specificity across all age groups compared to traditional methods. Authors in [10] also proposed a comprehensive ML framework incorporating four feature scaling techniques: quantile transformer, power transformer, normalizer, and max abs scaler, with eight classifiers, including AdaBoost and Linear Discriminant Analysis (LDA). High accuracy was observed on four age-specific ASD datasets, with AdaBoost achieving values of 99.25% for toddlers and 97.95% for children. However, the authors highlighted that larger datasets are critical for the generalization of the model.

Other studies have focused on improving the feature selection, Authors in [5], applied methods, such as Boruta and Recursive Feature Elimination (RFE) to identify the optimal feature subsets for ASD detection. They used SHAP to provide insights into the importance of specific features, though they acknowledged that further validation across diverse datasets is needed. Authors in [11] developed an integrated ML framework for ASD detection in toddlers, emphasizing the importance of data transformation and feature selection. Techniques, such as the Synthetic Minority Over-sampling Technique (SMOTE) and RFE, were employed to address the class imbalance and optimize the feature subsets. AdaBoost was highlighted as the optimum classifier, whose hyperparameter tuning further boosted its accuracy. Accordingly, authors in [12] conducted a comparative analysis on classifiers, demonstrating that RF achieved the highest accuracy for adolescent (93.69%) and toddler (93.33%) datasets.

Recent works have further advanced the explainable AI in autism and mental health. Authors in [13] proposed an interpretable IoT-based EfficientNet model for emotion recognition in autistic children, integrating deep learning with local interpretability methods (LIME, Grad-CAM) to highlight the influential regions in facial and physiological data, achieving high accuracy and transparent results. Similarly, authors in [14] introduced the Explainable Mental Health Disorders (EMHD) model, which combined ensemble feature selection with SHAP explanations to classify the mental health disorders in young children, including toddlers. These efforts underscored the growing importance of transparency in ML, especially for clinical decision support.

Based on previous studies, this research presented a rigorous and unified framework that addresses persistent challenges in early ASD detection for toddlers. Specifically, Chi-square statistical tests were applied to systematically identify the most relevant behavioral and demographic features, utilize SVM-SMOTE to create a more balanced and representative training set, and apply MinMaxScaler for consistent feature scaling. The current methodology compared six advanced classifiers both before and after feature selection, with Bayesian hyperparameter optimization refining the model performance. Until now, this is the first study to systematically combine Chi-square feature selection, SVM-SMOTE balancing, Bayesian optimization, and SHAP-based explanation for toddler ASD screening.

III. THE PROPOSED METHOD

A. Data Preprocessing

Preprocessing is referred to the process of data cleaning and transformation that prepares them for use as inputs into ML models. The first step involved the removal of non-informative attributes, such as 'Case no'. and 'Who is completing the test', which were considered irrelevant to the predictive modeling objectives. The current dataset involved no missing values and duplicates, so there was no need for data cleaning.

Subsequent preprocessing steps included data balancing and feature transformation. For data balancing, the SVM-SMOTE method was utilized [15]. This method seeks to optimize the synthetic sample generation process by utilizing SVM for guidance. Unlike standard SMOTE, which interpolates minority class instances, SVM-SMOTE recognizes support vectors - critical data points defining the class boundaries - and generates synthetic samples quite close to these vectors. This targeted approach helps the model better capture the class boundaries and improves the classification performance on imbalanced datasets.

Feature transformation was applied in two ways: encoding and scaling. Categorical variables, such as Sex, Jaundice, Ethnicity, Region, Family_mem_with_ASD, and Class/ASD_Traits were label-encoded, converting categories into consistent numerical values. This ensured that the categorical attributes could be processed by ML models while maintaining equal importance across the categories. For scaling, the MinMaxScaler technique was performed, which re-scaled the possible values according to the maximum and minimum values within each feature, typically [0, 1].

After the preprocessing steps, the dataset was randomly partitioned into training, validation, and testing subsets using a stratified split to preserve the class distribution. Specifically, 70% of the data were used for training, 10% for validation, and 20% for final testing. The training set was employed for model fitting, the validation set for hyperparameter tuning and model selection, while the independent test set was held out and used exclusively to report the final performance metrics presented in this study.

