



A Feature-Optimized, Imbalance-Aware Ensemble Classifier for Identifying Autism in Toddlers, Children, and Adults

Vittalraju Chetan Kumar^{*}, Dadadahalli Ramu Umesh[†]

Department of Computer Science and Engineering, P.E.S. College of Engineering, Affiliated to Visvesvaraya Technological University, Mandya 571401, India

Corresponding Author Email: vcetanpdh@gmail.com

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/mmep.120830>

ABSTRACT

Received: 4 June 2025

Revised: 3 August 2025

Accepted: 8 August 2025

Available online: 31 August 2025

Keywords:

Autism Spectrum Disorder (ASD), class imbalance, ensemble learning, Extreme Gradient Boosting (XGB), feature selection, machine learning, Multilayer Perceptron (MLP)

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition that impacts social behavior, communication, and cognitive functions. Accurate early detection is essential for timely intervention and improved developmental outcomes. However, ASD identification remains challenging due to feature relevance, class imbalance, noisy features, misclassification errors, and variability in symptom presentation across age groups. To address these issues, this paper proposes a novel ensemble classification framework, Feature-Optimized Imbalanced Data Ensemble (FOIDE). FOIDE incorporates a cost-sensitive variant of Extreme Gradient Boosting (XGB) to extract weighted features by penalizing misclassification of minority classes. These features are further refined through a Multilayer Perceptron (MLP) to capture complex non-linear patterns. Additionally, a probabilistic nested cross-validation mechanism is introduced to rank and select the most discriminative features, improving generalization across diverse age groups, including toddlers, children, and adults. FOIDE is evaluated on four benchmark ASD screening datasets representing toddlers, toddler-merged, children, and adults. Experimental results demonstrate FOIDE's strong performance, achieving 99.98% accuracy on the toddler dataset, 99.12% on the imbalanced toddler dataset, 99.97% on the toddler-merged dataset, 99.78% on the children's dataset, and 99.81% on the adult dataset. These results outperform existing methods, highlighting FOIDE's effectiveness in handling class imbalance and enhancing age-generalized ASD prediction.

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is a lifelong neurological and developmental condition that affects how individuals communicate, behave, and interact socially. The disorder manifests in early childhood, with symptoms that vary significantly in severity and expression across different individuals [1]. Early diagnosis is essential, as timely intervention can lead to substantial improvements in cognitive, emotional, and social outcomes. However, conventional diagnostic approaches are often manual, reliant on expert observation, and time-consuming, making early identification difficult, especially in resource-constrained environments [2]. Furthermore, ASD behaviors evolve with age, and many individuals are either diagnosed late or misdiagnosed entirely, which limits access to appropriate support mechanisms [3].

Modern machine learning (ML) techniques have shown considerable potential in addressing this gap by automating the analysis of high-dimensional behavioral and clinical data [4, 5]. Through data-driven models, patterns indicative of ASD can be learned from assessment datasets, enabling earlier and more consistent diagnoses. However, despite this promise, existing ML-based ASD detection frameworks face several significant limitations that prevent their widespread and

reliable application in real-world scenarios.

One major challenge lies in the lack of generalizability across age groups. Many prior models are tailored to specific cohorts, such as children or adults, and do not generalize well across diverse developmental stages, including toddlers, adolescents, and adults [6, 7]. Behavioral traits and cognitive indicators associated with ASD can differ significantly with age, and models trained on a narrow subset of the population tend to perform poorly when applied to other groups [8].

Another critical limitation is the presence of class imbalance in ASD datasets. Typically, datasets contain far fewer positive ASD samples than non-ASD controls, causing learning algorithms to be biased toward the majority class. This leads to high false-negative rates, where individuals with ASD are wrongly classified as non-ASD, which is particularly problematic in clinical settings [9, 10]. Most conventional classifiers, including logistic regression (LR), support vector machines (SVMs), and decision trees (DTs), do not incorporate mechanisms to address this imbalance adequately, resulting in misleadingly high accuracy but poor sensitivity [11].

Additionally, feature selection and interpretability remain underdeveloped in most existing frameworks. ASD diagnosis often involves a wide array of behavioral, sensory, and

cognitive variables. Selecting the most relevant features not only improves model performance but also aids clinicians in understanding which factors contribute most significantly to diagnostic decisions. However, many models use static or filter-based feature selection techniques, which may not account for the interdependencies between features or their relevance across age groups [12, 13].

To overcome these challenges, several ensemble learning approaches have been proposed in the literature. Ensemble models such as Random Forests (RF), AdaBoost (AB), and Extreme Gradient Boosting (XGB) aggregate predictions from multiple base learners to improve overall accuracy and robustness [14, 15]. These methods have been shown to outperform individual classifiers in various ASD prediction tasks. For instance, studies using XGB and RF have demonstrated strong performance in binary classification tasks involving children or adult datasets [16]. In other cases, combinations of classifiers like LR, SVMs, and naive Bayes (NBs) were evaluated on benchmark datasets such as those from the University of California, Irvine (UCI), and Kaggle, achieving classification accuracies ranging from 94% to 99% for isolated age cohorts [17, 18].

Optimization techniques such as Genetic Algorithms, Particle Swarm Optimization (PSO), and Cuckoo Search have also been employed to reduce feature dimensionality and enhance model accuracy [19, 20]. Additionally, some studies have incorporated feature transformation methods like Quantile Transform, Power Transform, and normalization to improve data consistency across samples [21]. These methods, while effective to an extent, do not comprehensively address the problems of class imbalance, interpretability, and age group diversity.

