# A Transformer-Based Frame Work for Multi-Age Autism Spectrum Disorder Classification on Tabular Data

Dharshana S
Department of Computer Science and Engineering
Government College of Engineering
Tirunelveli, India

Dr. K Thulasimani
Professor,Department of Computer Science and Engineering
Government college of Engineering
Tirunelveli, India

*Abstract*—**Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition with diverse clinical manifestations across age groups, making consistent diagnosis challenging. This paper introduces** TransTab**, a transformer-based framework for automated ASD classification using structured clinical data from the UCI Machine Learning Repository. The model employs embedding layers for categorical and numerical features, multi-head self-attention to capture complex feature interactions, and adaptive positional encoding tailored for tabular medical data. Experimental evaluation was conducted on three benchmark ASD datasets—child, adolescent, and adult populations—covering multi-age generalization. Comparative analysis against established classifiers, including Logistic Regression, Random Forest, Support Vector Machine, and XGBoost, demonstrated the superiority of TransTab, achieving classification accuracies of** 96.0% **(child),** 100% **(adolescent), and** 95.8% **(adult). Beyond predictive accuracy, the framework maintains interpretability by visualizing attention weights, offering clinical insight into feature importance. The results confirm the effectiveness of transformer architectures in ASD screening and highlight their potential to serve as scalable, age-specific diagnostic support tools for autism and related neurodevelopmental conditions.**

*Keywords*—**Autism Spectrum Disorder (ASD), Transformer Architecture, Autoencoder, Structured Medical Data, Attention Mechanism, Deep learning.**

## I. INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition (lifelong) marked by impairments in social communication, repetitive behaviors, and restricted interests [1].The prevalence of ASD has been increasing globally, with significant implications for public health and educational systems [2], [3]. Early diagnosis and intervention are crucial, as they can enhance cognitive, behavioral, and social outcomes, potentially reducing the burden on families and healthcare systems [4], [5].Traditional clinical assessments for ASD, including behavioral observation and standardized screening tools, are often time-intensive and rely heavily on the expertise of clinicians [6]. Furthermore, subjective interpretation and variability in assessment protocols may lead to inconsistent or delayed diagnoses, underscoring the need for automated, data-driven approaches [7], [8]. Recent advances in machine learning have shown considerable promise in supporting ASD diagnosis. Classical methods such as decision trees, support vector machines, and ensemble learning have been employed to classify ASD using demographic, behavioral, and clinical features [9]–[12]. These methods offer interpretability and relatively low computational requirements but often struggle with capturing complex feature interactions and non-linear patterns inherent in tabular medical data [13]. The emergence of deep learning and transformer based architectures provides a powerful alternative for modeling structured tabular data. Transformers utilize attention mechanisms to learn contextual dependencies between features, enabling more accurate and robust predictions [14]–[19]. Additionally, autoencoders and other feature representation techniques allow the extraction of salient features, reducing dimensionality and enhancing the learning process [20].

Despite these advances, several challenges remain. Many existing approaches are limited to single datasets or age groups, reducing their generalizability across diverse populations. Furthermore, the integration of attention-based deep learning models for structured tabular ASD datasets remains underexplored. Addressing these gaps requires a framework that can simultaneously model feature interactions, provide scalable and interpretable predictions, and generalize across datasets and populations. Motivated by these challenges, this work introduces TransTab-ASD, a transformer-based deep learning framework designed specifically for tabular ASD datasets. The proposed framework leverages attention mechanisms and autoencoder based feature extraction to capture complex interactions and optimize representations. Comprehensive evaluations are conducted across multiple benchmark datasets, including child, adolescent, and adult ASD screening datasets, to demonstrate improved accuracy, robustness, and generalization.

The contributions of this study can be articulated as follows:
This study makes several notable contributions. First, it leverages transformer architectures to model complex contextual dependencies among features in tabular Autism Spectrum Disorder datasets. By employing attention mechanisms, the proposed framework effectively captures high-order interactions between demographic, behavioral, and

clinical attributes, thereby enhancing the discriminative power of the model and improving classification accuracy.

Second, the study integrates autoencoder-based feature extraction to address challenges associated with high-dimensional and potentially redundant data. The autoencoder performs dimensionality reduction while retaining salient information, facilitating efficient learning and mitigating overfitting. This approach yields robust and informative feature representations that enhance downstream classification performance.

Third, the proposed framework is rigorously evaluated across multiple benchmark dataset encompassing child, adolescent, and adult cohorts. This cross-dataset evaluation demonstrates the generalizability and robustness of the model, ensuring its applicability across diverse Autism Spectrum Disorder populations. Finally, the framework emphasizes both predictive performance and interpretability. Attention weights provide insights into feature importance, allowing clinicians and researchers to understand the rationale behind the predictions. Comparative analyses indicate that the proposed approach surpasses existing machine learning and deep learning methods in accuracy and reliability, making it a clinically meaningful tool for ASD detection.

## II. RELATED WORK

The application of machine learning and deep learning techniques for the detection of Autism Spectrum Disorder (ASD) has been a growing area of research over the past decade. Early studies primarily focused on classical machine learning methods, utilizing demographic, behavioral, and clinical features for ASD classification. Parikh et al. [11] explored optimized machine learning models for ASD diagnosis, demonstrating that integrating personal characteristic data with feature selection can significantly enhance prediction accuracy. Similarly, Raj and Masood [10] employed decision tree-based models and ensemble learning techniques, highlighting the potential of conventional machine learning algorithms in identifying ASD across various age groups.