B. Chi-Square Feature Selection (Chi^2)

In predictive modeling, irrelevant or noisy attributes can lead to underfitting or overfitting. Underfitting occurs when the model can be recognized by too little learning from the training data; while overfitting happens when a model learns too much from the training data and can no longer generalize to new or unseen data.

The Chi-square feature selection technique addresses this need by statistically assessing the relationship between each categorical input feature and the target variable. For each feature, the method compares the observed (actual) frequencies of the values across classes with the expected frequencies. A Chi-square score closer to 1 suggests that the feature is most dependent on the predicted class; hence, it would be classified as relevant. By applying this method, data dimensionality was reduced, potential bias was minimized, and more effective classifications of the ASD and non-ASD cases were facilitated.

C. Machine Learning Classifiers

1) Lightweight Gradient Boosting Machine

LGBM is a gradient boosting framework optimized for speed and efficiency. It employs histogram-based techniques to reduce the memory usage and computational time. Innovations, such as leaf-wise tree growth, enable LGBM to handle large, high-dimensional datasets while maintaining high accuracy [16].

2) Extra Trees

The ET algorithm is an ensemble method that constructs a collection of unpruned decision trees using the entire training dataset. Unlike traditional algorithms, ETs introduce randomness by selecting split points at random for each feature, rather than searching for the optimal split. ETs are computationally efficient, capable of handling large datasets, and often outperform other tree-based models in both classification and regression tasks by aggregating the predictions of multiple diverse trees [17].

3) Decision Tree

DT models create a tree-like structure where the internal nodes represent feature-based decisions, and the leaf nodes denote class labels. While DTs are computationally efficient and interpretable, they are prone to overfitting, particularly when the tree depth is not constrained.

4) K-Nearest Neighbors

KNN is a non-parametric algorithm that predicts class labels based on the majority class of the k-nearest data points in the feature space. Although it does not require a training phase, its performance depends on the choice of k, the distance metric (e.g., Euclidean or Manhattan), and feature scaling [18]. Despite these dependencies, KNN is widely used in text classification, image processing, and recommendation systems [19-21].

5) Support Vector Machine

SVMs are supervised learning models that aim to identify the optimal hyperplane, separating different classes in a high-dimensional space. Through kernel functions (e.g., linear, polynomial, or radial basis), SVMs can efficiently handle both linear and non-linear data. The margin maximization principle gives SVM strong generalization capability. However, their performance can be sensitive to the kernel choice and regularization parameters [22].

6) Multi-Layer Perceptron

MLP is a type of feedforward artificial neural network consisted of an input layer, one or more hidden layers, and an output layer. Each layer comprises interconnected neurons, and non-linear activation functions are applied to introduce complexity to the learned representations. MLPs are trained using backpropagation, allowing them to approximate complex non-linear mappings between the inputs and outputs [23].

D. Hyperparameter Optimization

The robustness of ML models depends highly on careful hyperparameter tuning and well-considered architectural choices. In this study, Bayesian optimization was employed for hyperparameter tuning, as it is proven to effectively explore complex parameter spaces and accelerates the convergence towards optimal configurations [24]. This approach used Gaussian process-based surrogate modeling to guide the search, balancing he exploration and exploitation of high-performing settings.

Table I provides an overview of the key hyperparameters and architectural configurations used for each classifier. For the LGBM, the boosting type was set to 'gbdt' with a maximum tree depth of 10, 100 estimators, a minimum of 100 samples per child, and a subsample ratio of 0.7. The ET classifier was configured with a maximum depth of 16 and 184 estimators. The DT used the 'gini' criterion, a maximum depth of 10, and required at least six samples to split an internal node. For KNN, the Manhattan distance metric was used with four neighbours and distance-based weighting. The SVM employed an RBF kernel, with C = 635.23 and gamma set to 'scale'. The MLP architecture included four hidden layers with 148 neurons each, ReLU activation, the Adam optimizer, an adaptive learning rate, and an alpha value of 0.0002816.