Some recent work has also considered the use of ensemble models in clinical and therapeutic settings. For example, multi-class ensemble classifiers have been used to monitor therapy effectiveness and evaluate short-term factors contributing to ASD progression [22]. These studies emphasize the importance of integrating behavioral assessments with intelligent diagnostic tools. However, such models often lack a robust mechanism for feature selection and ranking, particularly in the presence of imbalanced data.

A few studies have explored federated learning approaches for ASD detection, where data collected from multiple clients (e.g., hospitals or edge devices) are used to train local models without centralizing sensitive information [23]. These frameworks demonstrate the feasibility of privacy-preserving ASD diagnosis, especially in mobile and remote healthcare contexts. However, the core classification models in these studies still suffer from the same limitations, namely, insufficient handling of imbalanced datasets and a lack of cross-age adaptability.

Motivated by these challenges, this paper proposes a comprehensive machine learning framework termed Feature-Optimized Imbalanced Data Ensemble (FOIDE), specifically designed for early and accurate detection of ASD across multiple age groups. Unlike previous approaches, FOIDE combines cost-sensitive learning, deep neural ensemble modeling, and nested cross-validation-based feature selection into a unified pipeline.

The first core component of FOIDE is a weighted XGBoost classifier, in which the loss function is modified to impose greater penalties on false negatives than on false positives.

This cost-sensitive approach is particularly valuable in ASD detection, where failing to identify a true ASD case can have serious consequences [24]. By embedding imbalance awareness directly into the learning objective, the model becomes more sensitive to minority class instances, improving recall without compromising overall accuracy.

The second component of the FOIDE framework is a Multilayer Perceptron (MLP) that receives the output of the weighted XGBoost feature transformation. This enables the model to learn complex, nonlinear relationships among behavioral indicators, which may not be adequately captured by tree-based methods alone. The MLP acts as a high-level representation learner that refines the decision boundaries informed by XGB [25].

The third innovation is a nested cross-validation (CV) mechanism that is used for both feature selection and model generalization. In this approach, the dataset is split into multiple folds, and within each fold, an inner-loop validation is used to rank features based on their predictive importance. This hierarchical evaluation ensures that the selected features are robust and not overfitted to any specific partition of the data. Moreover, features that consistently appear across multiple outer folds are prioritized, improving both model interpretability and generalization across age groups [26, 27].

Compared to prior work, FOIDE offers several methodological advancements. Unlike reference [16], which used a two-stage ensemble model [28] without feature ranking, FOIDE integrates feature selection, imbalance handling, and ensemble learning within a single, adaptive framework. Unlike references [20, 21], which focused on optimization-based feature reduction but lacked age group adaptability, FOIDE specifically targets cross-age generalization using stratified datasets. It also overcomes the limitations of reference [24], where misclassification cost tuning was absent, and ensemble learners were evaluated in isolation.

The key contributions of this research are summarized below:

- Introduction of a cost-sensitive XGBoost loss function tailored to improve sensitivity for ASD detection in imbalanced data environments.
- Development of a hybrid XGB–MLP ensemble model that captures both linear and nonlinear patterns in high-dimensional behavioral data.
- Implementation of a nested cross-validation strategy for robust feature ranking and selection, improving generalizability across toddlers, children, and adults.
- Comprehensive evaluation across multiple age-stratified ASD datasets, demonstrating superior accuracy, reduced false negatives, and improved feature interpretability compared to existing baselines.

This work addresses key gaps in the field by providing an ensemble-based, interpretable, and imbalance-resilient framework for ASD detection. By aligning learning objectives with clinical priorities and ensuring that models generalize across age groups, FOIDE contributes both technically and practically to the evolving field of computational autism research.

The rest of the paper is organized as follows: Section 2 details the proposed FOIDE methodology, Section 3 presents experimental results and comparative analysis, and Section 4 concludes with insights, limitations, and future research directions.

2. PROPOSED METHODOLOGY

This section introduces a novel framework named FOIDE for effectively detecting ASD among young and adults, as shown in Figure 1. In designing an effective framework, the work first introduces a novel weight optimization mechanism for the XGB algorithm. Then, the extracted feature is trained using MLP. Finally, XGB-MLP employs an enhanced cross-validation mechanism to effectively select and rank the features contributing to ASD and build an effective ASD predictive classifier.

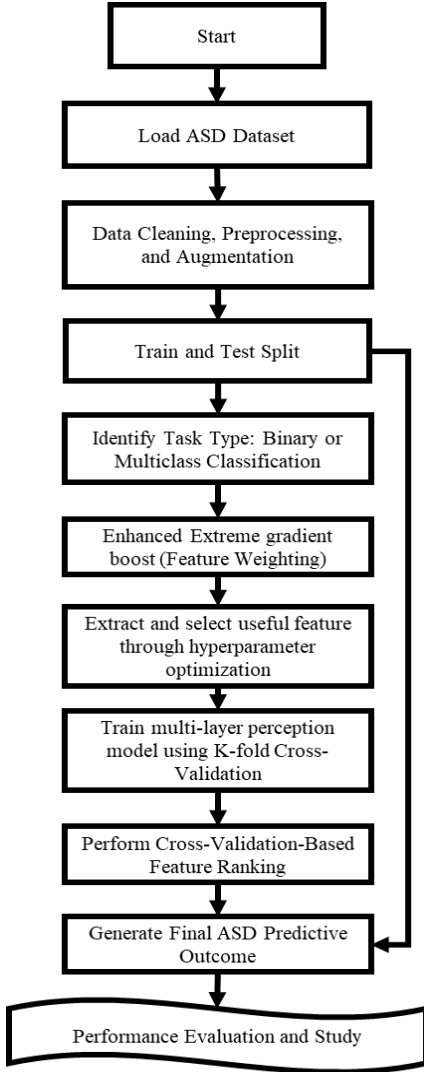


Figure 1. Framework for detecting ASD using feature-optimized imbalanced data-aware ensemble classifier