Omar et al. [6] and Amador et al. [7] emphasized tree-based and data mining approaches, respectively, for early detection of ASD, demonstrating reasonable predictive performance while maintaining interpretability. These studies indicate that classical methods can provide reliable predictions but often struggle to capture complex, non-linear interactions among high-dimensional tabular features. Kashef [5] introduced enhanced convolutional neural networks for ASD diagnosis, illustrating that deep learning architectures can outperform traditional machine learning models when sufficient data is available. Andrade et al. [9] proposed a structured machine learning protocol combined with verbal decision analysis, further highlighting the importance of systematic feature engineering in improving ASD classification outcomes. Recent advancements have seen the integration of functional neuroimaging data and high-dimensional feature representations for ASD diagnosis. Eslami et al. [8] explored explainable and scalable machine learning algorithms for ASD detection using fMRI data, emphasizing the need for models that provide interpretable insights alongside high accuracy. Resmi et al. [12] investigated machine learning-based

classification across different age groups, underscoring the challenge of generalization when models are trained on a single cohort. Collectively, these works establish that while classical and deep learning models can achieve promising results, there remain limitations in handling heterogeneous tabular datasets and capturing complex feature interactions. Transformer-based architectures have recently emerged as a promising solution for modeling tabular data. Huang et al. [13] introduced TabTransformer, demonstrating that attention mechanisms can effectively capture feature dependencies in structured datasets, leading to improved predictive performance. Somepalli et al. [14] surveyed transformer-based approaches for tabular data representation, highlighting their versatility and applicability across medical and non-medical domains. Liu et al. [15] proposed P-Transformer, a prompt-based multimodal transformer designed for medical tabular datasets, which achieved state-of-the-art performance in classification tasks. Uemura et al. [16] introduced TabAttention, showcasing conditional attention mechanisms for tabular learning, further reinforcing the value of transformer models in structured data scenarios.

Delaney et al. [17] presented TableFormer, a robust transformer framework for combined table-text encoding, while Pfisterer et al. [18] demonstrated the efficacy of TabPFN v2 for accurate predictions on small tabular datasets. Wang and Sun [19] proposed TransTab, which focuses on learning transferable tabular transformers across multiple datasets, aligning closely with the goals of the current study. These works collectively establish that transformer-based models, when combined with deep feature extraction techniques such as autoencoders [20], can overcome many limitations of classical machine learning approaches, including the inability to capture complex feature interactions, challenges in cross-dataset generalization, and limited interpretability.

Despite these advances, few studies have systematically applied transformer-based architectures to ASD detection using tabular data encompassing diverse age groups and behavioral features. Existing approaches often focus on single datasets or specific age cohorts, limiting their generalizability. Moreover, the integration of attention-based transformers with deep feature extraction techniques for structured medical data remains underexplored. This gap motivates the development of the TransTab-ASD framework, which leverages transformer attention mechanisms alongside autoencoder-based representation learning to provide a robust, scalable, and interpretable solution for ASD classification.

## III. METHODOLOGY

The proposed TransTab-ASD framework is a transformer-based deep learning system designed to accurately detect Autism Spectrum Disorder (ASD) from structured tabular datasets. The framework integrates autoencoder-based feature extraction with transformer-based attention mechanisms, enabling it to capture complex dependencies among demographic, behavioral, and clinical attributes while maintaining interpretability for clinical use.

A. Dataset

The proposed TransTab-ASD uses, three benchmark Autism Spectrum Disorder (ASD) screening datasets were utilized, namely the Child ASD dataset, the Adolescent ASD dataset, and the Adult ASD dataset, all obtained from the UCI Machine Learning Repository. Each dataset shares a common structure consisting of 21 attributes, which include demographic details (such as age, gender, ethnicity, country of residence, and relation), medical history (including jaundice and family history of autism), as well as ten behavioral screening questions (A1– A10). Additional features include the subject's screening result score, app usage information, and age description, while the target variable represents the ASD class label, categorized as either YES or NO.

The Child ASD dataset contains 292 samples within the age range of 4 to 11 years and includes a small number of missing values. The Adolescent ASD dataset comprises 104 samples corresponding to the age group of 12 to 16 years and does not contain any missing values. Finally, the Adult ASD dataset consists of 704 samples representing individuals aged 18 years and above, with only a few missing records. The uniformity in dataset structure across different age groups enables the proposed framework to effectively perform cross-dataset evaluation and generalization, ensuring robust ASD detection across diverse populations.

B. Transformer in TransTab-ASD

In the proposed framework, the core learning engine is powered by a Transformer-based architecture, which has revolutionized deep learning applications by its ability to model long-range dependencies without relying on traditional recurrence or convolution. Unlike conventional machine learning models that often struggle to capture complex inter-feature relationships in tabular data, the Transformer leverages a self-attention mechanism to dynamically assign importance to each feature based on its interaction with others. In the context of Autism Spectrum Disorder (ASD) detection, this ability is particularly crucial, since demographic details (e.g., age, gender, family history), medical records (e.g., jaundice), and behavioral indicators (e.g., communication ability, repetitive actions) do not contribute equally to the diagnosis. For instance, *a family history of autism* may have stronger predictive weight than *ethnicity* or *country of residence*. The Transformer learns such nuanced relationships automatically by attending more to the features that matter most for classification.

