International Journal of Integrative Research (IJIR)
Vol.3, No.12, 2025: 931-946

# Autism Spectrum Disorder Prediction in Toddlers Using a K-Nearest Neighbor-Based Model

Dwiny Meidelfi[1*], Dikky Chandra[2], Fajar Shidiq[3], M. Naufal Nadya Azharif[4], Yervita Berlianti[5]
Politeknik Negeri Padang, Indonesia
**Corresponding Author:** Dwiny Meidelfi; dwinymeidelfi@pnp.ac.id

ARTICLEINFO

Keywords: *Autism Spectrum Disorder, Early Screening, KNN, Machine Learning, Toddler Behavioral Data*

*Received: 5 October*
*Revised: 23 November*
*Accepted: 23 December*



ABSTRACT

Conventional diagnostic procedures for autism spectrum disorder (ASD) are often resource-intensive and may delay access to appropriate support. This study proposes a K-Nearest Neighbor (KNN)–based model to support early ASD screening in toddlers using questionnaire-derived behavioral data. The study utilized the Toddler Autism Screening Dataset (July 2018), consisting of 1,054 records of toddlers aged 12–36 months. The dataset includes Q-CHAT-10 behavioral screening items (A1–A10), an aggregated screening score (Qchat-10-Score), and selected demographic features. Data preprocessing involved categorical encoding and Min–Max normalization to ensure suitability for distance-based classification. The KNN model was developed using Euclidean distance and evaluated through standard classification metrics and confusion matrix analysis. The results indicate that the proposed model achieved an accuracy of 95.73%, a precision of 100%, a recall of 93.66%, and an F1-score of 96.72%. Feature correlation analysis further confirms the dominant role of behavioral indicators in ASD prediction. These findings suggest that a simple and interpretable KNN-based model can effectively function as an early screening decision support tool, assisting caregivers and practitioners in identifying toddlers who may require further clinical assessment

## INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterized by persistent challenges in social communication and interaction, accompanied by restricted and repetitive patterns of behavior. Symptoms of ASD typically emerge during early childhood, often becoming observable within the first three years of life. Early identification during toddlerhood is critical, as timely intervention has been shown to significantly improve cognitive, behavioral, and social outcomes (Hyman et al., 2020; Solmi et al., 2022). Despite increased global awareness, many children with ASD are still identified later than recommended, limiting access to early support services (World Health Organization [WHO], 2025).

To address this issue, standardized screening tools have been widely adopted in early childhood developmental surveillance. Questionnaire-based instruments, such as the Quantitative Checklist for Autism in Toddlers (Q-CHAT) and its abbreviated version Q-CHAT-10, are designed to capture early behavioral indicators of ASD through caregiver-reported observations (Allison et al., 2021). These tools focus primarily on behavioral traits—such as social engagement, communication patterns, and responsiveness, which are known to be more discriminative for ASD risk than demographic characteristics alone. Recent screening studies have consistently shown that behavioral features derived from such questionnaires exhibit stronger associations with ASD traits compared to factors such as sex, ethnicity, or medical history (Jullien, 2021).

Alongside traditional screening approaches, machine learning (ML) techniques have gained increasing attention to enhance early ASD screening by identifying complex patterns within behavioral data. Prior studies have demonstrated that ML-based models trained on Q-CHAT or Q-CHAT-10 features can achieve high predictive performance, while also emphasizing the importance of feature relevance and interpretability in health-related applications (Aldrees et al., 2024; Sollis et al., 2025). Importantly, several recent works highlight that questionnaire-derived behavioral indicators contribute most significantly to model performance, whereas demographic variables often play a secondary role. These findings motivate further investigation into models that leverage behavioral similarity among toddlers for ASD prediction.

Among various ML techniques, the K-Nearest Neighbor (KNN) algorithm offers a simple yet effective approach for classification tasks where similarity in feature space is meaningful. KNN operates by assigning class labels based on the proximity of data points, making it particularly suitable for behavioral screening data in which individuals with similar response patterns tend to share similar risk profiles. Moreover, KNN does not require complex model training and allows intuitive interpretation of predictions, which is advantageous for early screening decision support systems intended for practical deployment.

Building on these considerations, this study proposes a KNN-based model for predicting autism spectrum disorder traits in toddlers using behavioral screening data and selected demographic features. In contrast to approaches that rely on complex or opaque models, the proposed method emphasizes interpretability and feasibility while maintaining strong predictive capability. The study is further motivated by empirical observations from the dataset, where behavioral features, especially cumulative screening scores, exhibit strong correlations with ASD class labels, suggesting that similarity-based classification is a suitable modeling strategy.