2.1 Architecture for detection of autism

As shown in Figure 1, the general ASD dataset is defined in Eq. (1):

$$X = \{(a_1, b_1), (a_2, b_2), \dots, (a_m, b_m)\} \quad (1)$$

where, the parameter $j = 1, 2, 3, \dots, m$, depicts the total ASD samples present within the dataset, parameter a_j defines an n -dimensional vector defining the respective features of j , and parameter $b_j \in \{0, 1\}$ represents j^{th} row output. During preprocessing, any rows containing missing data are removed

before training the model. To increase the dataset size and balance the distribution, data augmentation is then applied. Subsequently, the model determines whether the classification task is binary or multi-label based on the dataset labels. The ASD dataset is composed of various assessments featuring multi-dimensional features. In analyzing and designing ASD classifier \hat{G} , for classifying the actual value of original data G , considering a specific label like the presence of ASD or not is expressed in Eq. (2):

$$g : A \rightarrow B \quad (2)$$

where, A represents the observer feature j and B defines its outcome in j^{th} row. In this work, an ASD predictive model is constructed by minimizing a carefully designed objective function, utilizing a hybrid XGBoost-MLP classifier. The model incorporates an enhanced cross-validation mechanism to ensure robust feature selection and optimal generalization performance.

2.2 Feature extraction using cost-sensitive Extreme Gradient Boosting (XGB)

XGB algorithm is a well-known gradient tree boosting methodology used by various standard models for solving various classification problems [24]. To handle class imbalance and capture initial feature importance, we employ a cost-sensitive version of the XGB algorithm. XGB is a robust ensemble technique based on gradient tree boosting, which constructs a series of decision trees to approximate complex functions. The cumulative output for a given input sample Y_j is expressed as:

$$\hat{Z}_j = G(Y_j) = \sum_{l=1}^L g_l(Y_j), \quad g_l \in \alpha \quad (3)$$

where, \hat{Z}_j defines the classification outcomes of a multi-label classification model with a certain dimension, l^{th} dimension describes the probability that it will be classified as belonging to the l^{th} class, and α defines a set of decision trees as described below:

$$\alpha = \{g(y) = x_{t(y)}\} \quad (4)$$

where, every tree $g(y)$ agree with respect to leaf weight vector x and structure parameter t . The objective of the XGB classification model is to minimize the regularized loss parameter:

$$M(G) = \sum_j \mathcal{L}(\hat{z}_j, z_j) + \sum_l \beta(g_l) \quad (5)$$

where,

$$\beta(g_l) = \delta U + \mu \|x\|^2 \quad (6)$$

The first parameter $\mathcal{L}(\hat{z}_j, z_j)$ in Eq. (5), defines the prediction loss function between actual and classified outcomes. The second parameter $\beta(g_l)$ in Eq. (5) depicts the penalizing term; U depicts the leaf size within a tree, δ and μ depict the controlling parameters used for controlling computational complexity. The conventional predictive model

using the binary cross-entropy model is obtained below:

$$\mathcal{L}(\hat{z}_j, z_j) = - \sum_{j=1}^o [z_j \log(\hat{z}_j) + (1 - z_j) \log(1 - \hat{z}_j)] \quad (7)$$

However, conventional loss functions do not account for class imbalance, which is critical in ASD datasets. Thus, we introduce a weighted cost-sensitive negative log probabilistic loss function to penalize misclassification more appropriately.

In addressing these problems in this section, a modified feature extraction strategy for XGB is presented in Eq. (8):

$$\mathcal{L}_b(\hat{z}_j, z_j) = - \sum_{j=1}^o [b \cdot z_j \log(\hat{z}_j) + (1 - z_j) \log(1 - \hat{z}_j)] \quad (8)$$

where, b is a feature weight parameter that increases the penalty for misclassifying minority class samples (ASD-positive cases). When $b > 1$, false negatives are penalized more heavily, improving sensitivity. The result of this phase is a feature-weighted output vector encoding both prediction strength and feature relevance, which serves as input for the next stage.

In Section 2.3, acknowledge that Eq. (8) is related to cost-sensitive learning, but clarify that your enhancement lies in joint optimization with deep learning and ranked cross-validation, not cost-sensitive classification alone.

2.3 Nonlinear feature learning using MLP

The feature outputs generated from the modified XGB model are passed into an MLP [29] to model complex, nonlinear patterns within the data. MLP further refines the representation learned by the XGB, allowing richer abstractions across multiple layers. The MLP prediction for a given input sample with F features is expressed as:

$$\hat{z} = \mu_o \left(\sum_{i=1}^F w_{jp}^o \cdot \mu_H \left(\sum_{i=0}^n w_{ij}^H \cdot x_i \right) \right) \quad (9)$$

where, variables w_{ij}^H and w_{jp}^o denote the weights associated with the hidden and output layers, respectively. The output-layer activation function is denoted as μ_o , while the hidden-layer activation function is denoted as μ_H . The variable x_i represents the i^{th} input feature. This architecture captures nonlinear dependencies that are difficult for tree-based models to represent, particularly in multi-age ASD behaviors spanning toddlers to adults.

2.4 Enhanced cross-validation-based feature selection and ranking

To further improve model robustness and identify the most influential features contributing to ASD classification, we introduce a two-phase nested CV mechanism for feature ranking and selection.

Phase 1: Nested Cross-Validation for Feature Evaluation

The ASD dataset \mathcal{E} is partitioned into K outer folds. For each fold k , one part E^k is used for testing, while the remaining \mathcal{E}^{-k} is used for training.