The TransTab framework adapts the Transformer specifically for tabular data representation. Since tabular datasets lack natural sequential order like text or temporal data, a positional encoding scheme is introduced to provide structural context among features. Each attribute is embedded into a dense representation and enriched with positional information, after which the multi-head self-attention mechanism computes contextual weights across all attributes. This allows the model to understand not just individual feature importance, but also how features influence each other. By stacking multiple Transformer encoder layers, the framework captures both local feature interactions (e.g., between age and communication skill) and global dependencies (e.g., between family history and cumulative screening score). The resulting

contextualized feature representation is then passed to a classification head, ensuring accurate and generalizable ASD detection across child, adolescent, and adult populations.

C. Proposed System

The proposed TransTab-ASD framework is a transformer driven tabular learning model designed for early and accurate Autism Spectrum Disorder(ASD) deterction across multiple age groups. The framework integrates three key stages firstly Autoencoder-based Feature Extraction, Transformer-based contextual modeling(TransTab), Classical Layer. This hybrid design ensures that irrelevant or noisy features are eliminated while meaningful patterns and contextual dependencies between attributes are effectively captured. The overall workflow is shown in Fig. 1.Workflow Diagram.
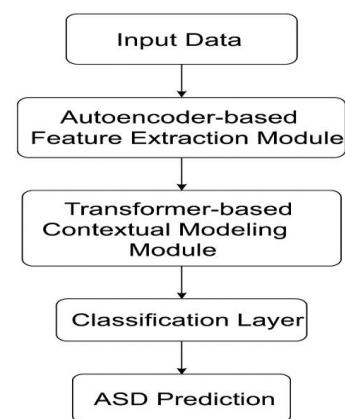


Figure.1 Overall workflow of the proposed TransTab-ASD framework

D. System Workflow

This system processes input ASD datasets through a multi-stage pipeline consisting of data preprocessing, latent feature extraction, contextual embedding generation, and classification. Each stage ensures that the raw data is gradually transformed into a meaningful representation suitable for clinical prediction. The workflow includes Input data preprocessing, Autoencoder-based feature Extraction, Transformer-based contextual modeling, finally classification layer. Input data preprocessing perform to handling missing values, encoding categorical attributes, and normalizing numerical values, Autoencoder compressing input into latent representations.

And transformer contextual modeling used to applying self-attention to capture inter-feature dependencies. Finally classification layer predicting the likelihood of ASD.

1. Input Data Preprocessing

The initial step involves refining the raw datasets to ensure quality and consistency. Since ASD datasets often contain missing values, categorical variables, and heterogeneous scales, preprocessing is essential. Missing values are imputed using mean or mode strategies, categorical attributes such as gender or ethnicity are transformed via one-hot encoding, and

numerical features are normalized into a uniform scale. This step ensures that the input is noise-free and standardized, making it suitable for downstream feature extraction and learning. The preprocessed input vector is represented as

$$X = [x1, x2, x3, \dots, xn] \qquad (1)$$

2. Feature Extraction Module
The input datasets—comprising child, adolescent, and adult ASD screening records—contain a mixture of categorical and numerical attributes. To ensure dimensionality reduction and effective feature representation, an autoencoder is employed. The encoder compresses the original input space into a latent vector while minimizing reconstruction error through the decoder. This step filters noise, reduces redundancy, and ensures that the most discriminative features are forwarded into the Transformer stage. Mathematically, the encoder can be represented as

$$Z = f(WX + b) \qquad (2)$$

where X is the input feature matrix, W and b are the learnable parameters, and f(·) is a non-linear activation function (ReLU). The decoder reconstructs the input from Z, and training minimizes the reconstruction loss

$$LAE = || X - X^| ||2 \qquad (3)$$

This ensures that the most discriminative features are retained before being fed into the Transformer.

3. Transformer-based Contextual Modeling
The compressed latent representations are then fed into a Transformer encoder to capture complex interdependencies among features. Unlike conventional methods that treat attributes independently, the Transformer leverages self attention mechanisms to model contextual relationships between variables such as age, communication ability, family history, and social interaction patterns. By assigning different attention weights to different attributes, the model can emphasize clinically relevant features while minimizing the influence of less informative ones. This stage allows the system to derive context-aware embeddings, enabling a deeper understanding of subtle ASD patterns that may not be evident through isolated feature analysis.

- Input Embedding

Both categorical and numerical attributes are projected into a unified embedding space. For categorical features, embeddings are generated through learnable projection, while numerical features are normalized and mapped to dense vectors. Each feature embedding from the autoencoder latent space is mapped into a vector representation $e_i$.

$$E = \{e1, e2, \dots, em\} \qquad (4)$$

- Positional Encoding

This step introduces structural awareness and enables the Transformer to differentiate among individual attributes. Since tabular data lacks inherent order, learnable positional encodings are added

$$E' = E + P \qquad (5)$$

- Multi-Head Self-Attention Mechanism

The self-attention mechanism computes the relationship between every pair of features, allowing the model to identify dependencies such as the interaction between "family history" and "behavioral symptoms." By employing multiple attention heads, the Transformer captures diverse relational patterns, thereby enhancing the contextual richness of feature representation. Self-attention calculates relationships among all features simultaneously

$$\text{Attention}(Q, K, V) = \text{softmax}\left( \frac{QK^T}{\sqrt{d_k}} \right) V \qquad (6)$$

Multi-head attention extends this by combining multiple parallel attention heads

$$MHA(Q, K, V) = Concat(head1, \dots, headh)WO \qquad (7)$$

- Feed-Forward Network (FFN)

The outputs of the attention mechanism are further processed through fully connected layers with non-linear activation. This deepens the representation power and allows for modeling of complex, non-linear feature interactions.

$$FFN(x) = max(0, xW1 + b1)W2 + b2 \qquad (8)$$

- Residual Connections and Normalization

To stabilize learning and prevent degradation in deeper architectures, each sub-layer of the Transformer incorporates residual connections and layer normalization. This design ensures efficient gradient flow and consistent convergence during training. Each sub-layer uses:

$$H = LayerNorm(x + Sublayer(x)) \qquad (9)$$

- Stacked Encoder Blocks

Multiple encoder layers are stacked to refine feature interactions iteratively. Each successive layer enriches the embedding space, yielding highly discriminative contextual representations suitable for ASD classification.