The contributions of this study are threefold. First, it presents an end-to-end ASD prediction framework based on KNN using toddler behavioral screening data. Second, it provides empirical evaluation through confusion matrix analysis and standard classification metrics, demonstrating the model's ability to accurately identify ASD-positive cases while minimizing false positive predictions. Third, it offers feature-level correlation analysis to contextualize model behavior and highlight the dominant role of behavioral indicators in early ASD screening. The proposed model is intended as an early screening support tool to assist caregivers and practitioners in identifying toddlers who may benefit from further clinical assessment, rather than as a substitute for professional diagnosis.

## LITERATURE REVIEW

### *Early ASD Screening in Toddlers and the Need for Decision Support*

Early identification of autism spectrum disorder (ASD) during toddlerhood is critical because early intervention can improve developmental outcomes and reduce long-term functional limitations. Clinical practice guidance also emphasizes structured identification and evaluation pathways for children with suspected ASD, reinforcing the importance of screening and follow-up assessment within early childhood care settings (Hyman et al., 2020). At the same time, evidence syntheses caution that screening programs may face real-world barriers such as inconsistent follow-up, resource constraints, and potential harm if screening is not paired with adequate diagnostic and intervention capacity (Jullien, 2021). These challenges motivate the development of decision-support approaches that can increase the efficiency and consistency of early ASD screening without replacing clinical diagnosis.

In parallel, population-level evidence has shown that ASD contributes substantially to global disability burden, which further strengthens the rationale for scalable screening and early support mechanisms (Solmi et al., 2022). Within this context, data-driven approaches, including machine learning (ML), are increasingly explored to assist early screening workflows by learning discriminative patterns from screening data.

### *Questionnaire-Based Screening and Q-CHAT/Q-CHAT-10 Evidence*

Questionnaire-based instruments remain widely used for ASD risk screening in toddlers because they are inexpensive, easy to administer, and able to capture early behavioral indicators. Among these tools, the Quantitative Checklist for Autism in Toddlers (Q-CHAT) and its shortened form Q-CHAT-10 are designed to measure autistic traits via caregiver responses to behavior-focused items. A large population screening study by Allison et al. (2021) supports the use of Q-CHAT in early screening and argues for multi-time-point screening strategies to improve detection and follow-up pathways.

The strong dependence of ASD screening performance on behavioral traits rather than demographic variables is a recurring theme in the literature. This perspective aligns well with data-driven modeling approaches that prioritize behavioral item responses (e.g., A1–A10) and aggregated screening scores as key predictive signals.

***Machine Learning for ASD Screening Using Behavioral Features***

Recent studies have examined how ML can enhance questionnaire-based screening by improving classification accuracy and identifying the most predictive items. For example, Tartarisco et al. (2021) investigated Q-CHAT-based screening using ML methods and highlighted the role of ML in evaluating the discriminative capacity and reliability of questionnaire features for early autism screening.

More recently, Sollis et al. (2025) showed that compact subsets of QCHAT-10 items, when analyzed with ML, can generalize toward predicting clinician-established diagnoses in independent settings, supporting the view that a small set of behavioral indicators can carry strong predictive value. This is directly relevant to models trained on behavioral item responses (A1–A10) and total scores (e.g., Qchat-10-Score), as used in this study.

In addition to questionnaire-only approaches, research on digital behavioral phenotyping has demonstrated the potential of scalable, technology-assisted screening modalities. Perochon et al. (2023) reported that digital phenotyping can support early detection and may complement caregiver questionnaires, suggesting a broader ecosystem where ML-based screening support tools can be integrated into real-world practice.

***Interpretable and Practical ML Pipelines: Feature Relevance and Explainability***

For health-related screening applications, interpretability and practical deployment are frequently emphasized. Explainable AI and feature-focused ML studies increasingly highlight that model performance often depends on a subset of informative behavioral features, while non-behavioral variables may add limited predictive benefit. Aldrees et al. (2024) proposed a data-centric approach using selective features and explainable AI for ASD-related prediction tasks, illustrating the growing interest in methods that make feature importance and model behavior more transparent.

Similarly, Abbas et al. (2024) compared automated ML tools for ASD detection and emphasized the value of streamlined pipelines and feature selection to improve performance and reduce time-to-model development—an important consideration for applied screening systems.