Given the very high performance scores observed—particularly with the SVM and MLP models—potential overfitting was a valid concern. To mitigate this, several measures were implemented. First, all hyperparameters were optimized using cross-validation within the training data, ensuring that performance metrics reflect generalization rather than memorization. Second, regularization parameters were carefully tuned to control model complexity. Third, Chi-square feature selection and the use of ensemble methods reduced the influence of irrelevant attributes and enhanced model robustness. Finally, all reported results were obtained on a separate hold-out test set that was never used during training or hyperparameter tuning, thereby providing an unbiased evaluation of model performance.

TABLE I. BAYESIAN-BASED HYPERPARAMETER TUNING FOR ML CLASSIFIERS

Classifier	Parameters
LGBM	boosting_type='gbdt', max_depth=10, *n*_estimators: 100, min_child_samples: 100, num_leaves: 127, subsample: 0.7
ET	max_depth: 16, min_samples_split: 2, *n*_estimators: 184
DT	criterion: 'gini', max_depth: 10, min_samples_leaf: 1, min_samples_split: 6, splitter: 'best'
KNN	metric: 'manhattan', n_neighbors: 4, weights: 'distance'
SVM	C: 635.2340, gamma: 'scale', kernel: 'rbf'
MLP	activation: 'relu', alpha: 0.0002816, learning_rate: 'adaptive', *n*_layers: 4, *n*_neurons: 148, solver: 'adam'

E. Explainable Artificial Intelligence

XAI methods are utilized to support technical experts and non-expert audiences to understand the AI decisions. Two concepts that are central to XAI are: interpretability, which refers to rules that govern AI decisions that can be understood by a human, and explainability, which refers to interfaces designed to make these rational processes accessible. In healthcare, interpretability is needed for clinical adoption, risk reduction, ethical compliance, informed consent, and patient engagement. Clinicians must be able to assess the pace of AI model reasoning with the existing medical research base and apply that reasoning in practice, knowing that it will be reliable and free of bias.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Dataset and Evaluation Metrics

This study utilized the ASD Toddler dataset [25], available on Kaggle [26], which comprised 1,054 cases and 17 categorical variables. Of these, 10 behavioral features were obtained from the Q-CHAT-10 screening tool, while the remaining variables captured demographic features. Among the total cases, 735 were diagnosed with ASD, whereas 319 were labeled as non-ASD.

Given the imbalance between the two classes, the SVM-SMOTE technique was applied during preprocessing to oversample the minority class in a boundary-aware manner (see above). As a result, the dataset was expanded to 1,456 instances with a more balanced distribution of ASD and non-ASD cases. For model development, this balanced dataset was randomly partitioned into training, validation, and testing subsets using a stratified split to preserve class distribution. Specifically, 70% of the data was used for training, 10% for validation, and 20% (292 cases: 148 ASD and 144 non-ASD) for final testing. The training set was employed for model fitting, the validation set for hyperparameter tuning and model selection, while the independent test set was held out and used exclusively for reporting the final performance metrics presented in this study.

A comprehensive collection of evaluation metrics was used for measuring the performance of the ML classifiers for ASD screening:

- Accuracy: Measures the overall correctness of the classifications across both positive and negative cases.
- Precision: Used to evaluate the reliability of positive ASD predictions. This is very critical in clinical screening because the major consideration is to avoid the false positives.
- Recall (or Sensitivity): Assesses the model's ability to identify the true cases, keeping the false negatives low. This is a very relevant consideration for early intervention.
- F1 score: To harmonize precision and recall, F1 score was utilized, providing a balanced measure through their geometric mean. This metric proved particularly valuable given the class imbalance in the dataset, where ASD and non-ASD cases were not equally represented.
- Log Loss (logistic loss or cross-entropy loss): Measures the accuracy of the model's probabilistic predictions. Unlike simple accuracy, log loss penalizes both incorrect and overly confident predictions. Lower values indicate more reliable and well-calibrated predictions, which is especially important for early ASD detection.