Within \mathcal{E}^{-k} , a second-level (inner) partitioning into H folds

is performed. Each inner fold h generates training and validation splits \mathcal{E}^{-kh} and \mathcal{E}^{kh} , respectively. For each hyperparameter configuration σ_l considering grid l , a candidate model is trained and evaluated:

$$\hat{g}_{\sigma_l} = \hat{g}(b_j, \hat{g}(\mathcal{E}^{-kh}; \sigma_l)) \quad (10)$$

The validation error is computed as:

$$\mathbb{E}_{\sigma_n} = \sum_{j \in \mathcal{E}^{-kh}} P(b_j, \hat{g}(\mathcal{E}^{-kh}; \sigma_l)) \quad (11)$$

The average cross-validation error experienced across inner folds for different grid l is:

$$CV(\hat{g}; \sigma_l) = \frac{1}{M_h} \sum_{h=1}^H \sum_{j \in \mathcal{E}^{-kh}} P(b_j, \hat{g}(\mathcal{E}^{-kh}; \sigma_l)) \quad (12)$$

The cross-validation error is further optimized in a repeated manner for different grid l , considering S -times, considering rows M_h to ensure stability:

$$CV_s(\hat{g}; \sigma_l) = \frac{1}{M_h S} \sum_{s=1}^S \sum_{h=1}^H \sum_{j \in \mathcal{E}^{-kh}} P(b_j, \hat{g}(\mathcal{E}^{-kh}; \sigma_l)) \quad (13)$$

The optimal ASD predictive model is selected by minimizing error for different l as:

$$\hat{\sigma}_n = \arg \min_{\sigma \in \{\sigma_1, \sigma_l\}} CV_s(\hat{g}; \sigma_l) \quad (14)$$

A binary ranking vector $r(a)$ is created to indicate whether each feature n_j contributes significantly to the final ASD predictive model:

$$r(a) = \begin{cases} 0 & \text{if } n_j \text{ is not chosen} \\ 1 & \text{if } n_j \text{ is chosen as} \\ & \text{the final predictive model } j = 1, 2, 3, \dots, n \end{cases} \quad (15)$$

The initial feature set selected contributing to ASD is obtained as:

$$F_s = \{r(n_1), r(n_1), \dots, r(n_n)\} \quad (16)$$

Then, considering K -folds, the feature with the highest rank is selected and chosen as a very important feature contributing to ASD, as follows:

$$F_{s_k} = \{r(n_1), r(n_1), \dots, r(n_n)\} \quad (17)$$

Phase 2: Final Feature Filtering

In this phase, only the features with the highest rank are used in designing the final predictive model, as shown in Eq. (18).

$$F_{s_{final}} = \{f_s(p_1), f_s(n_2), \dots, f_s(n_n)\} \quad (18)$$

where, $f_s(\cdot)$ defines whether the selected n^{th} feature should be selected or not, as:

$$F_s(a) = \begin{cases} 0 & \text{if } q_j \text{ is chosen lesser than } \frac{K}{2} \text{ times, } j = 1, 2, 3, \dots, n \\ 1 & \text{if } q_j \geq \frac{K}{2} \text{ times, } j = 1, 2, 3, \dots, n \end{cases} \quad (19)$$

Using Eq. (19) assures that only best feature is chosen for constructing the ASD predictive model using the XGB-MLP ensemble classifier. This two-phase CV ensures that only stable and highly discriminative features are retained, improving model robustness across age-diverse and imbalanced datasets.

2.5 FOIDE pipeline model

To address the detection of ASD across diverse age groups and imbalanced datasets, the FOIDE framework integrates three major components in a unified learning pipeline: cost-sensitive XGBoost, MLP, and enhanced CV. Figure 2 illustrates the FOIDE pipeline.

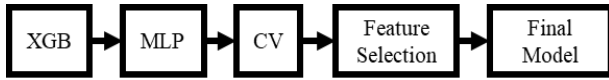


Figure 2. Pipeline model of FOIDE

The step-by-step interaction among the components is as follows:

(1) Feature weight optimization using modified XGB: A cost-sensitive variant of XGBoost is employed to handle data imbalance using a weighted loss function (Eq. (8)), where feature importance is adjusted via the penalty term b . This XGB model extracts initial representations and calculates weighted feature contributions considering false positive against false negative trade-offs.

(2) Deep learning enhancement using MLP: The intermediate feature vectors generated from the weighted XGB trees are passed into an MLP model to capture non-linear, high-order correlations. The MLP acts as a refiner and nonlinear enhancer, ensuring that interactions among features (especially for ASD-specific behaviors) are learned robustly. Eq. (9) describes the MLP transformation applied to XGB feature outputs.

(3) Feature selection and ranking via enhanced cross-validation: A two-phase nested cross-validation strategy is used to evaluate and rank features: The outer folds (K-fold) validate the generalizability of feature selections. Then, Inner folds (H-fold) tunes hyperparameters and tracks feature frequency. Eqs. (10)-(19) provide a systematic way to quantify feature stability and importance across multiple folds and repetitions (S). Features appearing in $\geq K/2$ outer folds are retained for the final predictive model (Eq. (19)).

(4) Final ensemble construction: The final XGB-MLP ensemble is trained on the optimally selected features

$F_{s_{final}}$ using the previously tuned architecture. This ensemble model is designed to improve generalization and robustness across imbalanced, multi-age ASD datasets, combining the interpretability of XGB and the nonlinearity modeling power of MLP.

Unlike traditional cost-sensitive classifiers, FOIDE uniquely integrates weight-adjusted tree boosting with deep learning and nested validation-based feature ranking in a tightly coupled pipeline. This ensures the final model is both interpretable (via XGB) and expressive (via MLP), while also

being resilient to class imbalance and data noise for detecting ASD across different age groups with enhanced accuracy.