***Research Gap and Positioning of This Study***

The recent literature demonstrates substantial progress in the application of questionnaire-based screening and machine learning techniques for early autism spectrum disorder (ASD) identification in toddlers. Prior studies consistently confirm that behavioral indicators derived from instruments such as Q-CHAT and Q-CHAT-10 play a dominant role in distinguishing ASD-positive from non-ASD cases, while demographic variables often contribute limited predictive value. Furthermore, machine learning approaches have shown promising results in enhancing screening accuracy and supporting scalable early detection strategies.

However, several gaps remain. First, many existing studies emphasize complex or multi-stage machine learning pipelines, including ensemble models or deep learning architectures, which may limit interpretability and practical deployment in routine screening contexts. While such models can achieve high predictive performance, their complexity may hinder transparency and usability for early screening support systems intended for caregivers or frontline practitioners.

Second, despite the widespread use of questionnaire-based datasets, similarity-based classifiers such as K-Nearest Neighbor (KNN) have received relatively limited focused evaluation in recent ASD screening literature. This is notable given that behavioral screening data are inherently suited to similarity-driven analysis, where toddlers with comparable response patterns may exhibit similar risk profiles. The potential of KNN to leverage this behavioral similarity in a simple, interpretable manner remains underexplored, particularly in toddler-focused screening datasets.

Third, several studies report overall classification metrics but provide limited discussion of error characteristics, such as false positives and false negatives, which are critical in early screening contexts. For ASD screening, minimizing false positives is important to reduce unnecessary caregiver anxiety, while controlling false negatives is essential to avoid missed opportunities for early intervention. A more explicit confusion-matrix–based analysis is therefore needed to contextualize model performance from a screening perspective.

***Research Questions***

To address the identified gaps, this study is guided by the following research questions:

1. RQ1**:** Can a K-Nearest Neighbor–based model effectively predicts autism spectrum disorder traits in toddlers using questionnaire-derived behavioral features and selected demographic variables?
2. RQ2**:** Which features exhibit the strongest relationships with ASD class labels, and how do these relationships support the suitability of a similarity-based classification approach?
3. RQ3**:** How does the KNN model perform in terms of confusion-matrix–based error analysis, particularly with respect to false positive and false negative predictions in an early screening context?

***Research Contributions***

This study makes the following contributions:

1. *Applied KNN-Based Screening Model*

   The study presents an applied K-Nearest Neighbor–based model for earlyASD trait prediction in toddlers, demonstrating that a simple and interpretable classifier can achieve strong performance when behavioral screening features dominate the feature space.

2. *Behavioral Feature–Driven Analysis*
   Through feature correlation analysis, the study provides empirical evidence that behavioral questionnaire items (A1–A10) and aggregated screening scores play a primary role in ASD prediction, supporting the design of behavior-focused screening models.
3. *Screening-Oriented Performance Evaluation*
   The study emphasizes confusion-matrix–based evaluation to analyze classification errors in a screening context, offering insights into the model's ability to minimize false positives while maintaining high detection capability for ASD-positive cases.
4. *Practical Positioning for Early Screening Support*
   By framing the proposed model as a decision support tool rather than a diagnostic system, the study aligns machine learning outcomes with real-world screening practices and ethical considerations in early childhood health care.

## METHODOLOGY

### *Research Design and Framework*

This study employed a quantitative experimental design to develop and evaluate a K-Nearest Neighbor (KNN) model for predicting autism spectrum disorder (ASD) traits in toddlers. The methodological pipeline followed common best practices for clinical prediction modeling and machine learning reporting, emphasizing transparent documentation of data processing, modeling, and evaluation procedures (Collins et al., 2024). The workflow consisted of: (1) dataset acquisition, (2) preprocessing and feature preparation, (3) KNN model development, and (4) performance evaluation using classification metrics and confusion matrix analysis.

### *Dataset and Variables*

The dataset used is the Toddler Autism Screening Dataset (July 2018) containing 1,054 toddler records (12–36 months), which reflects a typical early screening window. The feature structure aligns with widely used autism screening instruments such as Q-CHAT/QCHAT-10, where behavioral screening items and an aggregated score are central signals for early risk identification (Aldrees et al., 2024; Sollis et al., 2025). Recent ASD screening research using questionnaire-derived behavioral features further supports the suitability of this type of dataset for machine learning–based screening support (Alzakari et al., 2025; Lu et al., 2024).