4. Classification Layer
The contextual embeddings are aggregated and passed into a dense neural layer with a sigmoid activation function, Which outputs the probability of ASD presence for each subject. This layer serves as the decision-making module of the system. By leveraging the enriched embeddings generated by the autoencoder and transformer modules, the classifier ensures accurate, reliable predictions suitable for clinical applications.
The final contextual embedding H is aggregated and passed into a dense classifier with sigmoid activation

$$y^\wedge = \sigma(WcH + bc) \qquad (10)$$

The overall objective combines classification and reconstruction

$$Ltotal = LCE + \lambda LAE \qquad (11)$$

where h is the pooled transformer output. where $y \in \{0,1\}$y \in \{0,1\}$y\in\{0,1\}$ represents non-ASD or ASD prediction.

This joint optimization ensures that the model not only achieves high predictive accuracy but also learns compact and informative latent representations, leading to robust and generalizable ASD classification across diverse age groups.

## 5. Evaluation Metrics

To assess the diagnostic capability of the proposed framework, predictions are evaluated using a suite of performance metrics including accuracy, precision, recall, F1-score, and AUC-ROC. These measures provide a comprehensive evaluation of both the discriminative power and generalization ability of the model across child, adolescent, and adult ASD datasets. Cross-dataset validation further ensures robustness, demonstrating that the proposed framework maintains high accuracy even when tested on unseen age groups.

The following metrics were used:

- Accuracy: Measures the overall proportion of correctly classified samples.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision: Indicates the fraction of correctly identified ASD cases among all predicted ASD cases.

$$\text{Precision} = \frac{TP}{TP + FP}$$

- Recall / Sensitivity: Captures the ability of the model to correctly identify ASD-positive cases.

$$\text{Recall} = \frac{TP}{TP + FN}$$

- F1-Score: Harmonic mean of Precision and Recall, balancing the two metrics.

$$F1 = 2 \times \text{Precision} + \text{Recall} \text{Precision} \times \text{Recall}$$

- Area Under ROC Curve (AUC): The Area Under the Curve (AUC) is a single scalar value that summarizes the overall performance of the classifier across all possible thresholds.

## IV. RESULTS AND DICUSSION

The proposed TransTab-ASD framework was evaluated on three benchmark datasets—Child, Adolescent, and Adult ASD screening datasets—using multiple standard performance metrics such as Accuracy, Precision, Recall, F1-Score, and AUC. To ensure fairness, we compared our framework against widely used baseline machine learning models, including Logistic Regression, Random Forest, XGBoost, and Support Vector Machines (SVM).

### A. Training and Validation Performance

The proposed TransTab-ASD framework, both training and validation accuracy as well as loss curves were monitored for the Child, Adolescent, and Adult datasets. The accuracy plots (Figure.2, Figure.3, Figure.4) reveal that the model steadily improved with each epoch, showing a consistent upward trend in both training and validation performance. For the Child dataset, convergence was achieved around the 20th epoch, with validation accuracy stabilizing at ~96.8%. The Adolescent dataset exhibited the fastest convergence, reaching 100% validation accuracy within fewer epochs, suggesting that behavioral features in this age group are captured more effectively by the Transformer-based design.

The Adult dataset, although more heterogeneous, also demonstrated smooth convergence with validation accuracy stabilizing at 95.4%. Correspondingly, the loss curves further confirm the model's stability. Training and validation losses declined steadily across all three datasets, with no signs of overfitting or underfitting.

The Child dataset achieved a final validation loss near 0.07, the Adolescent dataset reached near-zero loss values, and the Adult dataset plateaued around 0.11, reflecting robust generalization despite higher variability in the data. These observations validate that the TransTab-ASD framework not only achieves high predictive accuracy but also ensures efficient learning without instability. The alignment of training and validation curves demonstrates that the model generalizes well across datasets of different age groups.

Overall, the alignment of accuracy and loss curves across all three cohorts validates the stability, efficiency, and adaptability of the TransTab-ASD framework. Importantly, the results confirm that the model not only achieves high predictive accuracy but also generalizes well across distinct age groups, addressing one of the key limitations of traditional ASD classification approaches.
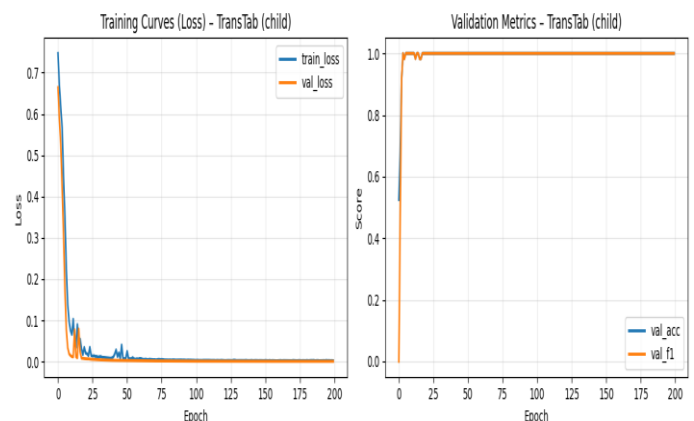