3. SIMULATION ANALYSIS AND RESULT

This section studies the performance of the FOIDE classifier over other existing approaches for detecting autism across different age groups such as toddlers, children, and adults. All the datasets are collected for UCI and Kaggle like work presented in references [23, 24]. The proposed ASD predictive and existing ASD predictive are implemented using the Anaconda Python 3 framework. The accuracy, precision, recall, and F1-score are metrics used for validating ASD predictive models using Eqs. (20)-(23), respectively.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (20)$$

$$Precision = \frac{TP}{TP + FP} \quad (21)$$

$$Recall = \frac{TP}{TP + FN} \quad (22)$$

$$F - Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (23)$$

3.1 Dataset details

This study evaluates the performance of the proposed FOIDE classifier using four publicly available ASD screening datasets sourced from Kaggle and the UCI Machine Learning Repository. A summary of these datasets is presented in Table 1.

Table 1. Dataset description

Category	Number of Instances	Number of ASD	Number of Non-ASD	Number of Attributes
Toddler [30]	1050	130	920	17
Toddler-Saudi Arabia [31]	506	341	265	16
Toddler-merged	1556	471	1085	16
Children [32]	292	66	226	20
Adult [33]	704	261	443	20

The Kaggle Toddler dataset [30] contains 1050 instances, with 130 positive ASD cases and 920 non-ASD cases, resulting in a class imbalance ratio of approximately 1:7.1. The dataset includes 17 attributes, encompassing behavioral assessments and demographic features.

The Saudi Arabia Toddler dataset [31], collected via clinical screening, comprises 506 samples, of which 341 are ASD-positive and 165 are non-ASD, yielding an inverse but still imbalanced ratio of approximately 2:1. This dataset includes 16 features, mostly binary responses to ASD-related behavioral indicators.

The toddler dataset was created by merging two prominent publicly available sources from Kaggle [30, 31]. The resulting merged toddler dataset contains 1556 instances, of which 471 are ASD-positive and 1085 are non-ASD, resulting in a class imbalance ratio of approximately 1:2.3. This dataset spans 16 attributes with jaundice factor being eliminated as it is not available in Saudi Arabia Toddler dataset [31], incorporating both behavioral indicators (e.g., A1-A10 responses) and demographic features like age, gender, and caregiver relationship.

The UCI Children dataset [32] includes 292 instances, with 66 ASD-positive and 226 ASD-negative samples, reflecting an imbalance ratio of 1:3.4. This dataset contains 20 attributes, covering social, communication, and motor behaviors relevant to ASD diagnosis.

Finally, the UCI Adult dataset [33] consists of 704 samples, including 261 ASD-positive and 443 non-ASD entries, with an imbalance ratio of approximately 1:1.7. Like the children’s dataset, it contains 20 attributes and is derived from standardized screening tools.

Class imbalance is a critical challenge in these datasets, especially for the toddler and children’s categories, where underrepresentation of ASD-positive cases can lead to biased model training and poor sensitivity. The FOIDE framework addresses this challenge by integrating weighted loss functions, ensemble learning, and enhanced cross-validation to improve robustness and generalization across all age groups and imbalance levels.

3.2 Classifier performance study for toddler and children dataset

The initial evaluation involves the Kaggle dataset, which is notably imbalanced with a 1:7.1 ASD to non-ASD ratio. FOIDE is compared against conventional classifiers such as AdaBoost [23] and SVM-RIPPER [25]. As shown in Figure 3, FOIDE outperforms baseline models in accuracy, precision, recall, and F1-score, particularly excelling in recall, which is crucial in identifying ASD-positive cases under severe class imbalance.

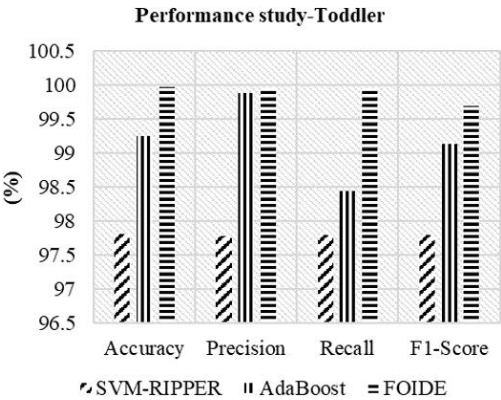


Figure 3. Classification performance for the ASD toddler dataset

Case study 2. In this experiment, the Kaggle toddler and Saudi Arabia toddler datasets are combined to simulate more diverse screening conditions. We compare FOIDE with an ensemble-based classifier combining RF and XGB with meta-feature selection [24]. The results in Figure 4 highlight FOIDE's consistent improvement across all performance metrics. Notably, FOIDE achieves higher recall and F1-score,

indicating better handling of minority class prediction. This is primarily due to the cost-sensitive loss adjustment in the enhanced XGB module and the two-phase feature selection mechanism integrated with cross-validation.

Case study 3. FOIDE is evaluated on the UCI children dataset and benchmarked against logistic regression [18], Cuckoo Search-based Feature Reduction (FR-CS) [19], AdaBoost [23], and Machine Learning Autism Spectrum Disorder (ML-ASD) [27]. The ML-ASD employed adaptive boosting with feature optimization to design the ASD predictive model. As presented in Figure 5, FOIDE demonstrates superior generalization, especially under class imbalance, achieving improved detection of minority ASD cases. This is attributed to its ensemble architecture and feature importance ranking based on repeated nested cross-validation, which reduces overfitting to dominant non-ASD classes.