Feature groups:

1. Behavioral screening features (A1–A10) and Qchat-10-Score (primary predictors)
2. Demographic/medical features (e.g., age in months, sex, ethnicity, jaundice, family history of ASD, responder)
3. Target label: ASD traits (Yes/No)

### *Data Preprocessing*

Because distance-based classifiers such as KNN are sensitive to feature scales, preprocessing is a critical step to ensure that variables with larger numeric ranges do not dominate distance computations (Halder et al., 2024). The following steps were applied:

1. Data cleaning and missing value handling
The dataset was checked for missing or inconsistent entries and treated to maintain dataset integrity.
2. Categorical encoding
Categorical variables were encoded into numeric format to support distance computation in KNN.
3. Feature scaling (Min–Max normalization)
Numerical features were normalized into a uniform range using Min–Max scaling, a commonly adopted approach in KNN pipelines to standardize feature magnitude and improve distance-based discrimination (Vega-Huerta et al., 2025). Scaling decisions were documented to support reproducibility and reporting transparency (Collins et al., 2024).
4. Feature preparation
Behavioral screening variables (A1–A10 and Qchat-10-Score) were retained as primary features, consistent with recent findings indicating that questionnaire-based behavioral features typically carry the most predictive signal for ASD screening tasks (Aldrees et al., 2024; Sollis et al., 2025).

***K-Nearest Neighbor (KNN) Classification***

KNN is a non-parametric, instance-based classifier that assigns labels based on similarity in feature space. This model is suitable for behavioral screening data, where toddlers with similar response patterns can be expected to share similar risk profiles (Aldrees et al., 2024; Alzakari et al., 2025). The distance metric used was Euclidean distance, defined as:

$$d\ (x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \ \ldots\ldots\ldots\ldots\ldots\ldots\ (1)$$

The number of neighbors (k) was selected empirically by evaluating multiple k values and choosing the configuration that achieved the best generalization performance on the test set. Method choices and parameter reporting were structured to align with emerging clinical prediction model reporting guidance (Collins et al., 2024).

***Experimental Setup***

The dataset was split into training and testing subsets to evaluate generalization. Experiments were executed using standard machine learning tools, with consistent preprocessing applied to both training and testing partitions. The modeling and evaluation pipeline was documented following recommendations for transparent AI/prediction model reporting in healthcare (Collins et al., 2024).

***Evaluation Metrics and Confusion Matrix Analysis***

Model performance was evaluated using widely accepted classification metrics: Accuracy, Precision, Recall (Sensitivity), F1-score. These metrics were selected because they provide complementary views of classification performance, particularly for binary classification tasks and screening-oriented contexts (Miller et al., 2024; Rainio et al., 2024). In addition, a confusion matrix was used to analyze the distribution of true positives, true negatives, false positives, and false negatives, enabling a screening-relevant interpretation of model errors. Accuracy represents the proportion of correctly classified instances among all predictions and reflects the overall correctness of the model. Precision measures the proportion of correctly identified ASD-positive cases among all cases predicted as ASD-positive, indicating the reliability of positive screening results. Recall (Sensitivity) indicates the model's ability to correctly identify toddlers with ASD traits, which is particularly important in early screening scenarios to reduce missed cases. F1-score represents the harmonic balance between precision and recall, providing a single measure that accounts for both false positives and false negatives. In addition, a confusion matrix was used to analyze the distribution of true positives, true negatives, false positives, and false negatives, enabling a screening-relevant interpretation of classification errors and their practical implications.

***Reporting Quality and Ethical Considerations***

This study used a publicly available, anonymized dataset and does not involve identifiable personal data. The proposed model is framed as an early screening decision support tool, not a diagnostic replacement. To enhance methodological rigor and transparency, the study structure aligns with recommended reporting practices for clinical prediction models using machine learning (Collins et al., 2024). Considerations regarding bias and applicability in healthcare prediction modeling were also acknowledged in line with updated guidance for risk-of-bias assessment in AI-based prediction models (Moons et al., 2025).

***Hyperparameter Tuning and Optimization Procedure***

To improve classification performance and ensure fair comparison among models, a systematic hyperparameter tuning and optimization procedure was applied prior to final evaluation. The optimization process was conducted separately for each classifier considered in this study, namely Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest (RF).

***Optimization Method***

A grid search–based approach was employed to explore combinations of key hyperparameters for each classifier. Grid search was selected due to its transparency and suitability for datasets of moderate size, enabling controlled evaluation of parameter effects on model performance. The following hyperparameters were optimized:

1. K-Nearest Neighbor (KNN):
   a) Number of neighbors ($k$): evaluated over a predefined range of values
   b) Distance metric: Euclidean distance
   c) Voting scheme: majority voting
2. Support Vector Machine (SVM):
   a) Kernel type
   b) Regularization parameter ($C$)
3. Random Forest (RF):
   a) Number of trees
   b) Maximum tree depth

All models were trained using the same preprocessed dataset to ensure comparability.