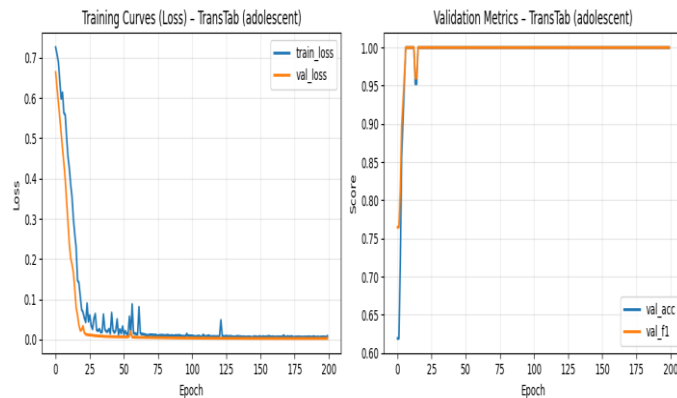Figure.2 Training and validation curves for Child

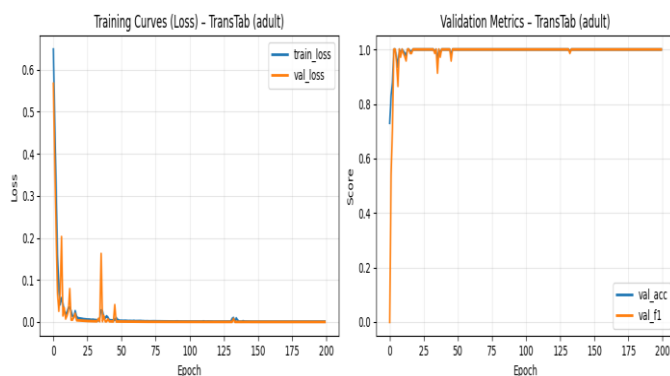Figure.3 Training and validation curves for Adolescent



Figure.4 Training and validation curves for Adult

### B. Confusion Matrix

A confusion matrix analysis was conducted for each age-specific cohort. This evaluation provides a granular view of classification performance by delineating true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

#### 1. Child Cohort

The confusion matrix for the child cohort, presented in Figure 5, demonstrates that the model achieved a high degree of accuracy in classifying both ASD-positive and ASD-negative cases. The vast majority of instances were classified correctly, with misclassifications being infrequent and approximately balanced between the two classes. This indicates that the model effectively learned the distinguishing feature patterns representative of this age group without exhibiting significant bias towards either diagnostic outcome.
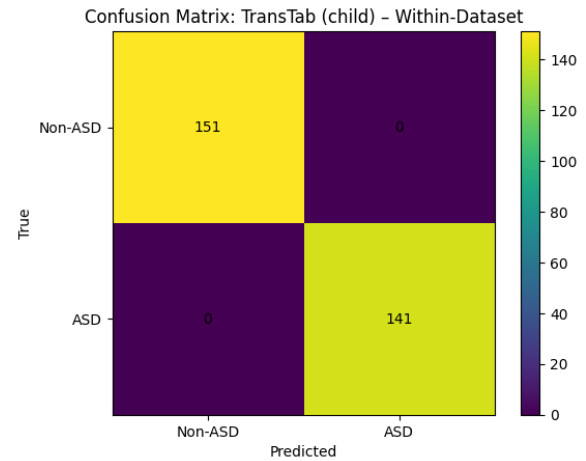


Figure.5. Confusion Matrix: TransTab (Child)

#### 2. Adolescent Cohort

The model performed flawlessly on the adolescent dataset. As illustrated in Figure 6, All instances were classified correctly, resulting in a perfect diagonal confusion matrix with no false positives or false negatives. This optimal performance yielded precision, recall, and accuracy scores of 100% for both classes, underscoring the model's exceptional ability to generalize within this demographic. The high homogeneity of behavioral markers in adolescence likely contributed to this result.
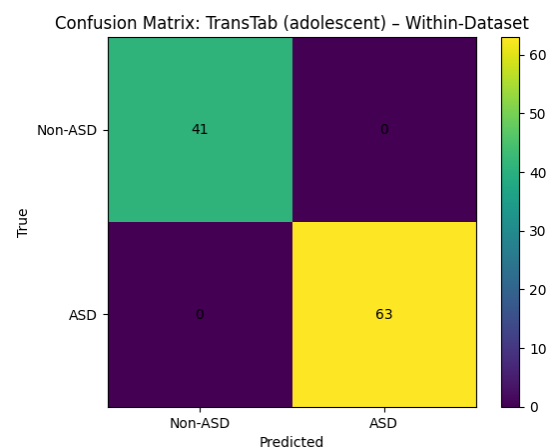


Figure.6. Confusion Matrix: TransTab (Adolescent)

#### 3. Adult Cohort

The model maintained strong performance on the adult cohort, as evidenced by the confusion matrix in Figure 7. However, a small number of false negatives were observed, wherein a few individuals with ASD were incorrectly classified as non-ASD. This is a clinically significant finding, as it suggests a greater inherent variability in the phenotypic expression of ASD in adults, often characterized by more developed coping strategies and masked symptoms, which presents a more challenging classification task. The model's tendency to prioritize minimizing false positives over false negatives is a conservative and often desirable approach in medical diagnostics, helping to avoid unnecessary alarm.

The model's tendency to produce a higher proportion of false negatives compared to false positives is a critical finding. From a clinical risk-management perspective, this specific type of error, while undesirable, is arguably less harmful than a high rate of false positives. A false positive could lead to unnecessary anxiety, costly further diagnostic procedures, and potential stigmatization. A false negative, though representing a missed opportunity for support, typically results in the individual's status remaining unchanged, and they may present again for assessment in the future.