B. Experimental Results

To evaluate the impact of Chi-square feature selection on the early ASD detection, six classifiers were tested: LGBM, ET, DT, KNN, SVM, and MLP. Each model was estimated on the original feature set and on a reduced set determined by the Chi-square technique, with all key metrics (Table II).

TABLE II. PERFORMANCE COMPARISON OF ML CLASSIFIERS USING ALL FEATURES VERSUS CHI-SQUARE SELECTED FEATURES

Classifier	Features selected	Accuracy	Precision	Recall	F1-Score	Log Loss
LGBM	ALL	0.976	0.977	0.976	0.976	0.072
	Chi^2	0.980	0.980	0.980	0.979	0.066
ET	ALL	0.986	0.986	0.986	0.986	0.063
	Chi2	0.990	0.990	0.990	0.990	0.053
DT	ALL	0.938	0.939	0.938	0.938	2.223
	Chi^2	0.969	0.969	0.969	0.969	1.113
KNN	ALL	0.959	0.960	0.959	0.959	0.310
	Chi^2	0.973	0.973	0.973	0.973	0.418
SVM	ALL	1.000	1.000	1.000	1.000	0.011
	Chi^2	1.000	1.000	1.000	1.000	0.008
MLP	ALL	0.993	0.993	0.993	0.993	0.012
	Chi^2	0.997	0.997	0.997	0.997	0.006

Across all classifiers, incorporating Chi-square feature selection led to consistent improvements. For example, ET improved from 98.6% to 99% in accuracy, precision, recall, and F1-score, while its log loss reduced from 0.063 to 0.053, reflecting enhanced reliability in probabilistic predictions. LGBM similarly improved its accuracy and precision from 97.6% to 98%, with a drop in log loss from 0.072 to 0.066. DT showed marked gains, with accuracy rising from 93.8% to 96.9% and a substantial reduction in log loss from 2.223 to 1.113, demonstrating that effective feature selection is critical for mitigating overfitting and boosting generalizability. KNN benefited from the increased accuracy and recall (from 95.9% to 97.3%), while MLP reached an accuracy of 99.7%, accompanied by a lower log loss. The SVM classifier was the best performing model achieving 100% in accuracy, precision, recall, and F1-score both before and after feature selection while displaying further improvement in log loss (from 0.011 to 0.008) post-selection.

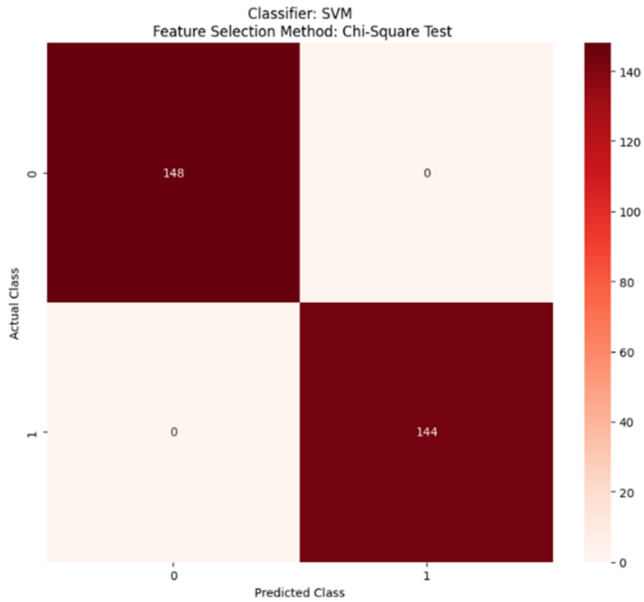

Fig. 1. Confusion matrix of SVM with chi-square feature selection.

To further validate the results, the confusion matrix of the best-performing model (SVM with Chi-square feature selection) was developed, as presented in Figure 1. The classifier achieved accurate separation between ASD and non-ASD cases on the independent test set. Out of 148 ASD instances and 144 non-ASD instances, all were classified correctly, yielding zero false positives and zero false negatives. This indicated that the model attained 100% sensitivity and 100% specificity, confirming its suitability for reliable clinical screening applications.