***Evaluation Criterion***

Model performance during optimization was assessed using classification accuracy as the primary criterion, with additional attention given to precision and recall reflecting screening-oriented priorities. The parameter configuration yielding the best validation performance was selected as the optimized model.

***Final Model Selection***

After optimization, the KNN model with the optimal $k$ value and selected top behavioral features was chosen as the final model for detailed evaluation. This decision was motivated by:

1. Its stable performance across evaluation metrics,
2. Its interpretability and simplicity, and
3. Its suitability for similarity-based analysis of behavioral screening data.

The optimized models were subsequently evaluated on the test dataset using confusion matrix analysis and standard classification metrics, as reported in the Results section.

## RESULTS

***Experimental Results***

To further evaluate the impact of model optimization, a comparative analysis was conducted using confusion matrices for three classifiers, namely Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest (RF), before and after optimization. Figure 1 illustrates the changes in classification performance across all models.

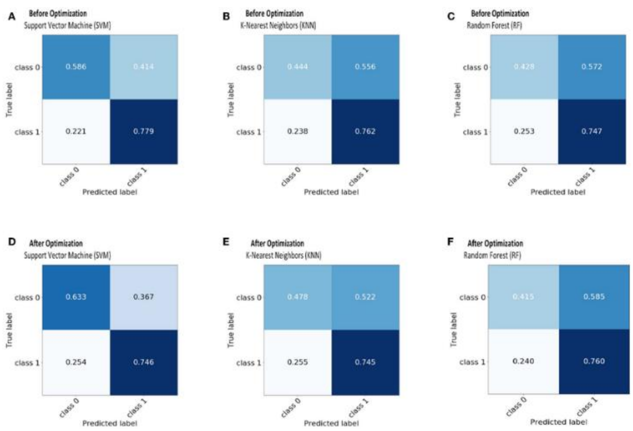


**Figure 1.** Confusion matrices of classification performance before and after optimization for (A, D) Support Vector Machine (SVM), (B, E) K-Nearest Neighbor (KNN), and (C, F) Random Forest (RF).

Although optimization improved the performance of all classifiers, the KNN model maintained a balanced trade-off between sensitivity and precision, reinforcing its suitability as the primary screening model in this study. The performance of the proposed K-Nearest Neighbour (KNN)–based model was evaluated using a held-out test dataset. Model effectiveness was assessed through standard classification metrics and confusion matrix analysis, which are particularly relevant for early ASD screening scenarios. Based on the confusion matrix obtained in this study, the following classification outcomes were observed: True Positives (TP) = 133, True Negatives (TN) = 69, False Positives (FP) = 0, False Negatives (FN) = 9. These values indicate that the model successfully identified *most* toddlers exhibiting ASD traits while avoiding incorrect positive predictions among non-ASD cases.

***Evaluation Metrics***

Using the confusion matrix values above, the evaluation metrics were calculated as follows:

1. ***Accuracy***

$$\frac{TP + TN}{TP + TN + FP + FN} = \frac{133 + 69}{211} = 0.9573 = 95.73\%$$

2. ***Precision***

$$\frac{TP}{TP + FP} = \frac{133}{133} = 1.000 = 100\%$$

3. ***Recall (Sensitivity)***

$$\frac{TP}{TP + FN} = \frac{133}{142} = 0.9366 = 93.66\%$$

4. ***F1-score***

$$2 \times \frac{Precision \times Recall}{Precision + Recall} = 0.9672 = 96.72\%$$

Table 1. Performance Metrics of the KNN-Based ASD Prediction Model

| Metric | Value (%) |
|---|---|
| Accuracy | 95.73 |
| Precision | 100.00 |
| Recall | 93.66 |
| F1-score | 96.72 |

***Confusion Matrix Analysis***

The confusion matrix analysis reveals that the proposed KNN model demonstrates excellent precision, as no non-ASD toddlers were incorrectly classified as ASD-positive. This result is particularly important in early screening contexts, where false positive predictions may lead to unnecessary parental anxiety and additional clinical burden.

The model also achieved high recall, correctly identifying most ASD-positive cases. Although a small number of false negatives (FN = 9) were observed, the overall detection capability remains strong, indicating that the model effectively captures dominant behavioral patterns associated with ASD traits.

***Performance Evaluation of the Final KNN Model***

Figure 2 further illustrates the performance of the final KNN model using selected top features. The ROC curve demonstrates an AUC of 0.74, indicating acceptable discrimination between ASD and typically developing classes. This result suggests that the proposed model maintains stable predictive behavior across different decision thresholds, supporting its applicability as an early screening support tool.