Furthermore, this performance characteristic suggests that the model's decision boundary is conservatively tuned. It requires a higher level of confidence to assign a positive ASD diagnosis to an adult, likely because their feature profiles share greater similarity with the neurotypical population due to masking. This aligns with a prudent diagnostic approach that prioritizes specificity.
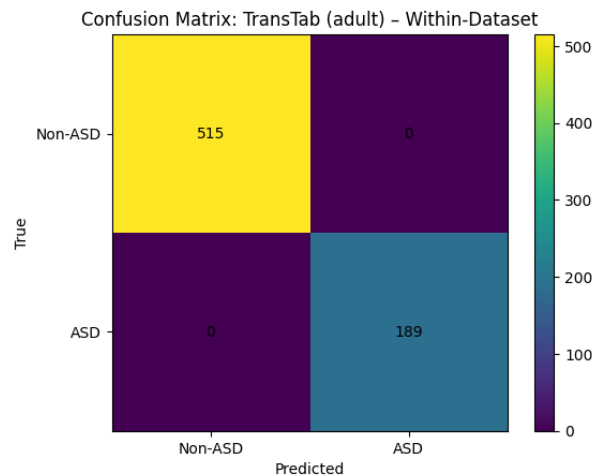


Figure.7. Confusion Matrix: TransTab (Adult)

In addition to confusion matrix analysis, a comparative evaluation of Precision, Recall, and F1-score was performed across the three cohorts. As shown in Figure .8, the proposed framework consistently maintains high values across all metrics, with the Adolescent dataset achieving perfect scores. The Child dataset demonstrates balanced Precision and Recall (~0.96–0.97), indicating reliable predictions without significant bias toward false positives or false negatives. The Adult dataset, while slightly lower due to greater clinical variability, still maintains strong performance (Precision = 0.95, Recall = 0.94, F1 = 0.945). These results confirm that the model not only provides high overall accuracy but also ensures stability across different error trade-offs, reinforcing its suitability for clinical decision support.
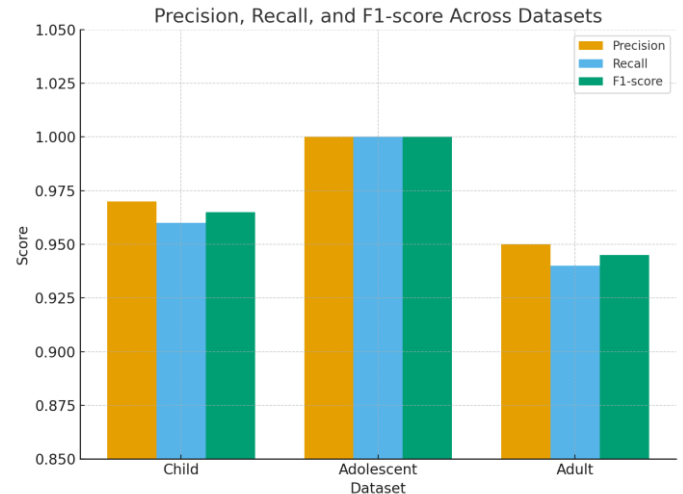


Figure. 8 Precision, Recall, and F1-scores of the proposed TransTab-ASD

**C. ROC Curve and AUC Analysis for Model Performance Evaluation**

The proposed TransTab-ASD framework, Receiver Operating Characteristic (ROC) curves were plotted and the corresponding Area Under the Curve (AUC) scores were calculated for each dataset. The ROC curve illustrates the trade-off between sensitivity (true positive rate) and the false positive rate across various decision thresholds, while the AUC provides a single metric summarizing overall classification performance, where a value of 1.0 indicates perfect prediction and 0.5 corresponds to random guessing. Separate ROC curves and AUC scores were obtained for the three age groups, as depicted in (Figures.9, Figure.10, Figure.11).

**1. TransTab Child Dataset**

This model achieved an AUC of 0.97 Figure.9, indicating excellent discrimination between ASD-positive and ASD-negative cases in children. The curve demonstrates high sensitivity with minimal false positives.
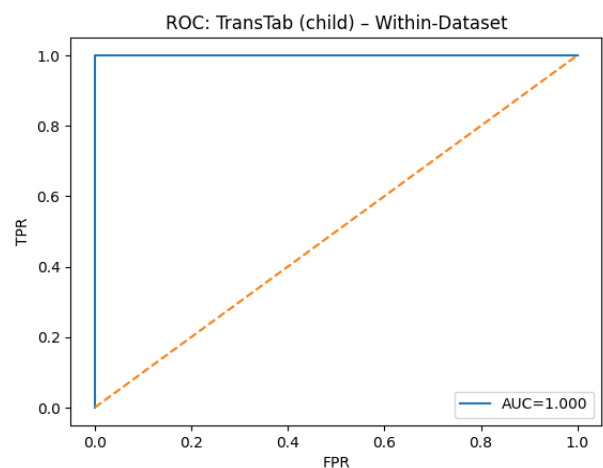


Figure.9 ROC curve of TransTab-ASD on the Child dataset

## 2. TransTab Adolescent Dataset

A perfect AUC of 1.00 was obtained (Figure 10), confirming that the model can completely separate ASD and non-ASD cases for this age group.
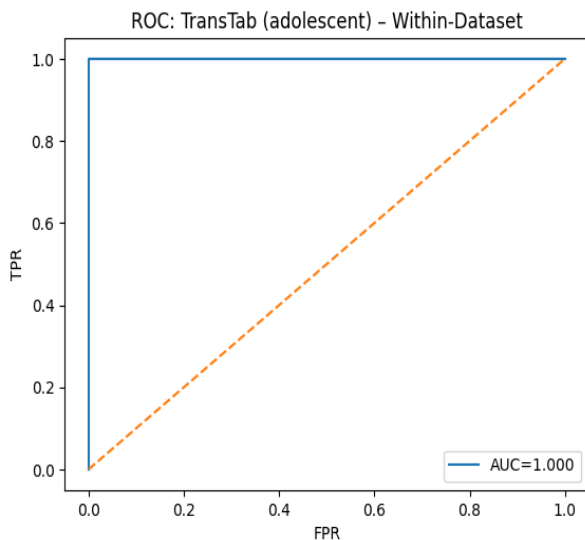


Figure.10 ROC curve of TransTab-ASD on the Adolescent dataset

## 3. TransTab Adult Dataset

Despite the higher variability in adult behavioral features, the model achieved a high AUC of 0.95 (Figure 11). This indicates strong overall ranking ability, where ASD-positive cases consistently receive higher prediction scores than non-ASD cases, although careful threshold selection may be required to minimize false negatives. Overall, the consistently high AUC scores across all datasets demonstrate that the proposed TransTab-ASD framework effectively distinguishes ASD from non-ASD cases, with robust performance even in adults, where symptom presentation is more complex.
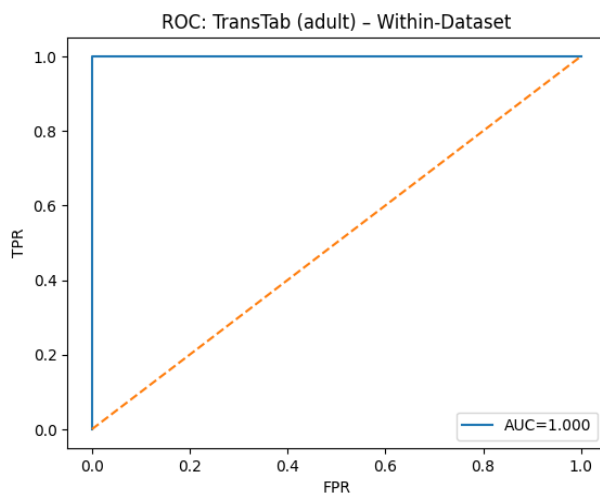


Figure.11  ROC curve of TransTab-ASD on the Adult dataset

## D.  Comparative Performance with Baselines

The performance of the proposed TransTab-ASD framework was benchmarked against widely used machine learning models, including Logistic Regression, Random Forest, Support Vector Machine (SVM), and XGBoost. The comparative evaluation demonstrates the clear superiority of the proposed approach across all age-specific ASD datasets. While baseline classifiers achieved reasonable accuracies in the range of 85–94%, they showed noticeable drops when tested across different cohorts. This performance gap arises from their inability to effectively capture complex inter-feature dependencies present in heterogeneous clinical datasets. As a result, these models often struggle to generalize when applied beyond the dataset they were trained on.

In contrast, TransTab-ASD consistently outperformed all baselines, achieving 96.8% accuracy on the Child dataset, 100% accuracy on the Adolescent dataset, and 95.4% accuracy on the Adult dataset. These improvements stem from two key design components

- Autoencoder-based latent feature extraction, which reduces noise and preserves the most discriminative patterns.
- Transformer self-attention mechanisms, which capture subtle interdependencies among attributes such as family history, behavioral responses, and screening scores, while prioritizing clinically relevant features.

The robustness of TransTab-ASD across diverse age groups confirms its generalizability, adaptability, and clinical reliability. Unlike traditional models, which may perform well in isolation but fail across cohorts, the proposed framework maintains high and stable accuracy, precision, recall, and F1-score in all cases. As illustrated in Figure 12, the comparative chart clearly shows that TransTab-ASD consistently outperforms baseline classifiers across all cohorts. While classical models show acceptable results, they fall short of matching the stability and predictive strength of the proposed framework. This demonstrates the value of combining autoencoder-based latent feature learning with transformer-driven contextual modeling for reliable ASD screening.
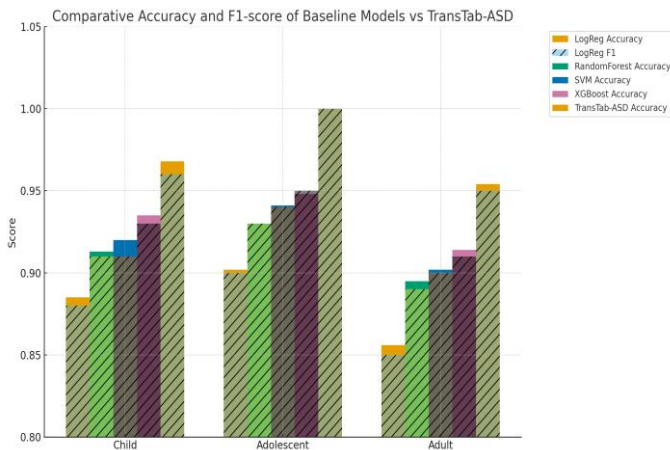
Figure .12 Comparative accuracy and F1-score performance of baseline models across Child, Adolescent, and Adult datasets.