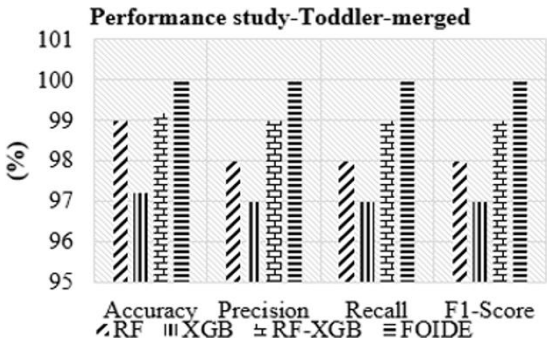


Figure 4. Classification performance for the ASD toddler-merged dataset

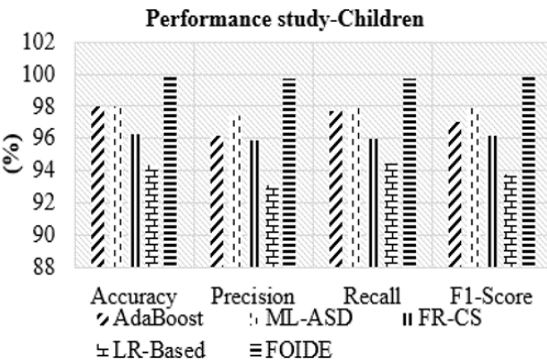


Figure 5. Classification performance for the ASD children dataset

Table 2 presents a comparative evaluation of classification performance across multiple ASD detection approaches on both the original and imbalanced versions of the toddler dataset. The proposed FOIDE model consistently demonstrates superior performance over existing methods [23-26], achieving a peak accuracy of 100% on the balanced dataset and an average accuracy of 99.12% on the imbalanced dataset. These results were obtained using 5-fold nested cross-validation, considering statistical tests conducted with $p < 0.05$, and performance is reported as mean accuracy across multiple runs to account for variability and model stability. A paired t-test between FOIDE and the next-best model [24] confirms that the performance difference is statistically significant ($p < 0.05$).

The FOIDE model’s ability to maintain high performance under imbalanced conditions highlights its robustness and

generalizability, especially compared to baseline models like AdaBoost, RF-XGB, and SVM-based ensembles. While the peak accuracy is 100%, the average accuracy and performance consistency (as indicated by standard deviation) provide a more realistic evaluation, mitigating concerns of overfitting or bias due to class imbalance.

Table 2. Comparative study of the toddler dataset

Method	Feature Selection	Approach	Accuracy (±SD)	Accuracy Imbalanced Data
[23]	No	AdaBoost, LDA	99.25 ± 0.14	96.5
[24]	Yes	RF-XGB with MFS	98.6 ± 0.13	97.5
[25]	Yes	SVM, AdaBoost, Glmboost	97.82 ± 0.21	94.1
[26]	Yes	AdaBoost	99.85 ± 0.17	96.5
Proposed model	Yes	FOIDE	99.91 ± 0.8	99.12

4. CLASSIFIER PERFORMANCE STUDY FOR ADULT DATASET

The final evaluation is conducted using the UCI adult ASD dataset, which has a moderate class imbalance (1:1.7). The proposed FOIDE model is compared against prior methods, including LR-based [18] and Linear Discriminant Analysis (LDA)-based [23]. Although AdaBoost [23] was also considered, it showed subpar performance compared to LDA and was excluded from graphical comparison for clarity. Figure 6 presents the comparative results. FOIDE exhibits superior performance across all evaluation metrics, including precision, recall, and F1-score. This improvement is especially pronounced in recall, affirming FOIDE’s ability to correctly identify ASD-positive cases. This robustness stems from its enhanced XGB component that incorporates weighted loss functions, along with a feature selection mechanism that mitigates bias caused by class imbalance.

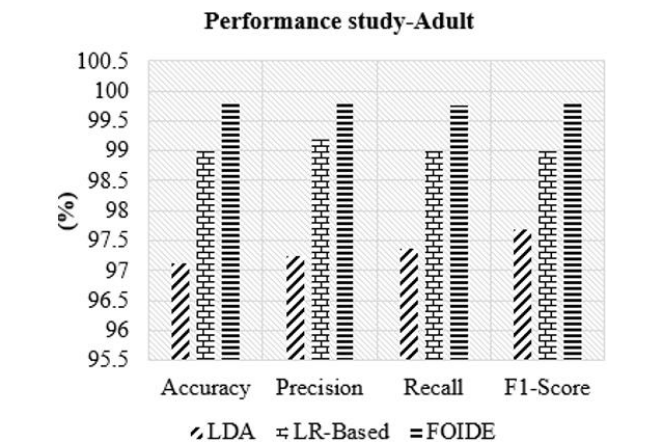


Figure 6. Classification performance for the ASD adult dataset

The consistently high performance across all age groups and datasets reinforces FOIDE’s capacity to generalize well, especially when trained on imbalanced datasets a frequent

characteristic in real-world ASD screening. Key enhancements that justify FOIDE's superiority are as follows:

Imbalance handling: Unlike conventional models, FOIDE explicitly incorporates a cost-sensitive optimization strategy in XGB to reduce false negatives, a common issue in minority ASD class detection.

Cross-validation aware feature ranking: The nested K-fold cross-validation ensures robust feature importance estimation and generalization across folds, preventing overfitting to dominant class samples.

MLP-based classification with selected features: By feeding XGB-derived features into an MLP with refined weights, FOIDE achieves better nonlinear separation and class discrimination, enhancing prediction on subtle ASD patterns.