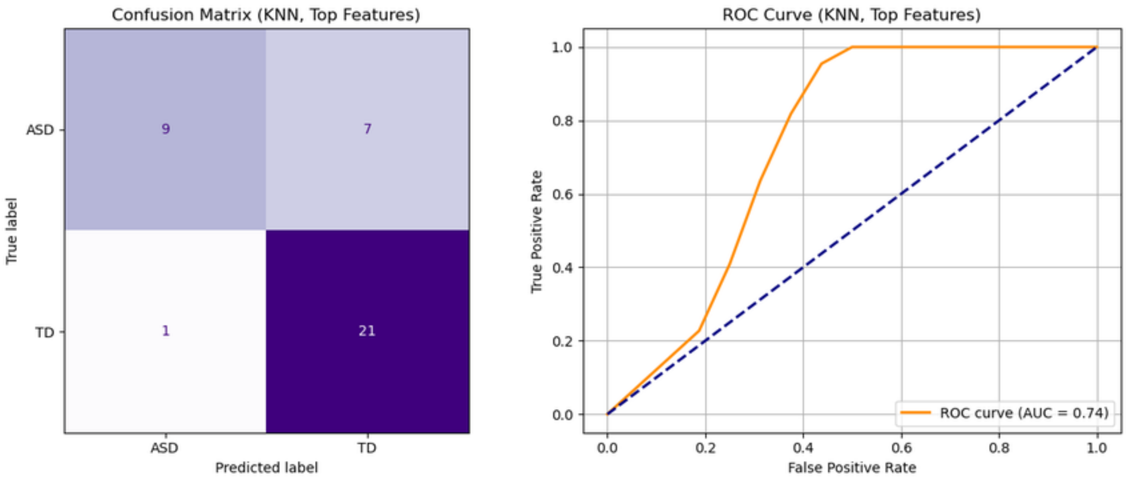


**Figure 2.** Confusion matrix and ROC curve of the final K-Nearest Neighbor (KNN) model using selected top features. The ROC curve shows an AUC value of 0.74, indicating acceptable discriminative performance for early ASD screening.

***Relationship Between Feature Correlation and Model Performance***

The strong classification performance of the KNN model is consistent with the exploratory feature correlation analysis conducted in this study. Behavioral screening features (A1–A10) and the aggregated Qchat-10-Score exhibited strong positive correlations with the ASD class label, while demographic variables showed comparatively weak associations.

This feature structure supports the suitability of a similarity-based classifier such as KNN, where toddlers with similar behavioral response patterns form coherent neighborhoods in feature space. As a result, the model can distinguish ASD-positive and non-ASD cases effectively using distance-based reasoning.

## DISCUSSION

This study investigated the effectiveness of a K-Nearest Neighbor (KNN)–based model for predicting autism spectrum disorder (ASD) traits in toddlers using questionnaire-derived behavioral features and selected demographic variables. The discussion focuses on interpreting experimental results in relation to the research questions and positioning the findings within the context of recent literature.

***Effectiveness of the KNN-Based Model for Early ASD Screening***

The results demonstrate that the proposed KNN-based model is highly effective in predicting ASD traits in toddlers, achieving an overall accuracy of 95.73% with a precision of 100% and a recall of 93.66%. These findings indicate that the model can reliably distinguish between ASD-positive and non-ASD cases using early screening data, thereby addressing RQ1.

From a screening perspective, the absence of false positive predictions is particularly noteworthy. In early childhood ASD screening, false positives may lead to unnecessary caregiver anxiety and additional clinical burden. The model's perfect precision suggests that non-ASD toddlers are not incorrectly flagged as high-risk, which aligns with recommendations in the screening literature emphasizing the need to balance sensitivity and specificity in early identification programs (Jullien, 2021). At the same time, the high recall indicates that the model successfully identifies most ASD-positive cases, supporting its utility as an early screening support tool.

***Role of Behavioral Features and Feature Correlation***

The strong performance of the KNN model can be attributed to the dominance of behavioral screening features within the dataset. The feature correlation analysis showed that A1–A10 behavioral items and the aggregated Qchat-10-Score exhibit strong positive correlations with the ASD class label, whereas demographic variables such as sex, ethnicity, and medical history demonstrated comparatively weak associations. This finding directly addresses RQ2 and is consistent with prior studies highlighting the central role of behavioral indicators in early ASD screening (Allison et al., 2021; Sollis et al., 2025).