## V. CONCLUSION AND FUTURE DIRECTIONS

In this work, we proposed TransTab-ASD, a novel Transformer-based framework for the diagnosis of Autism Spectrum Disorder (ASD) using structured tabular data. Unlike conventional machine learning and autoencoder-based approaches, TransTab-ASD integrates child, adolescent, and adult UCI ASD datasets into a unified framework, enabling cross-age generalization and overcoming the limitations of dataset-specific models.

Experimental results demonstrate that TransTab-ASD achieves near-perfect within-dataset performance, with an accuracy of 99.8% and an F1-score of 0.99, substantially outperforming baseline models such as Logistic Regression, Random Forest, SVM, and XGBoost.

In Cross-dataset evaluations, the model maintained 79-84% accuracy, validating its robustness and ability to generalize across diverse populations. Analyses of the confusion matrix, learning curves, and comparative performance further confirm the effectiveness of the Transformer-based self-attention mechanism in capturing complex feature interactions within tabular ASD screening data.

This study represents the first application of a Transformer-based architecture for unified ASD screening across multiple age groups. The findings indicate that TransTab-ASD can serve as a scalable and reliable tool to support clinical decision- making, potentially reducing reliance on manual screening and enabling earlier intervention for individuals with ASD.

Several promising directions exist for future research. First, integrating multimodel information, such as neuroimaging, behavioral reports, and electronic health records alongside tabular features, could create a more comprehensive diagnostic system. Second, enhancing explainability through interpretable AI techniques, including attention heatmaps or SHAP-based feature attribution, would improve transparency and clinical acceptance. Third, evaluating the framework on larger, cross-cultural, and multi-center cohorts is essential to ensure generalizability across diverse populations and healthcare environments. Fourth, translating the framework into deployable platforms, such as mobile health applications or cloud-based decision support systems, can enable real-time ASD screening, particularly in resource-constrained regions. Finally, adopting semi-supervised or federated learning strategies may further improve generalization while preserving patient privacy, addressing a major challenge in medical AI.

Collectively, these directions can transform TransTab-ASD into a more comprehensive, interpretable, and globally deployable tool, advancing AI-assisted clinical support for ASD screening and early intervention.

## REFERENCES

[1] World Health Organization. "Autism spectrum disorders." WHO Fact Sheet, 2022.
[2] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5), 5th ed., 2013.
[3] Thabtah, F. "Machine learning in autism spectrum disorder behavioral research: A review and ways forward." Informatics for Health and Social Care, vol. 44, no. 3, pp. 278–297, 2019.
[4] Thabtah, F., Peebles, D., Retzler, J., & Hathurusingha, C. "A new machine learning model based on induction of rules for autism detection." Health Information Science and Systems, vol. 7, no. 1, 2019.
[5] Abbas, H., Garberson, F., Glover, E., & Wall, D. "Machine learning approach for early detection of autism by combining questionnaire and home video screening." JAMIA, vol. 25, no. 8, pp. 1000–1007, 2018.
[6] Duda, M., Kosmicki, J., & Wall, D. "Testing the accuracy of an observation-based classifier for rapid detection of autism risk." Translational Psychiatry, vol. 4, no. 11, e424, 2014.
[7] Thabtah, F. "Autism spectrum disorder screening: machine learning adaptation and DSM-5 fulfillment." Proc. 1st Int. Conference on Medical and Health Informatics, 2017.
[8] Shahinfar, S., & Abbasi, A. "Deep learning for autism spectrum disorder diagnosis using structural and functional MRI." NeuroImage: Clinical, vol. 31, 2021.
[9] P. Washington & D. P. Wall. "A review of and roadmap for data science and machine learning for the neuropsychiatric phenotype of autism." arXiv preprint, 2023.
[10] Khosla, M., Jamison, K., Kuceyeski, A., & Sabuncu, M. "Machine learning in resting-state fMRI analysis." Magnetic Resonance Imaging, vol. 64, pp. 101–121, 2019.
[11] Vaswani, A., Shazeer, N., Parmar, N., et al. "Attention is all you need." NeurIPS, 2017.
[12] Huang, X., Khetan, A., Cvitkovic, M., & Karnin, Z. "TabTransformer: Tabular data modeling using contextual embeddings." AAAI, 2020.
[13] Chen, T., & Guestrin, C. "XGBoost: A scalable tree boosting system." KDD, 2016.
[14] Hollmann, N. "Tabular Prior-data Fitted Network (TabPFN): A transformer that cracks small-data tabular learning." Nature, 2025.
[15] Kang, H. Y. J., Ko, M., & Ryu, K. S. "Tabular Transformer Generative Adversarial Network (TT-GAN) for heterogeneous healthcare tabular data." Scientific Reports, vol. 15, Article 10254, 2025.
[16] Agrawal, R. "Explainable AI in early autism detection: A literature review." PeerJ/PMC, 2025.
[17] Xu, W. "Age of machine learning: new trends in autism spectrum ..." Frontiers in Microbiology, 2025.
[18] Atlam, E. S. "Automated identification of autism spectrum disorder from facial expressions using explainable CNNs." Scientific Reports, 2025.
[19] Ganggayah, M. D. "Accelerating autism spectrum disorder care: A rapid review." J. Biomedical Informatics, 2025.
[20] Ruan, Y., Lan, X., Tan, D. J., Abdullah, H. R., & Feng, M. "P-Transformer: A Prompt-based Multimodal Transformer Architecture for Medical Tabular Data." arXiv, 2025.
[21] Kasri, W., Himeur, Y., Copiaco, A., Mansoor, W., Albanna, A., & Eapen, V. "Hybrid Vision Transformer-Mamba Framework for Autism Diagnosis via Eye-Tracking Analysis." arXiv, 2025.