5. CONCLUSION

Numerous studies have sought to enhance ASD prediction, yet challenges such as class imbalance and generalizability across age groups remain unresolved. This study addressed these gaps by compiling and analyzing ASD screening datasets spanning toddlers, children, and adults. We introduced the FOIDE model, which integrates a cost-sensitive XGBoost-based feature extractor with an MLP classifier, guided by an enhanced nested cross-validation mechanism for effective feature selection and ranking. Our experimental results demonstrate that FOIDE consistently achieves competitive accuracy and robust generalization across all age categories when compared to existing machine learning baselines, particularly in handling imbalanced datasets. However, certain limitations must be acknowledged. FOIDE’s computational complexity, due to its ensemble architecture and repeated cross-validation, may pose deployment challenges in resource-constrained environments such as low-power edge devices or real-time systems. Additionally, the current model relies solely on structured screening data and does not incorporate other potentially rich modalities such as speech, facial expression, or wearable sensor data.

Future research will explore extending FOIDE to multi-modal ASD detection, integrating behavioral signals (e.g., eye tracking, voice patterns, physiological data) with structured questionnaire responses to improve prediction accuracy and explainability. Additionally, efforts will focus on developing lightweight variants of FOIDE suitable for real-time or mobile health (mHealth) applications in clinical and home environments [34, 35].

ACKNOWLEDGMENT

We sincerely thank the PES, Mandya, Karnataka, faculty, and friends for their support in all aspects, which contributed to the completion of this work and paper. We also thank our parents and the almighty for all the moral support and encouragement.

REFERENCES

[1] Basri, M.A.F.A., Ismail, W.S.W., Nor, N.K., Tohit, N.M., et al. (2024). Development of a social communication skills intervention using video modeling and spherical video-based virtual reality for high

- functioning autism spectrum disorder youth: A preliminary study. *IEEE Access*, 12: 77976-77987. <https://doi.org/10.1109/ACCESS.2024.3407114>
- [2] Qureshi, M.S., Qureshi, M.B., Asghar, J., Alam, F., et al. (2023). Prediction and analysis of autism spectrum disorder using machine learning techniques. *Journal of Healthcare Engineering*, 2023(1): 4853800. <https://doi.org/10.1155/2023/4853800>
 - [3] Koehler, J.C., Dong, M.S., Song, D.Y., Bong, G., et al. (2024). Classifying autism in a clinical population based on motion synchrony: A proof-of-concept study using real-life diagnostic interviews. *Scientific Reports*, 14(1): 5663. <https://doi.org/10.1038/s41598-024-56098-y>
 - [4] Chistol, M., Turcu, C., Danubianu, M. (2023). Autism assistant: A platform for autism home-based therapeutic intervention. *IEEE Access*, 11: 94188-94204. <https://doi.org/10.1109/ACCESS.2023.3310397>
 - [5] Haque, N., Islam, T., Erfan, M. (2025). An exploration of machine learning approaches for early autism spectrum disorder detection. *Healthcare Analytics*, 7: 100379. <https://doi.org/10.1016/j.health.2024.100379>
 - [6] Silva, B., Santos, L., Barata, C., Geminiani, A., et al. (2024). Attention analysis in robotic-assistive therapy for children with autism. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 32: 2220-2229. <https://doi.org/10.1109/TNSRE.2024.3411299>
 - [7] Wei, Q., Xu, X., Xu, X., Cheng, Q. (2023). Early identification of autism spectrum disorder by multi-instrument fusion: A clinically applicable machine learning approach. *Psychiatry Research*, 320: 115050. <https://doi.org/10.1016/j.psychres.2023.115050>
 - [8] Yaneva, V., Eraslan, S., Yesilada, Y., Mitkov, R. (2020). Detecting high-functioning autism in adults using eye tracking and machine learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(6): 1254-1261. <https://doi.org/10.1109/TNSRE.2020.2991675>
 - [9] Jamwal, I., Malhotra, D., Mengi, M. (2023). A systematic study of intelligent autism spectrum disorder detector. *International Journal of Computational Vision and Robotics*, 13(2): 219-234. <https://doi.org/10.1504/IJCVR.2023.129435>
 - [10] Hassan, I., Nahid, N., Islam, M., Hossain, S., et al. (2025). Automated autism assessment with multimodal data and ensemble learning: A scalable and consistent robot-enhanced therapy framework. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 33: 1191-1201. <https://doi.org/10.1109/TNSRE.2025.3546519>
 - [11] Lakhan, A., Mohammed, M.A., Abdulkareem, K.H., Hamouda, H., et al. (2023). Autism spectrum disorder detection framework for children based on federated learning integrated CNN-LSTM. *Computers in Biology and Medicine*, 166: 107539. <https://doi.org/10.1016/j.combiomed.2023.107539>
 - [12] Arafat, M.Y., Pan, S., Bak, E. (2023). Distributed energy-efficient clustering and routing for wearable IoT enabled wireless body area networks. *IEEE Access*, 11: 5047-5061. <https://doi.org/10.1109/ACCESS.2023.3236403>
 - [13] Kumar, V.C., Umesh, D.R. (2025). Effective autism spectrum disorder sensory and behavior data collection using internet of things. *Indonesian Journal of Electrical Engineering and Computer Science*, 37(2): 1274-1283. <http://doi.org/10.11591/ijeecs.v37.i2.pp1274-1283>
 - [14] Farooq, M.S., Tehseen, R., Sabir, M., Atal, Z. (2023). Detection of autism spectrum disorder (ASD) in children and adults using machine learning. *Scientific Reports*, 13(1): 9605. <https://doi.org/10.1038/s41598-023-35910-1>
 - [15] Khudhur, D.D., Khudhur, S.D. (2023). The classification of autism spectrum disorder by machine learning methods on multiple datasets for four age groups. *Measurement: Sensors*, 27: 100774. <https://doi.org/10.1016/j.measen.2023.100774>
 - [16] Shinde, A.V., Patil, D.D. (2023). A multi-classifier-based recommender system for early autism spectrum disorder detection using machine learning. *Healthcare Analytics*, 4: 100211. <https://doi.org/10.1016/j.health.2023.100211>
 - [17] Twala, B., Molloy, E. (2023). On effectively predicting autism spectrum disorder therapy using an ensemble of classifiers. *Scientific Reports*, 13(1): 19957. <https://doi.org/10.1038/s41598-023-46379-3>
 - [18] Bawa, P., Kadyan, V., Mantri, A., Vardhan, H. (2024). Investigating multiclass autism spectrum disorder classification using machine learning techniques. *E-Prime-Advances in Electrical Engineering, Electronics and Energy*, 8: 100602. <https://doi.org/10.1016/j.prime.2024.100602>
 - [19] Reghunathan, R.K., Venkidusamy, P. N. P., Kurup, R.G., George, B., et al. (2024). Machine learning-based classification of autism spectrum disorder across age groups. *Engineering Proceedings*, 62(1): 12. <https://doi.org/10.3390/engproc2024062012>
 - [20] J. Koehler, J.C., Dong, M.S., Bierlich, A.M., Fischer, S., et al. (2024). Machine learning classification of autism spectrum disorder based on reciprocity in naturalistic social interactions. *Translational Psychiatry*, 14(1): 76. <https://doi.org/10.1038/s41398-024-02802-5>
 - [21] Thapa, R., Garikipati, A., Ciobanu, M., Singh, N.P., et al. (2024). Machine learning differentiation of autism spectrum sub-classifications. *Journal of Autism and Developmental Disorders*, 54(11): 4216-4231. <https://doi.org/10.1007/s10803-023-06121-4>
 - [22] Rasul, R.A., Saha, P., Bala, D., Karim, S.R.U., et al. (2024). An evaluation of machine learning approaches for early diagnosis of autism spectrum disorder. *Healthcare Analytics*, 5: 100293. <https://doi.org/10.1016/j.health.2023.100293>
 - [23] Hasan, S.M., Uddin, M.P., Al Mamun, M., Sharif, et al. (2022). A machine learning framework for early-stage detection of autism spectrum disorders. *IEEE Access*, 11: 15038-15057. <https://doi.org/10.1109/ACCESS.2022.3232490>
 - [24] Hajje, F., Ayouni, S., Alohal, M.A., Maddeh, M. (2024). Novel framework for autism spectrum disorder identification and tailored education with effective data mining and ensemble learning techniques. *IEEE Access*, 12: 35448-35461. <https://doi.org/10.1109/ACCESS.2024.3349988>
 - [25] Bala, M., Ali, M.H., Satu, M.S., Hasan, K.F., et al. (2022). Efficient machine learning models for early stage detection of autism spectrum disorder. *Algorithms*, 15(5): 166. <https://doi.org/10.3390/a15050166>
 - [26] Uddin, M.J., Ahamad, M.M., Sarker, P.K., Aktar, S., et al. (2023). An integrated statistical and clinically applicable machine learning framework for the detection