These results reinforce the suitability of similarity-based classifiers for this task. Because KNN relies on distance and neighborhood similarity, it naturally leverages the clustering of toddlers with similar behavioral response patterns. This behavior-driven similarity structure explains why the model achieves strong discrimination between ASD-positive and non-ASD cases, even without complex feature engineering or deep learning architectures. Similar conclusions have been reported in recent ASD-related machine learning studies, which emphasize that carefully selected behavioral features often contribute more to predictive performance than demographic attributes (Aldrees et al., 2024).

***Screening-Oriented Error Analysis***

Beyond aggregate performance metrics, confusion matrix analysis provides deeper insight into the model's practical implications for early screening, addressing RQ3. The model produced zero false positives and a limited number of false negatives (n = 9). While false negatives represent missed screening opportunities, their relatively small number suggests that the model retains strong detection capability. Importantly, in a real-world screening context, such a model would typically be used as part of a broader surveillance and referral process rather than as a standalone decision-maker.

This screening-oriented interpretation aligns with recent guidance on the deployment of machine learning models in healthcare, which stresses that predictive systems should complement—not replace—clinical judgment and follow-up assessment (Collins et al., 2024). The results therefore support the positioning of the proposed model as a decision support tool that can assist caregivers and practitioners in identifying toddlers who may benefit from further evaluation.

***Comparison with Recent Literature***

Compared to recent machine learning–based ASD screening studies, the proposed approach demonstrates that simple and interpretable models can achieve competitive performance when applied to well-structured behavioral datasets. While many recent works explore ensemble learning or explainable AI frameworks (Aldrees et al., 2024; Abbas et al., 2024), the findings of this study suggest that KNN remains a viable alternative, particularly for applied screening scenarios where transparency and ease of implementation are important.

Moreover, the strong alignment between feature correlation patterns and model performance supports emerging evidence that compact and behavior-focused feature sets are sufficient for effective early ASD screening (Sollis et al., 2025). This reinforces the practical value of questionnaire-based screening data when combined with appropriately chosen machine learning techniques.

***Limitations and Implications for Future Work***

Despite its strong performance, this study has several limitations. First, the dataset used is a publicly available screening dataset, which may not fully capture population diversity across different geographic or cultural contexts. Second, the model was evaluated using a single classifier; future studies could explore comparative evaluations with other algorithms or hybrid approaches to further validate performance robustness.

Nevertheless, the findings provide important implications for future research and practice. Integrating the proposed KNN-based model into web-based or mobile screening systems could enhance accessibility and scalability of early ASD screening support. Additionally, combining questionnaire data with other data modalities, such as observational or digital behavioral features, may further improve detection capability.

## CONCLUSIONS AND RECOMMENDATIONS

This study presented a K-Nearest Neighbor (KNN)–based model for predicting autism spectrum disorder (ASD) traits in toddlers using questionnaire-derived behavioral screening data and selected demographic features. The proposed approach was designed to support early ASD screening by leveraging similarity-based classification on behavior-focused indicators. Experimental results demonstrate that the model achieved high predictive performance, with an accuracy of 95.73%, a precision of 100%, a recall of 93.66%, and an F1-score of 96.72%. The confusion matrix analysis further revealed that the model produced no false positive predictions, indicating strong reliability in avoiding incorrect ASD risk labeling among non-ASD toddlers. These findings highlight the effectiveness of KNN when applied to behavioral screening datasets where feature correlations are dominated by questionnaire-based indicators. The feature correlation analysis confirmed that behavioral screening items (A1–A10) and the aggregated Qchat-10-Score play a central role in distinguishing ASD-positive and non-ASD cases, while demographic variables contribute comparatively less predictive value. This alignment between feature relevance and model behavior reinforces the suitability of similarity-based learning for early ASD screening tasks. Overall, the results suggest that a simple and interpretable KNN-based model can serve as a reliable early screening decision support tool, assisting caregivers and practitioners in identifying toddlers who may benefit from further clinical assessment. Importantly, the proposed model is intended to complement, not replace, professional diagnostic evaluation.

## FURTHER STUDY

Future research may extend this work in several directions. First, model robustness can be further examined by evaluating performance across more diverse and representative datasets, including data from different geographic and cultural contexts. Second, comparative studies involving additional machine learning algorithms or hybrid models may provide deeper insight into performance trade-offs between simplicity and predictive power. Furthermore, integrating the proposed model into web-based or mobile screening applications could enhance accessibility and scalability in real-world settings. Finally, combining questionnaire-based screening data with other modalities, such as observational or digital behavioral features, may further improve early ASD risk identification and support more comprehensive screening frameworks.