Except these, a SHAP analysis was performed on the best-performing model to ensure the clinical trust. Both beeswarm and bar plots (Figures 2 and 3) confirmed that the most influential predictors for the ASD classification were behavioral features, such as joint attention ("Does your child follow where you're looking?"), use of simple gestures, and empathy-related behaviors. These features not only exhibited the highest SHAP values, but also corresponded to established clinical markers. Conversely, features, like sex and jaundice, revealed a minimal impact, further supporting the importance of robust feature selection.

By focusing on the most relevant features, the proposed approach enhanced the classification performance, while it simplified the screening process, reducing both the data collection time and model complexity. The alignment between the statistical significance and clinical interpretability further supports the use of this integrated ML framework as a practical tool for the early ASD risk assessment.

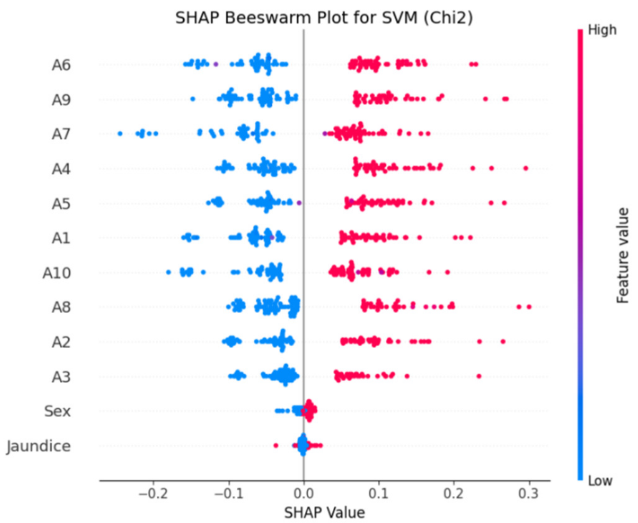

Fig. 2. SHAP beeswarm plot illustrating feature contributions for SVM classifier with Chi-square feature selection.

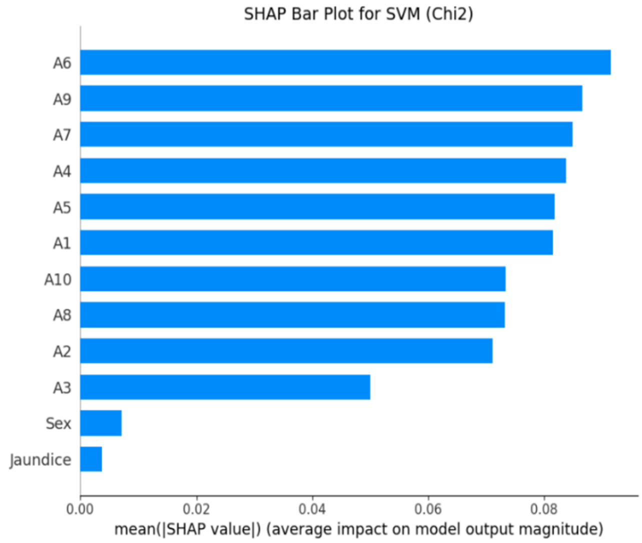

Fig. 3. SHAP bar plot displaying the mean absolute SHAP values of the features for the SVM classifier with Chi-square feature selection.

A direct comparison between some recent ASD detection models is presented in Table III. The proposed SVM-based framework achieved a perfect accuracy of 100%, outperforming other leading approaches, such as LR (97.15%,[9]), SVM (97.82%,[6]), and RF (93.69%,[13]).

Despite these strengths, the study has certain limitations. The results were based on a single, publicly available dataset, which may not capture the full diversity of ASD presentations across different populations and cultural contexts. This means that real-world deployment would require external validation on larger and more heterogeneous cohorts, as well as the potential integration of additional modalities, such as genetic or neuroimaging data, to further enhance the diagnostic accuracy.

TABLE III. COMPARISON AGAINST STATE-OF-THE-ART MODELS FOR ASD DETECTION

Study	Publication Year	Model	Accuracy (%)
[9]	2021	LR	97.15
[6]	2022	SVM	97.82
[13]	2023	RF	93.69
Proposed model	2025	SVM	100

V. CONCLUSION

In this study, an integrated Machine Learning (ML) framework was presented, aiming to identify the Autism Spectrum Disorder (ASD) in toddlers. It was distinguished by the systematic incorporation of the following four components:

- The Chi-square test to extract the most important behavioral indicators from a statistical and clinical standpoint,
- Support Vector Machine - Synthetic Minority Over-sampling Technique (SVM-SMOTE) to balance the dataset
- Bayesian hyperparameter optimization for improving the model's accuracy

- SHapley Additive (SHAP) explainability for effective and clinically interpretable decision-making.

The experimental findings indicated that combining SVM-SMOTE alongside Chi-square feature selection enhanced the prediction capabilities. The SVM classifier, in particular, achieved perfect scores across all evaluation metrics, while other models, including ET, DT, LGBM, KNN, and MLP also showed marked improvement.

Future research should prioritize the external validation on larger and more diverse cohorts, as well as the incorporation of multimodal data sources, to further strengthen the accuracy, generalizability, and clinical relevance of early ASD detection strategies.

ACKNOWLEDGMENT

This research was supported by Jadara University, Jordan. The author gratefully acknowledges the University's financial support, which made this study possible.

REFERENCES

[1] G. Dawson *et al.*, "Randomized, Controlled Trial of an Intervention for Toddlers With Autism: The Early Start Denver Model," *Pediatrics*, vol. 125, no. 1, pp. e17–e23, Jan. 2010, https://doi.org/10.1542/peds.2009-0958.

[2] L. Zwaigenbaum *et al.*, "Early Intervention for Children With Autism Spectrum Disorder Under 3 Years of Age: Recommendations for Practice and Research," *Pediatrics*, vol. 136, no. Supplement_1, pp. S60–S81, Oct. 2015, https://doi.org/10.1542/peds.2014-3667E.

[3] F. J. Rajam and B. R. K. Swakkin, "MAPLE: A Novel Processing Technique for Adult Autism Prediction," *Engineering, Technology & Applied Science Research*, vol. 15, no. 3, pp. 23901–23906, June 2025, https://doi.org/10.48084/etasr.10864.

[4] Q. Y. Shambour, M. M. Al-Zyoud, A. H. Hussein, and Q. M. Kharma, "A doctor recommender system based on collaborative and content filtering," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 1, pp. 884–893, Feb. 2023, https://doi.org/10.11591/ijece.v13i1.pp884-893.

[5] M. Bala, M. H. Ali, M. S. Satu, K. F. Hasan, and M. A. Moni, "Efficient Machine Learning Models for Early Stage Detection of Autism Spectrum Disorder," *Algorithms*, vol. 15, no. 5, May 2022, Art. no. 166, https://doi.org/10.3390/a15050166.

[6] D. Theng and K. K. Bhoyar, "Feature selection techniques for machine learning: a survey of more than two decades of research," *Knowledge and Information Systems*, vol. 66, no. 3, pp. 1575–1637, 2024, https://doi.org/10.1007/s10115-023-02010-5.

[7] Z. Sadeghi *et al.*, "A review of Explainable Artificial Intelligence in healthcare," *Computers and Electrical Engineering*, vol. 118, Aug. 2024, Art. no. 109370, https://doi.org/10.1016/j.compeleceng.2024.109370.

[8] 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, July 2021, Art. no. 386, https://doi.org/10.1007/s42979-021-00776-5.

[9] M. R. Alteneiji, L. M. Alqaydi, and M. U. Tariq, "Autism Spectrum Disorder Diagnosis using Optimal Machine Learning Methods," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 9, 2020, https://doi.org/10.14569/IJACSA.2020.0110929.

[10] S. M. Mahedy Hasan, M. P. Uddin, M. A. Mamun, M. I. Sharif, A. Ulhaq, and G. Krishnamoorthy, "A Machine Learning Framework for Early-Stage Detection of Autism Spectrum Disorders," *IEEE Access*, vol. 11, pp. 15038–15057, 2023, https://doi.org/10.1109/ACCESS.2022.3232490.

[11] M. 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, May 2023, Art. no. 92, https://doi.org/10.3390/computers12050092.

[12] J. Talukdar, D. K. Gogoi, and T. P. Singh, "A comparative assessment of most widely used machine learning classifiers for analysing and classifying autism spectrum disorder in toddlers and adolescents," *Healthcare Analytics*, vol. 3, Nov. 2023, Art. no. 100178, https://doi.org/10.1016/j.health.2023.100178.

[13] A. M. Almars, I. Gad, and E.-S. Atlam, "Unlocking autistic emotions: developing an interpretable IoT-based EfficientNet model for emotion recognition in children with autism," *Neural Computing and Applications*, vol. 37, no. 21, pp. 17129–17148, July 2025, https://doi.org/10.1007/s00521-025-11290-7.

[14] E.-S. Atlam *et al.*, "Explainable artificial intelligence systems for predicting mental health problems in autistics," *Alexandria Engineering Journal*, vol. 117, pp. 376–390, Apr. 2025, https://doi.org/10.1016/j.aej.2024.12.120.

[15] H. M. Nguyen, E. W. Cooper, and K. Kamei, "Borderline over-sampling for imbalanced data classification," *International Journal of Knowledge Engineering and Soft Data Paradigms*, vol. 3, no. 1, pp. 4–21, Apr. 2011, https://doi.org/10.1504/IJKESDP.2011.039875.

[16] G. Ke *et al.*, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," in *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA, 2017.

[17] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63, no. 1, pp. 3–42, Apr. 2006, https://doi.org/10.1007/s10994-006-6226-1.

[18] L. E. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, Feb. 2009, Art. no. 1883, https://doi.org/10.4249/scholarpedia.1883.

[19] Q. Y. Shambour, N. M. Turab, and O. Y. Adwan, "An Effective e-Commerce Recommender System Based on Trust and Semantic Information," *Cybernetics and Information Technologies*, vol. 21, no. 1, pp. 103–118, Mar. 2021, https://doi.org/10.2478/cait-2021-0008.

[20] J. Lu, Q. Shambour, and G. Zhang, "Recommendation technique-based government-to-business personalized e-services," in *Conference of the North American Fuzzy Information Processing Society - NAFIPS*, Cincinnati, OH, USA, June 2009, https://doi.org/10.1109/NAFIPS.2009.5156456.

[21] Q. Shambour and J. Lu, "A Framework of Hybrid Recommendation System for Government-to-Business Personalized E-Services," in *2010 Seventh International Conference on Information Technology: New Generations*, Las Vegas, NV, USA, Apr. 2010, https://doi.org/10.1109/ITNG.2010.114.

[22] D. A. Pisner and D. M. Schnyer, "Support vector machine," in *Machine Learning: Methods and Applications to Brain Disorders*, Amsterdam, Netherlands: Elsevier, 2020, pp. 101–121.

[23] R. Kruse, S. Mostaghim, C. Borgelt, C. Braune, and M. Steinbrecher, "Multi-layer Perceptrons," in *Computational Intelligence: A Methodological Introduction*, R. Kruse, S. Mostaghim, C. Borgelt, C. Braune, and M. Steinbrecher, Eds. Cham, Switzerland: Springer International Publishing, 2022, pp. 53–124.

[24] J. Wu, X.-Y. Chen, H. Zhang, L.-D. Xiong, H. Lei, and S.-H. Deng, "Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimizationb," *Journal of Electronic Science and Technology*, vol. 17, no. 1, pp. 26–40, Mar. 2019, https://doi.org/10.11989/JEST.1674-862X.80904120.

[25] F. Thabtah, F. Kamalov, and K. Rajab, "A new computational intelligence approach to detect autistic features for autism screening," *International Journal of Medical Informatics*, vol. 117, pp. 112–124, Sept. 2018, https://doi.org/10.1016/j.ijmedinf.2018.06.009.

[26] F. Thabtah, "Autism screening data for toddlers." Kaggle, 2018, https://www.kaggle.com/datasets/fabdelja/autism-screening-for-toddlers/data.