- of autism spectrum disorder. *Computers*, 12(5): 92. <https://doi.org/10.3390/computers12050092>
- [27] Farooq, M.S., Tehseen, R., Sabir, M., Atal, Z. (2023). Detection of autism spectrum disorder (ASD) in children and adults using machine learning. *Scientific Reports*, 13(1): 9605. <https://doi.org/10.1038/s41598-023-35910-1>
- [28] Ehsan, K., Sultan, K., Fatima, A., Sheraz, M., et al. (2025). Early detection of autism spectrum disorder through automated Machine Learning. *Diagnostics*, 15(15): 1859. <https://doi.org/10.3390/diagnostics15151859>
- [29] Jahedi, A., Salehi, M., Goltapeh, E.M., Safaie, N. (2023). Multilayer perceptron-genetic algorithm as a promising tool for modeling cultivation substrate of *Auricularia cornea* Native to Iran. *PloS One*, 18(2): e0281982. <https://doi.org/10.1371/journal.pone.0281982>
- [30] Fayes, F. (2022). Autism Screening for Toddlers, Kaggle. <https://www.kaggle.com/datasets/fabdelja/autism-screening-for-toddlers>.
- [31] ASD Screening Data for Toddlers in Saudi Arabia, Kaggle, 2022. <https://www.kaggle.com/datasets/asdpredictioninsaudi/asd-screening-data-for-toddlers-in-saudi-arabia>.
- [32] Thabtah, F.F. (2017). Autistic spectrum disorder screening data for children. *UCI Machine Learning Repository*, 10: C5659W. <https://doi.org/10.1016/j.jksus.2024.103468>
- [33] Autism Screening Adult 2017. <https://archive.ics.uci.edu/dataset/426/autism+screening+adult>.
- [34] Shukur, F., Mosa, S.J., Raheem, K.M.H. (2024). Optimization of fuzzy-PD control for a 3-DOF robotics manipulator using a Back-Propagation Neural Network. *Mathematical Modelling of Engineering Problems*, 11(1): 199-209. <https://doi.org/10.18280/mmep.110122>
- [35] Okokpujie, K., Okokpujie, I.P., Ayomikun, O.I., Orimogunje, A.M., Ogundipe, A.T. (2023). Development of a web and mobile applications-based cassava disease classification interface using Convolutional Neural Network. *Mathematical Modelling of Engineering Problems*, 10(1): 119-128. <https://doi.org/10.18280/mmep.100113>