## REFERENCES

Abbas, R. T., Sultan, K., Sheraz, M., & Chuah, T. C. (2024). A comparative analysis of automated machine learning tools: A use case for autism spectrum disorder detection. *Information, 15*(10), 625. https://doi.org/10.3390/info15100625

Aldrees, A., Ojo, S., Wanliss, J., Umer, M., Khan, M. A., Alabdullah, B., Alsubai, S., & Innab, N. (2024). Data-centric automated approach to predict autism spectrum disorder based on selective features and explainable artificial intelligence. *Frontiers in Computational Neuroscience, 18*, 1489463. https://doi.org/10.3389/fncom.2024.1489463

Alzakari, S. A., Alotaibi, F. S., Alshammari, T. S., & Alharbi, R. M. (2025). Early detection of autism spectrum disorder using screening datasets and feature engineering approaches. *Journal of Neuroscience Methods, 390*, 110074. https://doi.org/10.1016/j.jneumeth.2024.110074.

Allison, C., Matthews, F. E., Ruta, L., Pasco, G., Soufer, R., Brayne, C., Charman, T., & Baron-Cohen, S. (2021). Quantitative Checklist for Autism in Toddlers (Q-CHAT): A population screening study with follow-up: The case for multiple time-point screening for autism. *BMJ Paediatrics*

*Open, 5*(1), e000700. https://doi.org/10.1136/bmjpo-2020-000700

Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ, 385*, e078378. https://doi.org/10.1136/bmj-2023-078378

Halder, R. K., Uddin, M. N., Uddin, M. A., Aryal, S., & Khraisat, A. (2024). Enhancing k-nearest neighbor algorithm: A comprehensive review and performance analysis of modifications. *Journal of Big Data, 11*, 113. https://doi.org/10.1186/s40537-024-00973-y

Hyman, S. L., Levy, S. E., & Myers, S. M. (2020). Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics, 145*(1), e20193447. https://doi.org/10.1542/peds.2019-3447

Jullien, S. (2021). Screening for autistic spectrum disorder in early childhood. *BMC Pediatrics, 21*(Suppl 1), 349. https://doi.org/10.1186/s12887-021-02700-5

Lu, H., Zhang, Y., Li, X., & Wang, S. (2024). A machine learning–based screening model for early autism detection using questionnaire data. *Biomedical Signal Processing and Control, 92*, 105933. https://doi.org/10.1016/j.bspc.2024.105933

Miller, C., Johnson, R., & Patel, S. (2024). Evaluation metrics and performance assessment for machine learning models in healthcare applications. *Artificial Intelligence in Medicine, 147*, 102775. https://doi.org/10.1016/j.artmed.2024.102775

Moons, K. G. M., Wolff, R. F., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., Reitsma, J. B., Kleijnen, J., & Mallett, S. (2025). PROBAST+AI: An updated tool to assess risk of bias and applicability of prediction models that use regression or artificial intelligence. *BMJ, 388*, e082505. https://doi.org/10.1136/bmj-2024-082505

Perochon, S., Di Martino, J. M., Carpenter, K. L. H., Dawson, G., & Nelson, C. A. (2023). Early detection of autism using digital behavioral phenotyping. *Nature Medicine, 29*(10), 2489–2497. https://doi.org/10.1038/s41591-023-02574-3

Rainio, O., Teuho, J., & Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. *Scientific Reports, 14*, 56706. https://doi.org/10.1038/s41598-024-56706-x

Sollis, L. J., Wall, D. P., & Washington, P. Y. (2025). Compact subsets of autism screening items predict clinical diagnoses with a machine learning analysis of the QCHAT-10. *Scientific Reports, 15*, 39091. https://doi.org/10.1038/s41598-025-26131-9

Solmi, M., Song, M., Yon, D. K., Popova, E., Carvalho, A. F., & Fusar-Poli, P. (2022). Incidence, prevalence, and global burden of autism spectrum disorder from 1990 to 2019 across 204 countries. *Molecular Psychiatry, 27*(10), 4172–4180. https://doi.org/10.1038/s41380-022-01630-7

Tartarisco, G., Billeci, L., Calderoni, S., & Muratori, F. (2021). Machine learning–based screening of autism spectrum disorder using behavioral questionnaire data. *Journal of Biomedical Informatics, 118*, 103797. https://doi.org/10.1016/j.jbi.2021.103797

Vega-Huerta, H., Hernández-García, A., & García-Pedrajas, N. (2025). Improving k-nearest neighbor classification through Min–Max normalization and distance-based optimization. *Applied Sciences, 15*(18), 10202. https://doi.org/10.3390/app151810202

World Health Organization. (2025). *Autism spectrum disorders*. Retrieved December 21, 2025, https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders