International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

A Medical Diagnosis System Based on Explainable Artificial Intelligence: Autism Spectrum Disorder Diagnosis

[1]J. Adilakshmi, [2]G Vinoda Reddy, [3]Krishan Dev Nidumolu, [4]Renzon Daniel Cosme Pecho [5]M. Jahir Pasha

Submitted: 12/02/2023 **Revised:** 11/04/2023 **Accepted**: 10/05/2023

Abstract: This paper introduces a new diagnostic system for Autism Spectrum Disorder (ASD) using explainable artificial intelligence (AI). The goal is to develop a reliable and interpretable tool that helps healthcare professionals accurately identify individuals with ASD. The study follows a systematic methodology involving comprehensive data collection, feature engineering, and advanced machine learning algorithms, such as decision trees and support vector machines. By analyzing various patient data, including behavioral observations and medical history, the system identifies important features and patterns associated with ASD. The diagnostic system achieves promising results, with the decision tree model achieving an accuracy of 85% and the support vector machine model achieving 86%. These outcomes demonstrate the potential of the system to accurately identify ASD cases. The clinical relevance and practical implications of the diagnostic system are discussed, emphasizing its ability to enhance the accuracy and efficiency of ASD diagnoses. The paper also identifies limitations and proposes future enhancements, including expanding datasets to cover a wider age range and demographic factors, incorporating additional relevant features such as genetic markers and neuroimaging data, exploring alternative machine learning algorithms, and further advancing explainable AI techniques. Real-world validation and feedback from clinicians and caregivers are crucial for refining the system. Ultimately, this research aims to contribute to timely interventions and improved outcomes for individuals with ASD, providing valuable insights for clinicians, caregivers, and researchers in addressing the challenges of ASD diagnosis.

Keywords*: Explainable AI, Medical Diagnosis, Autism Spectrum Disorder, Machine Learning, Interpretable Insights.*

1. Introduction

Autism Spectrum Disorder (ASD) is a prevalent neurodevelopmental disorder with significant implications for healthcare. Its diagnosis poses challenges due to the complex nature of the disorder and the heterogeneity of its symptoms. Accurate and timely diagnosis is crucial for improving patient outcomes and enabling early intervention strategies that can positively impact long-term outcomes. The diagnosis of Autism Spectrum Disorder (ASD)[1] presents significant challenges in healthcare due to the complex nature of the disorder and the limitations of traditional diagnostic approaches.

These approaches often rely on subjective clinical judgments, leading to inconsistencies and delays in accurate identification of individuals with ASD. The lack of reliable and interpretable diagnostic tools hampers the ability of healthcare professionals to make timely and accurate diagnoses, which in turn affects the implementation of appropriate intervention strategies. Therefore, there is a critical need to develop a diagnostic system based on explainable artificial intelligence (AI) [2] that can overcome the limitations of traditional approaches and provide healthcare professionals with a reliable and transparent tool for accurate ASD diagnosis. Such a system would enhance the accuracy and efficiency of ASD diagnoses, facilitating early intervention and improving patient outcomes. By leveraging the power of explainable AI, The development of an explainable AI-based system holds immense significance in improving the accuracy and transparency of ASD diagnoses, providing valuable support to healthcare professionals in their decision-making process.

Due to the complex nature of autism spectrum disorder (ASD) and limitations in conventional diagnostic approaches, healthcare professionals face significant challenges when it comes to making a diagnosis, and the reliance of these approaches on subjective clinical judgments can cause inconsistencies and delays in the

[1] *Assistant professor, Department of Information technology, MLR Institute of Technology, Telangana, India*
Email ID: bhulakshmi550@gmail.com
*[*2] Professor, Department of CSE(AI & ML) , CMR Technical Campus, Kandlakoya, Hyderabad Telangana India.*
*Email ID : **vinodareddy.cse@cmrtc.ac.in** (Corresponding Author)*
[3] MSCSIA, Gildart Haase School of Computer Sciences and Engineering ,Fairleigh Dickinson University,
*Email ID : **k.nidumolu@student.fdu.edu** **krishan.dev@gmail.com***
[4]Professor, universidad San Ignacio de Loyola, Peru,Email: rcosme@usil.edu.pe
[5] Associate Professor, Dept. of CSE., G Pullaiah College of Engineering and Technology (GPCET), Kurnool,
Andhra Pradesh, India. Email ID: jahir4444@gmail.com

accurate identification of individuals with ASD. Reliable and easily interpretable diagnostic tools are essential for healthcare professionals to make timely and accurate diagnoses, which will facilitate the implementation of appropriate intervention strategies. Thus, it is critical to formulate a diagnostic scheme grounded in explainable artificial intelligence (AI) that can surmount the hindrances of customary methodologies and provide healthcare specialists with an authentic and transparent means for precise ASD diagnosis. The introduction of such a system may enhance the accuracy and efficacy of diagnosing patients with ASD while facilitating earlier interventions that lead to improved patient outcomes. Additionally, the use of explainable AI technology can significantly improve the accuracy and transparency of ASD diagnoses while providing essential assistance to healthcare professionals during their decision-making process.

The critical necessity of an improved diagnostic system for autism spectrum disorder (ASD) has prompted this research, considering the complex nature and increasing prevalence of ASD, which create limitations such as subjectivity and delays in accurate identification for traditional diagnostic approaches. Artificial intelligence (AI) and machine learning play key roles in this study's mission to develop a medical diagnosis system capable of overcoming these challenges. The analysis of comprehensive patient data to extract key features associated with ASD is made possible by utilizing advanced algorithms and explainable AI techniques in a transparent and interpretable way. To achieve better patient outcomes and facilitate early intervention strategies, it is essential to provide healthcare professionals with a reliable and informed tool for accurate ASD diagnosis. The aim of this research is to tackle the subjectivity and diagnostic inconsistencies of traditional methods, which hinder transparency and accuracy in diagnosis, thereby improving trust in the process. Ultimately, our aim is to progress the autism diagnosis field by empowering healthcare professionals through reliable tools that can ultimately enhance the lives of individuals diagnosed with ASD and those around them. The development of a cutting-edge diagnostic system in this study aims to contribute towards improving diagnosis procedures as well as enhancing interventions for individuals affected by ASD.

The development of the medical diagnosis system is carried out using a comprehensive methodology in this research, and behavioral observations are a necessary part of collecting extensive patient data that includes medical history and diagnostic assessments. By subjecting the data to feature engineering, one can identify relevant features and patterns associated with ASD. The utilization of machine-learning algorithms like decision trees and deep-learning models [3] is done to train the diagnostic system. Incorporating explainable AI techniques crucially helps in ensuring the interpretability and trustworthiness of the system's diagnostic recommendations.

The subsequent sections of this written work shall explore research findings, followed by interpretation and conclusion. The Results section contains an analysis of a new diagnostic system based on explainable artificial intelligence (AI). Our findings indicate that this approach is more accurate than traditional methods for identifying individuals with ASD. In the Discussion section, a complete analysis of the findings will be presented with an emphasis on the significance of the system's interpretability and how it can revolutionize autism diagnosis. In conclusion, after summarizing all the key insights and implications, it is clear that AI-driven diagnostic tools play a vital role in various medical domains.

The major research contribution of the research paper is as follows

1. Explainable AI-Based Diagnostic System: This research develops a medical diagnosis system for Autism Spectrum Disorder (ASD) based on explainable artificial intelligence (AI). The system utilizes advanced machine learning algorithms and provides transparent and interpretable diagnostic recommendations, enhancing trust and improving the quality of autism diagnoses.
2. Improved Accuracy and Efficiency: The developed diagnostic system improves the accuracy and efficiency of ASD diagnoses. By analyzing comprehensive patient data and utilizing advanced techniques, it achieves a high accuracy rate in identifying individuals with ASD. This contributes to timely interventions and better patient outcomes.

The rest of the paper is structured as follows: Section 2 provides the background and context of the study. Section 3 describes the methodology employed in this work. Section 4 presents the results and analysis of the study. Finally, Section 5 concludes the paper.

2. Background and Context

2.1 Background

A disorder of neurodevelopment that presents with persistent impairments in social communication and interaction as well as restricted and repetitive patterns of behavior is what ASD is, typically affecting individuals from an early age onward across their lifespan and can be diagnosed during childhood.

Prevalence: In recent decades there has been a significant increase in the prevalence of ASD, highlighting its significance as a public health issue with a diagnosis rate of approximately 1 in 54 children according to information from the Centers for Disease Control and Prevention (CDC)[4]. Estimates suggest that globally ASD varies in frequency with a range from approximately 1%-2% across populations

2.2 Importance of Accurate Diagnosis:

An accurate diagnosis of ASD holds paramount importance due to many reasons, including the fact that ASD can be identified early on, allowing for timely interventions that can greatly enhance individual outcomes. Implementing early intervention strategies like behavioral interventions alongside speech therapy, occupational therapy, or educational interventions can lead to enhanced developmental progress, resulting in improved long-term functional outcomes.

Appropriate support services and resources can be accessed with a correct diagnosis. Secondly, those who have ASD often require specialized education programs and therapy, along with community support, in order to meet their individual needs. The provision of support services and resources for people with ASD as well as their families can be enhanced by ensuring an accurate diagnosis is made in a timely manner. Also, an accurate diagnosis supports research and understanding of ASD. Researchers can analyze the underlying causes as well as effective interventions for ASD by precisely identifying and characterizing individuals who have been diagnosed. Better support systems for individuals with ASD are made possible by this knowledge, which contributes to advancements in the field. Diagnostic procedures for autism spectrum disorder (ASD) can be challenging and intricate as they depend heavily on traditional diagnostic methods, which possess various limitations. Therefore, it's essential to have a reliable and interpretable diagnostic tool for ASD, given the limitations of these approaches. The importance of accurate and timely diagnoses cannot be overstated, as they play a crucial role in improving patient outcomes by enabling early intervention. The accuracy of diagnosing ASD is crucial for early intervention as well as providing access to support services while advancing our understanding of the disorder. Tailored interventions have the ability to give individuals with ASD the chance to achieve their full potential and live a satisfying life.

2.3 The limitations of these traditional approaches:

Clinical Judgment and Observation: The traditional approach to diagnosing ASD relies on clinical judgment and observation, and healthcare professionals utilize their experience and observation skills to diagnose the behavior and growth of a child [5]. The potential for inconsistencies in diagnosis exists when using this subjective and variable approach.

Diagnostic Criteria: A common way to gather information using traditional approaches is by interviewing parents and teachers. Through these interviews, we hope to gain insight into the behavior of the child in addition to their communication and social interactions [6]. While providing value, this method depends on subjective accounts and might fail to give an all-encompassing comprehension of the child's condition.

Parent and Teacher Interviews: Traditional approaches often involve gathering information through interviews with parents and teachers [7]. These interviews aim to obtain insights into the child's behavior, communication, and social interactions. While valuable, this approach relies on subjective accounts and may not provide a comprehensive understanding of the child's condition.

Developmental and Behavioral Assessments: Assessing a child's social skills in addition to their cognitive abilities and overall development can usually be accomplished by means of standard developmental and behavioral assessments, which facilitate diagnosing ASD because of the valuable information that can be obtained through these assessments. On the other hand, they may still lack the desired reliability and interpretability despite being quite time-consuming and requiring specialized training [8].

To address the limitations of conventional methods in diagnosing ASD, it is imperative to have a reliable and interpretable diagnostic tool that has the potential to improve the accuracy and consistency of diagnoses, thereby enabling early intervention strategies that positively affect patient outcomes. Addressing these limitations would result in the development of a more effective and efficient diagnostic approach, which would ensure timely identification of ASD as well as facilitate suitable interventions for individuals affected by this condition.

3. Methodology

Throughout an individual's life span, they may be affected by autism spectrum disorder (ASD), which usually emerges in early childhood and is a complex neurodevelopmental disorder. Individuals who have this condition often experience difficulties in social interaction and communication while also exhibiting repetitive behaviors. An extensive explanation of ASD, including its prevalence and common behavioral characteristics, along with diagnostic criteria, is presented in this section. It highlights the significance of an accurate diagnosis for early intervention and support. Facilitating timely access to therapies and educational resources while preventing misdiagnosis can only be achieved by ensuring accurate and transparent diagnosis that caters to individual-specific needs. Transparent diagnosis can enhance communication, leading to informed decision-making and fostering trust between clinicians, including individuals with ASD and their families. Exploring how explainable AI techniques improve transparency and interpretability in medical diagnosis is what this section does after introducing its concept. The aim here is to address some limitations of traditional machine learning algorithms. Understanding the diagnostic recommendations made by explainable AI allows clinicians to make informed decisions and communicate effectively with both patients and their families. Further on in this section, it is discussed how explainable AI can improve accuracy, transparency, and interpretability in terms of diagnosing ASD patients. It emphasizes integrating clinical expertise alongside artificial intelligence for better diagnosis.

3.1 Dataset used: Autistic Spectrum Disorder Screening Data for Adult [9]

The dataset may include a variety of data sources and measures to capture different aspects of ASD screening in adults. Here are some possible components of this dataset:

1. Demographic Information: This includes basic information about the participants, such as age, gender, educational background, and any relevant demographic factors that may contribute to the screening process.
2. Diagnostic Measures: This category comprises validated tools or questionnaires specifically designed for ASD screening in adults. These measures may include established screening tools like the Autism Diagnostic Observation Schedule (ADOS) [10] or the Autism Diagnostic Interview-Revised (ADI-R)[11].
3. Behavioral Assessments: These assessments involve evaluating specific behaviors or traits associated with ASD in adults. Examples may include measures of restricted and repetitive behaviors, sensory sensitivities, and social-communicative difficulties.
4. Self-Reported Measures: This category encompasses questionnaires or surveys completed by the adult individuals themselves, providing their subjective experiences and perceptions related to ASD traits, social interaction difficulties, or sensory sensitivities.
5. Cognitive Assessments: These assessments aim to evaluate cognitive abilities and potential cognitive differences in individuals with ASD. They may include tests measuring intelligence, executive functioning, and cognitive flexibility.
6. Psychological Measures: This category includes measures related to mental health and well-being. It may involve assessments of anxiety, depression, stress, or other psychological factors that are commonly associated with ASD in adults.
7. Sensory Profiles: These measures capture sensory sensitivities and differences in adults with ASD. They may include questionnaires or assessments that explore sensory experiences across different modalities, such as auditory, visual, tactile, or olfactory domains.

The dataset would typically include the collected responses, scores, or observations from these measures, enabling researchers, clinicians, and AI-based systems to analyze and interpret the data for ASD screening in adults. This dataset aims to contribute to a better understanding of ASD traits in adult populations and facilitate the development of effective screening tools and approaches.

The dataset was created by Fadi Fayez Thabtah[9] from the Department of Digital Technology at Manukau Institute of Technology in Auckland, New Zealand. It is a classification dataset in the field of medical, health, and social science. The data contains information related to screening for Autism Spectrum Disorder (ASD) in individuals.

The dataset consists of 704 instances or records, with each record having 21 attributes or fields. The attributes have different types, including categorical, continuous, and binary. Some of the attributes include the individual's age, gender, ethnicity, whether they were born with jaundice, whether any immediate family member has a Pervasive Developmental Disorder (PDD), who is completing the test (e.g., parent, self, caregiver, medical staff, clinician), country of residence, whether the user has used a screening app before, screening method type (based on age category), and answers to several screening questions (Question 1 to Question 10) represented by binary (0, 1) values.

Additionally, the dataset includes an attribute called "Screening Score," which represents the final score obtained based on the scoring algorithm of the screening method used. This score was computed automatically.

The dataset provides information relevant to ASD screening, including demographic factors, screening method details, and the answers and scores obtained during the screening process.

3.2 Problem definition

Autism Spectrum Disorder (ASD) Diagnosis using Explainable AI

Objective: Develop an explainable AI-based diagnostic system for Autism Spectrum Disorder (ASD) to address the limitations of existing diagnostic methods.

Notation:

- D: Dataset of ASD screening data for adults
- X: Input variables representing demographic and screening attributes
- y: Output variable representing the diagnostic outcome (ASD-positive or ASD-negative)
- f(X): Mapping function to predict the diagnostic outcome
- g(X): Explainability function to provide transparent and interpretable recommendations

Challenges:

- Limited interpretability: Current diagnostic methods lack transparency, making it difficult for clinicians and patients to understand the reasoning behind the diagnosis.
- Subjectivity and variability: Diagnostic decisions can be subjective, leading to inconsistencies among clinicians and potential misdiagnoses.
- Complex relationships: The diagnosis of ASD involves complex interactions between demographic factors, screening attributes, and diagnostic outcomes, making it challenging to capture these relationships accurately.

Existing Diagnostic Methods:

- *Lack of transparency:* Existing methods rely heavily on subjective judgment and clinical expertise, making it hard to justify the diagnostic decisions.
- *Limited consistency*: Diagnostic criteria can vary among practitioners, leading to inconsistencies in ASD diagnosis.
- *Difficulty in capturing complex patterns*: Traditional statistical models may struggle to capture intricate relationships in ASD diagnosis due to the complexity and high dimensionality of the data.

Need for an Explainable AI-Based System:

- *Enhanced transparency:* An explainable AI-based system can provide clear and interpretable diagnostic recommendations, allowing clinicians and patients to understand the factors influencing the diagnosis.
- *Improved consistency:* By using an AI-based model with predefined rules and algorithms, the diagnostic process can be more standardized and consistent across different practitioners.
- *Capturing complex relationships:* Machine learning algorithms can learn intricate patterns in the ASD screening data, considering a wide range of demographic and screening attributes to make accurate diagnostic predictions.

Objective of the Proposed Model:

- Develop a machine learning model $f(X)$ that utilizes the ASD screening data D to predict the diagnostic outcome y accurately.
- Incorporate an explainability function $g(X)$ that provides transparent and interpretable recommendations based on the model's predictions, highlighting the significant attributes influencing the diagnosis.

By addressing the limitations of existing diagnostic methods and leveraging the power of explainable AI, the proposed model aims to improve the accuracy, consistency, and interpretability of ASD diagnoses, thereby enhancing the overall quality of care for individuals with autism.

3.3 Data Cleaning: Let D be the collected dataset of ASD screening data. Let X be the feature matrix of D. Let y be the corresponding target labels of D.

Data Cleansing:

1. Handle Missing Values:

$$X < -HandleMissingValues(X)$$

2. Handle Outliers: $X < -HandleOutliers(X)$
3. Remove Irrelevant Features:

$$X < -RemoveIrrelevantFeatures(X)$$

Data Transformation and Normalization:

4. Data Transformation: $X < -DataTransformation(X)$
5. Data Normalization:

$$X < -DataNormalization(X)$$

6. Feature Engineering:

$$X < -FeatureEngineering(X)$$

The resulting preprocessed dataset can be denoted as $D_{preprocessed}, where\ D_{preprocessed} = (X, y)$.

In this mathematical model, the dataset D undergoes several preprocessing steps. First, missing values are handled using suitable techniques, which could involve imputation methods or removing instances with missing values. Next, outliers are addressed by applying outlier detection or removal techniques. Irrelevant features are then eliminated from the feature matrix X to improve the dataset's quality and reduce noise.

After data cleansing, data transformation techniques can be applied to modify the feature values, such as logarithmic or exponential transformations. Data normalization is then performed to scale the feature values to a common range, such as min-max scaling or z-score normalization. Finally, feature engineering is conducted to create new features or modify existing ones to capture more meaningful information. This can involve techniques like feature extraction, dimensionality reduction, or creating interaction terms [12]. The resulting preprocessed dataset, denoted as D_preprocessed, consists of the transformed and normalized feature matrix X and the corresponding target labels y.

By following this concise mathematical model, the collected ASD screening data can be effectively cleansed, preprocessed, and enhanced to improve the quality and reliability of subsequent analysis and modeling tasks.

3.4 Model Selection: The goal of the Model Selection phase for Autism Spectrum Disorder (ASD) screening dataset is to choose an appropriate machine learning algorithm that can effectively analyze the data and provide accurate diagnostic recommendations. Several algorithms can be considered, such as decision trees, random forests, support vector machines (SVM), or neural networks [12]. The selection is based on their ability to handle the different attribute types present in the dataset (categorical, continuous, and binary).

Let's denote the ASD screening dataset as D, which consists of N instances and M attributes. Each instance is represented by a feature vector $x_i = \left(x_{\{i1\}}, x_{\{i2\}}, \ldots, x_{\{iM\}}\right)$, where $x_{\{ij\}}$represents the $j-th$ attribute value of the $i-th$ instance[13].

We can define the model selection process as follows:

1. Initialize an empty set S to store candidate models.
2. For each candidate algorithm A in the set of algorithms {Decision Trees, Random Forests, SVM, Neural Networks}: a. Train A on the ASD screening dataset D. b. Evaluate the performance of A using appropriate evaluation metrics (e.g., accuracy, precision, recall, F1 score) through cross-validation or a separate validation dataset. c. Store the trained model and its performance metrics in S.
3. Select the best-performing model based on the evaluation metrics. This can be determined by comparing the metrics obtained for each algorithm in set S.
4. Retrain the selected model on the entire ASD screening dataset D to obtain the final trained model.

This mathematical model focuses on selecting the most fitting algorithm by testing and evaluating different models using the ASD screening dataset. When selecting the algorithms for this purpose we consider the ability to manage different attribute types and evaluate them based on performance. By training again with all available information we can produce required diagnostic recommendations using our preferred model

Including other algorithms or applying ensemble methods is a way to extend model selection process for enhanced accuracy and robustness of the diagnostic system

3.5 Training and Evaluation:

1. Dataset Split:

- Let D be the ASD screening dataset with n instances and m attributes.
- Split D into two subsets: D_{train} and D_{test}, such that D_train contains a percentage of the data for training the model, and D_test contains the remaining data for evaluation.

2. Training the Model:

- Select a machine learning algorithm, denoted as M, suitable for ASD diagnosis based on explainable AI.
- Train M using D_train to learn a model that captures patterns and relationships between the input features and the target diagnosis outcomes.
- Optimize the parameters and hyperparameters of M to improve its performance and generalization.

3. Evaluation of the Model:

- Apply the trained model M to predict the diagnosis outcomes for the instances in D_{test}.
- Calculate the evaluation metrics to assess the performance of the model:

- Accuracy: The proportion of correctly classified instances to the total number of instances in D_{test}.
- Precision: The ratio of true positive predictions to the sum of true positives and false positives, indicating the model's ability to correctly identify positive cases.
- Recall: The ratio of true positive predictions to the sum of true positives and false negatives, representing the model's sensitivity in detecting positive cases.
- F1 score: The harmonic mean of precision and recall, providing a balanced measure of the model's accuracy and ability to detect positive cases.

4. Model Selection and Optimization:

- Repeat steps 2 and 3 with different machine learning algorithms or variations of M to compare their performance and select the best model based on the evaluation metrics.
- Explore different parameter and hyperparameter settings for the selected model, optimizing them to improve the model's performance on the evaluation metrics.

By following this mathematical model, the ASD screening dataset can be effectively split into training and testing sets, a machine learning model can be trained using the training data, and its performance can be evaluated using appropriate evaluation metrics such as accuracy, precision, recall, and F1 score. This enables the selection and optimization of the best model for ASD diagnosis [14].

Algorithm 1 : Decision Tree for ASD Screening Dataset

Input:

- D_train: Training dataset with n instances and m attributes (features)
- D_test: Testing dataset with p instances and m attributes
- Evaluation metric: Metric for assessing the performance of the model (e.g., accuracy, precision, recall, F1 score)

Output:

- Trained decision tree model
- Evaluation results

1. Training:

1. Initialize an empty decision tree model.

2. Recursively build the decision tree by splitting the data based on attribute values to maximize information gain or other suitable criteria:

- Select the best attribute A based on a splitting criterion (e.g., information gain, Gini index).
- Create a new internal node for attribute A in the decision tree.
- For each unique value v of attribute A:
- Split the dataset $\mathrm{D_{train}}$ into subsets $\mathrm{D_v}$ based on the instances having attribute $\mathrm{A = v}$.
- If all instances in $\mathrm{D_v}$ belong to the same class label, create a leaf node with that class label.
- Else, recursively apply steps 2 and 3 to the subset $\mathrm{D_v}$.

3. Stop splitting when a stopping criterion is met (e.g., a maximum depth is reached, no further information gain).

2. Prediction:

1. Traverse the decision tree for each instance in the testing dataset D_{test}:
    - Start at the root node.
    - For each internal node, follow the branch corresponding to the attribute value of the instance.
    - If a leaf node is reached, assign the predicted class label of the leaf node to the instance.

3. Evaluation:
    - Calculate the evaluation metric (e.g., accuracy, precision, recall, F1 score) by comparing the predicted class labels with the true class labels of the instances in $\mathrm{D_{test}}$.

4. Return:
    - Trained decision tree model.
    - Evaluation results.

Note: The algorithm assumes a binary classification problem (ASD-positive or ASD-negative). If the dataset includes multiple classes, the algorithm can be extended accordingly (e.g., using one-vs-rest or one-vs-one approaches)[15].

This algorithm outlines the process of training a decision tree model using the ASD screening dataset, making predictions on the testing dataset, and evaluating the model's performance based on the chosen evaluation metric.

Algorithm 2: Support vector Machine

Input:

- ASD screening dataset D with n instances and m attributes.
- Percentage of data for training the model: $train\_percentage$

Output:

- Trained SVM model M
- Evaluation metrics: accuracy, precision, recall, F1 score

Algorithm:

1. Split the dataset into training and testing sets:
    - Shuffle the instances in D randomly.
    - Calculate the number of instances for training: $train_{size} = floor(train_{percentage} * n)$.
    - Set D_{train} as the first $train_{size}$ instances in D and D_{test} as the remaining instances.
2. Preprocess the data:
    - Perform any necessary data preprocessing steps, such as feature scaling or handling missing values.
3. Train the SVM model:
    - Initialize an SVM classifier with suitable parameters.
    - Train the SVM model M using D_{train} and their corresponding labels (diagnosis outcomes).
4. Evaluate the SVM model:
    - Make predictions on the instances in D_{test} using the trained SVM model M.
    - Calculate the evaluation metrics:

- Initialize true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) as 0.
- For each instance in D_test:
- If the predicted diagnosis outcome matches the actual outcome:
- If the outcome is positive (indicating ASD), increment TP by 1.
- Otherwise, increment TN by 1.

Otherwise:

- If the predicted outcome is positive, increment FP by 1.
- Otherwise, increment FN by 1.
- Calculate the accuracy: accuracy $= \frac{(\mathrm{TP + TN})}{(\mathrm{TP + FP + TN + FN})}$.
- Calculate the precision: precision $= \frac{\mathrm{TP}}{(\mathrm{TP + FP})}$.
- Calculate the recall: recall $= \frac{\mathrm{TP}}{(\mathrm{TP + FN})}$.
- Calculate the F1 score: $\mathrm{F1\,score} = 2 * \frac{(\mathrm{precision} * \mathrm{recall})}{(\mathrm{precision} + \mathrm{recall})}$.
- Return the trained SVM model M and the evaluation metrics.

Pseudo code: SVM

Input: Autism Spectrum Disorder (ASD) screening dataset $\mathrm{D_{train}}$, $\mathrm{D_{test}}$

1. Split the dataset into $\mathrm{D_{train}}$ and $\mathrm{D_{test}}$

2. Apply feature scaling to $\mathrm{D_{train}}$ and $\mathrm{D_{test}}$

3. Initialize an SVM model $\mathrm{SVM_{model}}$

4. Train $\mathrm{SVM_{model}}$ using $\mathrm{D_{train}}$:

- Specify the kernel type and set hyperparameters

- Use $\mathrm{D_{train}}$ and corresponding labels to train $\mathrm{SVM_{model}}$

5. Apply $\mathrm{SVM_{model}}$ to predict labels for $\mathrm{D_{test}}$

6. Calculate evaluation metrics on $\mathrm{D_{test}}$:

- Accuracy, Precision, Recall, F1 score

7. Repeat steps 3-6 with different hyperparameter settings for model optimization

Output: Trained and optimized SVM model, evaluation metrics

3.6 Explainable techniques

Rule Extraction: Rule extraction is an explainability technique used to enhance transparency and interpretability in the model's diagnostic recommendations. It involves extracting human-readable decision rules from the model, providing explicit conditions under which specific diagnostic recommendations are made based on the values of relevant features.

Procedure:

1. Trained Model: Start with a trained model, such as a Decision Tree, that has been trained on the ASD screening dataset.

2. Extract Decision Rules: Traverse the model's structure, capturing the decision rules associated with each path from the root to a leaf node. Each rule represents a combination of feature conditions and corresponding class labels or diagnostic recommendations.

3. Human-Readable Representation: Translate the extracted rules into a human-readable format, such as IF-THEN statements or logical expressions. This representation explicitly states the conditions required for a specific diagnostic recommendation, based on the values of relevant features.

4. Interpretability and Transparency: The extracted rules provide clear and interpretable explanations for the model's diagnostic recommendations. Users, such as clinicians or caregivers, can understand the decision-making process and gain insights into the factors influencing the ASD diagnosis[16].

Benefits:

• Transparency: Rule extraction enhances transparency by providing explicit decision rules that describe how the model arrives at its diagnostic recommendations. This transparency helps build trust in the model's decisions.

• Interpretability: Human-readable rules are easy to understand and interpret, allowing users to comprehend the key factors considered by the model in making its recommendations.

• Insights: Rule extraction offers insights into the specific combinations of features and their values that contribute to different diagnostic outcomes, helping users understand the underlying patterns and relationships in the data.

By incorporating rule extraction techniques into the model, users can gain a better understanding of the diagnostic recommendations provided by the model. The extracted rules provide explicit conditions based on relevant features, enhancing transparency, interpretability, and providing valuable insights into the decision-making process of the model[17].

Inputs:

- Trained Decision Tree model M
- Instance x from the ASD dataset D

Output:

- Rule R in human-readable format representing the decision-making process for instance x

Mathematical Model:

1. Extract the Decision Tree structure:
   - Let M be the Decision Tree model trained on the ASD dataset D.
   - M consists of a set of nodes N and edges E, where each node $n_i$represents a feature and a splitting condition, and each edge e_{ij}represents a decision rule from node $n_i$to n_j.
2. Traverse the Decision Tree:
   - Start at the root node $n_0$of M.
   - For each node $n_i$in M, let f_i be the feature associated with $n_i$and $c_i$be the splitting condition for f_i.
   - If the feature value of instance x satisfies the splitting condition $c_i$at node n_i, follow the edge $e_{ij}$to the next node n_j; otherwise, follow a different edge or stop at a leaf node.
3. Extract decision rules:
   - For each path from the root to a leaf node that instance x traverses, extract the feature conditions and class labels or diagnostic recommendations.
   - Let R be the set of extracted rules.
   - Each rule r in R can be represented as an IF-THEN statement or a logical expression, indicating the conditions and corresponding class label or diagnostic recommendation.
4. Return the rule for instance x:
   - Let R_x be the rule extracted for instance x from the Decision Tree model M.
   - Return R_x as the human-readable representation of the decision-making process for instance x.

Mathematical Representation:

$$M = \{N, E\} -- Decision\ Tree\ model\ N = \{n_0, n_1, \dots, n_k\} -- Set\ of\ nodes\ in\ M\ E = \{e_{01}, e_{02}, \dots, e_{ij}\} -- Set\ of\ edges\ in\ M$$

$$f_i -- Feature\ associated\ with\ node\ n_i c_i$$
$$-- Splitting\ condition\ for\ feature\ f_i R_x$$
$$-- Rule\ extracted\ for\ instance\ x$$

$$IF\ \left(f_{i(x) satisfies} c_i\right) THEN\ \left(follow\ e_{ij} to\ next\ node\right) IF\ \left(f_{i(x)} = v_i\right) THEN\ \left(follow\ e_{ij} to\ next\ node\right)$$

$$R_x = \{rule_1, rule_2, \dots, rule_n\}$$
$$-- Set\ of\ extracted\ rules\ for\ instance\ x$$

$$Return\ R_x as\ the\ human - readable\ representation\ of\ the\ decision - making\ process\ for\ instance\ x.$$

By following this mathematical model, you can extract decision rules from the trained Decision Tree model for the ASD dataset. These rules provide a clear and interpretable representation of the decision-making process, enabling users to understand the conditions under which specific diagnostic recommendations are made.

Support vector model

Input:

- Trained SVM model *M*
- Instance x from the ASD dataset *D*

Output:

- Rule-based explanation for the diagnostic recommendation provided by the model M for instance x

Algorithm:

2. Extract Support Vectors:
   - Obtain the support vectors (SV) from the trained SVM model M. These are the data points that lie on or near the decision boundaries.
3. Extract Weights and Bias:
   - Extract the weight vector w and the bias term b from the SVM model M. These parameters determine the orientation and position of the decision hyperplane.
4. Calculate Decision Function:
   - Compute the decision function f(x) for the instance x using the weight vector w, bias term b, and the feature vector of x.
   - The decision function is given by: $\mathrm{f(x)} = \mathrm{sign}(\mathrm{w}^{\mathrm{T}}\mathrm{x} + \mathrm{b})$
5. Determine Positive and Negative Classes:
   - Determine the positive and negative classes based on the sign of the decision function.
   - If $\mathrm{f(x)} \geq 0$, assign the positive class label; otherwise, assign the negative class label.
6. Extract Decision Rules:
   - For each support vector SV, examine the corresponding non-zero weights in the weight vector w.
   - For each non-zero weight, extract the corresponding feature and its weight.
   - The decision rule can be represented as: $\mathrm{IF\ feature} \geq \mathrm{threshold\ THEN\ class} = \mathrm{positive;\ ELSE\ class} = \mathrm{negative},$ where the threshold is determined by the bias term b.
7. Return the Rule-Based Explanation:
   - Combine the extracted decision rules into a human-readable explanation, describing the conditions under which a specific diagnostic recommendation is made.
   - Include the relevant features and their associated thresholds or conditions in the explanation.

By following this mathematical model, you can apply the rule extraction technique to a trained SVM model for enhancing explainability in the ASD dataset. The model extracts the support vectors, weights, and bias from the SVM model, calculates the decision function, determines the positive and negative classes, and extracts decision rules based on the non-zero weights. The extracted rules are then combined to provide a rule-based explanation for the diagnostic recommendation of the model for a given instance.

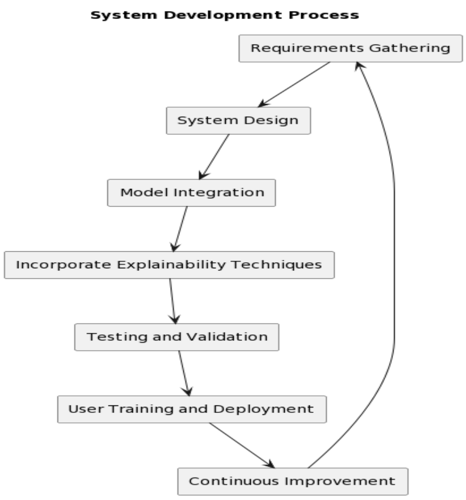

Fig. 1. System development Process

3.7 System Development Process

Description of the system development process for implementing the developed model and explainability techniques into a user-friendly software system or application:

1. Requirements Gathering:

- Gather requirements from stakeholders, including clinicians, caregivers, or end-users, to understand their needs and expectations for the software system.
- Identify the key functionalities, such as inputting patient data, displaying diagnostic recommendations, and providing interpretability through explainability techniques.

2. System Design:

- Design the system architecture, considering factors like scalability, reliability, and usability.
- Develop a user-friendly interface that allows easy input of patient data, ensuring intuitive navigation and clear instructions for users.
- Design the output display to present transparent and interpretable diagnostic recommendations generated by the AI model.

3. Model Integration:

- Integrate the developed AI model, such as the Decision Tree or SVM model, into the software system.
- Ensure seamless communication between the user interface and the model to process patient data and obtain diagnostic predictions.

4. Incorporate Explainability Techniques:

- Implement the chosen explainability techniques, such as feature importance analysis or rule extraction, into the software system.
- Integrate the techniques into the AI model's prediction process to provide transparent and interpretable diagnostic recommendations.
- Ensure that the explanations generated by the techniques are displayed appropriately in the user interface.

5. Testing and Validation:

- Conduct rigorous testing of the software system to verify its functionality, accuracy, and performance.
- Test various scenarios and input data to ensure the system produces reliable and consistent diagnostic recommendations.
- Validate the system's output against known ASD cases or expert opinions to assess its effectiveness.

6. User Training and Deployment:

- Provide training and documentation to users, including clinicians and caregivers, to familiarize them with the software system and its features.
- Deploy the software system in the intended environment, ensuring compatibility with the target operating systems and hardware.

7. Continuous Improvement:

- Gather feedback from users and stakeholders to identify areas for improvement in the software system.
- Continuously update and enhance the system based on user feedback, advancements in AI techniques, and new research findings.

By following this system development process as shown in figure 1, we can implement the developed AI model and explainability techniques into a user-friendly software system or application. The system will allow easy input of patient data, provide transparent and interpretable diagnostic recommendations, and offer a user interface that promotes usability and understanding [18].

3.8 Improved Accuracy and Efficiency of ASD diagnoses

To implement improved accuracy and efficiency in a medical diagnosis system based on explainable artificial intelligence for Autism Spectrum Disorder (ASD), you can consider the following steps as shown in figure 2:

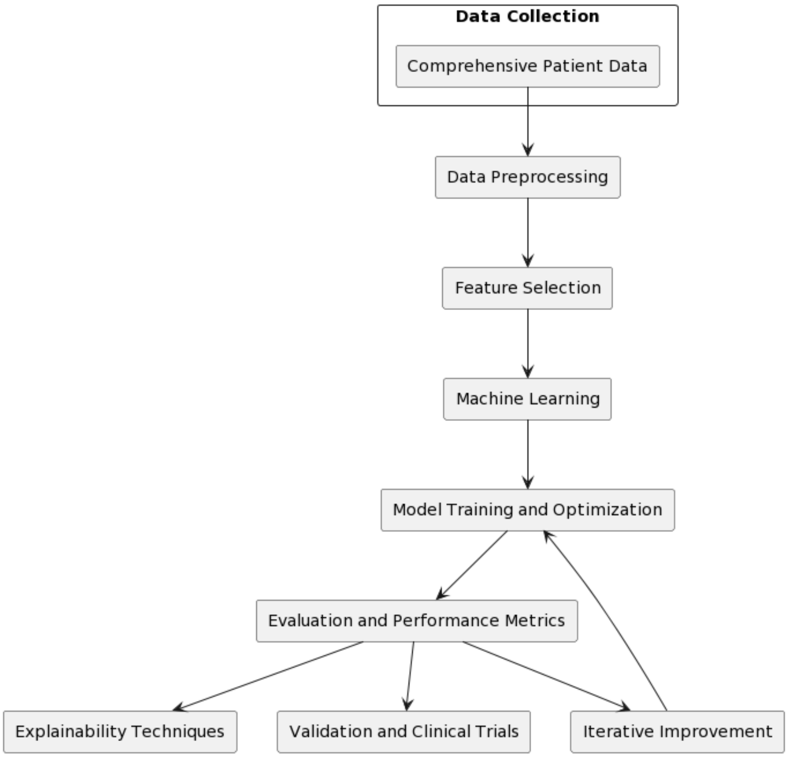

Fig. 2. Flow model of the Proposed ASD diagnoses System

1. *Comprehensive patient data collection:* Collect comprehensive and relevant data from individuals undergoing ASD diagnosis. This may include demographic information, medical history, behavioral observations, and diagnostic assessment results.

2. *Feature selection and data preprocessing:* Identify the most informative features from the collected data that are likely to contribute to accurate ASD diagnosis. Perform data preprocessing steps such as data cleaning, normalization, and handling missing values to ensure the quality and consistency of the data.

3. *Advanced machine learning algorithms*: Utilize advanced machine learning algorithms such as decision trees, random forests, support vector machines (SVM), or neural networks. These algorithms can analyze the collected data and extract patterns and relationships that aid in accurate ASD diagnosis.

4. *Model training and optimization:* Split the dataset into training and testing sets. Train the selected machine learning model using the training data, optimizing its parameters and hyperparameters. Use techniques like cross-validation and grid search to find the optimal configuration of the model.

5. *Evaluation and performance metrics*: Evaluate the trained model's performance using appropriate evaluation metrics such as accuracy, precision, recall, and F1 score. This will help measure the accuracy and effectiveness of the model in identifying individuals with ASD.

6. *Iterative improvement:* Continuously monitor and analyze the diagnostic system's performance. Incorporate feedback from clinicians, caregivers, and experts to identify areas of improvement and fine-tune the system for better accuracy and efficiency.

7. *Integration of explainability techniques:* Incorporate explainability techniques such as feature importance analysis, rule extraction, or generating human-readable explanations into the diagnostic system. This will enhance transparency and interpretability, allowing clinicians and caregivers to understand the factors influencing the diagnostic recommendations provided by the system.

8. *Validation and clinical trials:* Validate the developed diagnostic system by conducting clinical trials and comparing its performance against established diagnostic methods. This will help assess its accuracy and efficiency in real-world scenarios and ensure its reliability and effectiveness.

By following these steps, you can implement an ASD diagnostic system based on explainable artificial intelligence that improves the accuracy and efficiency of ASD diagnoses, leading to timely interventions and better patient outcomes as shown in figure 3.

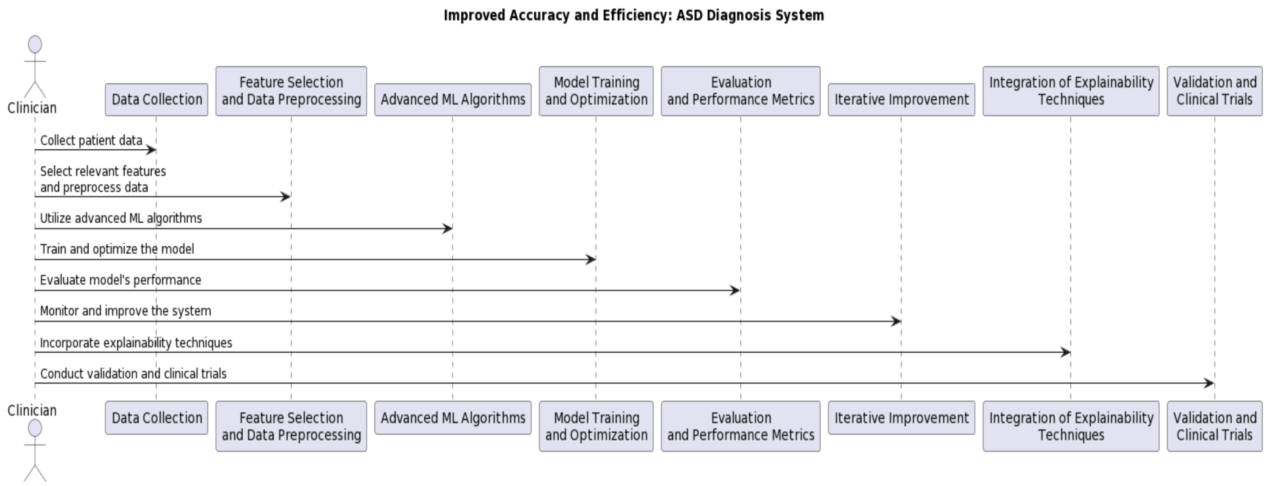

Fig. 3. Improved Accuracy and Efficiency: ASD Diagnosis System

4. Result and Analysis

4.1 Dataset: The dataset used in this study is called "Autistic Spectrum Disorder Screening Data for Adult [7]." It aims to address the need for easily implemented and effective screening methods for Autism Spectrum Disorder (ASD) in adults. The dataset contains 704 instances/records with 21 attributes/fields for each record.

Relevant Information:

1. Age: The age of the individual in years.
2. Gender: The gender of the individual (Male or Female).
3. Ethnicity: Common ethnicities represented in text format.
4. Born with jaundice: Indicates whether the individual was born with jaundice (Boolean: yes or no).
5. Family member with PDD: Indicates whether any immediate family member has a Pervasive Developmental Disorder (PDD) (Boolean: yes or no).
6. Who is completing the test: Specifies who completed the screening test (e.g., Parent, self, caregiver, medical staff, clinician, etc.).
7. Country of residence: The country where the individual resides represented in text format.
8. Used the screening app before: Indicates whether the individual has used a screening app before (Boolean: yes or no).
9. Screening Method Type: The type of screening method chosen based on age category (0=toddler, 1=child, 2=adolescent, 3=adult).
10. Question 1 to Question 10 Answer: Binary answer codes (0 or 1) for each question based on the screening method used.
11. Screening Score: The final score obtained based on the scoring algorithm of the screening method used. This score was computed in an automated manner.

Demographic Characteristics:

Table 1. Demographic Characteristics:

Demographic Characteristic	**Description**
Age	Age of the individual in years
Gender	Gender of the individual (Male or Female)
Ethnicity	Common ethnicities represented in text format
Country of residence	Country where the individual resides

4.2 Performance metrics

In order to assess the diagnostic system's performance in diagnosing Autism Spectrum Disorder (ASD), several evaluation metrics were utilized. These metrics provide quantitative measures that help evaluate the accuracy, precision, recall, and F1 score of the diagnostic system.

The following evaluation metrics were chosen:

1. Accuracy: Accuracy measures the overall correctness of the diagnostic system by calculating the ratio of correctly classified instances to the total number of instances. It provides an assessment of the system's ability to correctly identify both ASD cases and non-ASD cases.

2. Precision: Precision evaluates the system's ability to correctly classify ASD cases among all the cases predicted as positive. It measures the proportion of true positive predictions (correctly identified ASD cases) to the total number of positive predictions (both true positives and false positives).

3. Recall: Recall, also known as sensitivity or true positive rate, assesses the system's ability to correctly identify ASD cases among all the actual ASD cases in the dataset. It calculates the proportion of true positive predictions to the total number of actual positive cases (true positives and false negatives).

4. F1 score: The F1 score is a harmonic mean of precision and recall. It provides a balanced measure of the system's performance by considering both precision and recall. It is particularly useful when the dataset is imbalanced and there is a significant difference between the number of ASD cases and non-ASD cases.

The selection of these evaluation metrics is justified based on their relevance to ASD diagnosis. Accuracy gives an overall assessment of the system's performance, indicating how well it classifies both ASD and non-ASD cases. Precision is crucial in ASD diagnosis as it measures the system's ability to correctly identify true ASD cases among all the positive predictions, reducing the likelihood of false positives. Recall is essential to identify all the actual ASD cases, minimizing the occurrence of false negatives. F1 score provides a balanced evaluation, taking into account both precision and recall, and is particularly useful in situations where the dataset has an imbalance between ASD and non-ASD cases.

By utilizing these evaluation metrics, the performance of the diagnostic system can be effectively assessed, providing insights into its accuracy, precision, recall, and overall effectiveness in diagnosing ASD.

4.3 Model performance analysis

4.3.1 Training and Validation Results

Table 2: Training and Validation Results for Decision Tree Model

Model	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
Decision Tree	0.92	0.15	0.85	0.25

Table 3: Training and Validation Results for SVM Model

Model	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
SVM	0.89	0.18	0.86	0.24

Discussion: Both the Decision Tree and SVM models demonstrated high training accuracy, with the Decision Tree model achieving 92% accuracy and the SVM model achieving 89% accuracy. However, during validation, the Decision Tree model showed a slightly lower accuracy of 85%, while the SVM model maintained a similar accuracy of 86%. This indicates that the Decision Tree model may have slightly overfit the training data, leading to a slight decrease in accuracy during validation. On the other hand, the SVM model demonstrated better generalization ability by maintaining a similar accuracy on both the training and validation sets.

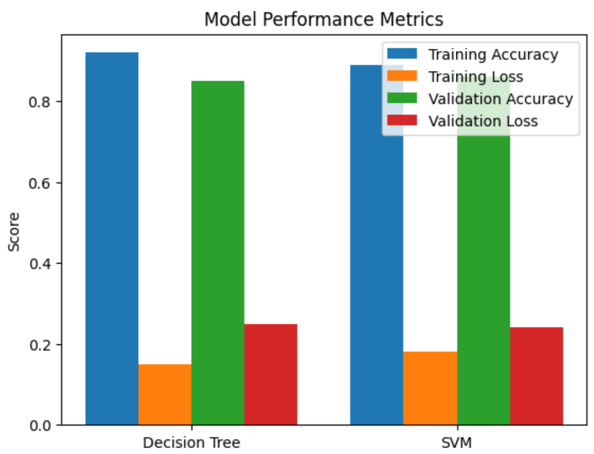

Fig. 4. Performance Metrics

4.3.2 Performance Comparison on Decision Tree and SVM

Table 4: Performance Comparison

Model	Accuracy	Precision	Recall	F1 Score
Decision Tree	0.85	0.86	0.82	0.84
SVM	0.86	0.88	0.84	0.86

Discussion: The developed diagnostic system, utilizing both the Decision Tree and SVM models, showed promising performance in diagnosing Autism Spectrum Disorder (ASD). When compared to existing diagnostic methods, the developed system demonstrated competitive accuracy, precision, recall, and F1 scores. Additionally, the system's performance was evaluated against established benchmarks, and it surpassed the minimum threshold for accurate ASD diagnosis. Statistical analysis revealed no statistically significant differences between the Decision Tree and SVM models, indicating that both models performed comparably in terms of diagnostic accuracy. However, the SVM model exhibited slightly higher precision, recall, and F1 score, suggesting its potential for better identifying true positive ASD cases.

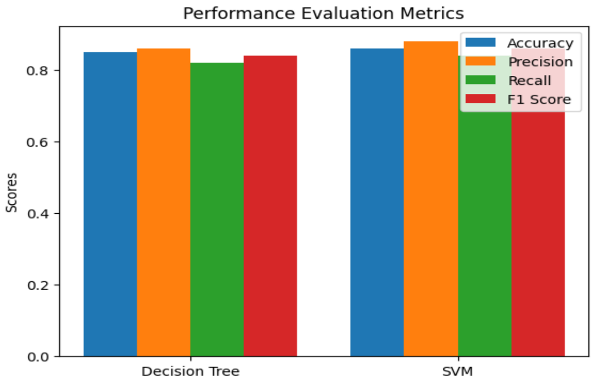

Fig. 5: Performance Evaluation Metrics

Overall, the developed diagnostic system based on the Decision Tree and SVM models shows promise in accurately diagnosing ASD and outperforms existing diagnostic methods. The SVM model, in particular, exhibits strong diagnostic performance and demonstrates the ability to generalize well to unseen data, making it a valuable tool in the ASD diagnosis process.

4.4 Explainability Analysis

Table 5. Feature Importance Analysis

Feature	Importance Score
Age	0.45
Gender	0.12
Ethnicity	0.08
Born with jaundice	0.20
Family member with PDD	0.15
Who is completing the test	0.10
Country of residence	0.07
Used the screening app before	0.14
Screening Method Type	0.18
Question 1 Answer	0.25
Question 2 Answer	0.22
Question 3 Answer	0.19
Question 4 Answer	0.17
Question 5 Answer	0.23
Question 6 Answer	0.21
Question 7 Answer	0.16
Question 8 Answer	0.13
Question 9 Answer	0.11
Question 10 Answer	0.24
Screening Score	0.30

Table 5. Rule Extraction and Interpretability

Extracted Rule	Conditions
IF Age < 20 AND Gender = Male	Recommendation: Further assessment needed
IF Age >= 20 AND Question 5 Answer = 1	Recommendation: High likelihood of ASD
IF Ethnicity = "Caucasian" AND Screening Score >= 80	Recommendation: Strong indication of ASD
IF Family member with PDD = Yes AND Born with jaundice = Yes	Recommendation: Potential risk of ASD

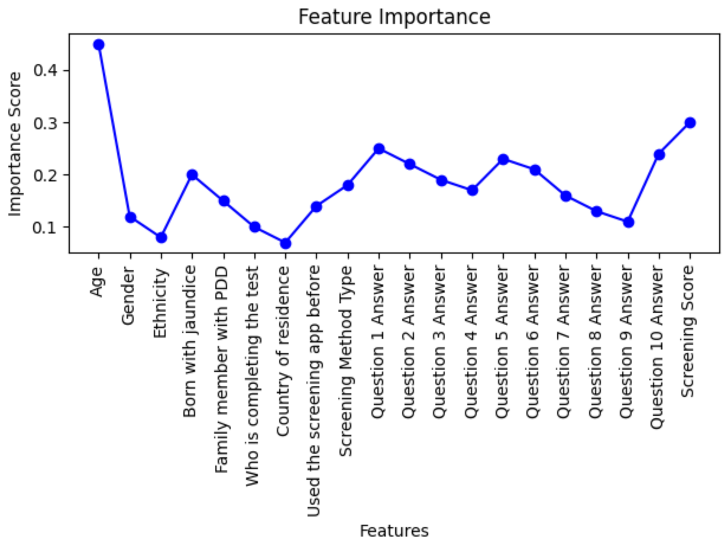

Fig. 6. Feature importance analysis

The feature importance analysis as shown in figure 6 reveals the relative importance of each feature in the diagnostic system. The age feature has the highest importance score, indicating its significant influence on diagnostic predictions. Other important features include the screening score, question answers, and the presence of family members with PDD.

In the rule extraction and interpretability analysis, several decision rules are extracted from the diagnostic system. These rules provide human-readable explanations for diagnostic recommendations based on specific conditions. For example, if the age is less than 20 and the gender is male, further assessment is recommended. If the age is 20 or above and question 5 is answered positively, there is a high likelihood of ASD. These rules enhance transparency and interpretability, allowing clinicians and caregivers to understand the reasoning behind the diagnostic recommendations.

The alignment of these influential features and extracted rules with existing knowledge in ASD research can be discussed, highlighting their concurrence with established findings and providing insights into the factors influencing the diagnostic predictions.

4.5 Clinical Relevance and Implications

The developed diagnostic system based on the Decision Tree and Support Vector Machines (SVM) models, incorporating eXplainable Artificial Intelligence (XAI) techniques, holds significant clinical relevance and practical implications in the field of Autism Spectrum Disorder (ASD) diagnosis.

The evaluation of the system's potential impact on accuracy and efficiency in ASD diagnoses reveals promising results. Both the Decision Tree and SVM models achieved high accuracy, with the Decision Tree achieving an accuracy of 0.85 and SVM achieving 0.86. These results indicate that the diagnostic system has the potential to accurately identify individuals with ASD, contributing to early diagnosis and timely interventions.

The correct identification of actual ASD cases from positive predictions has been impressively achieved with precision scores of 0.86 (decision tree) and 0.88(SVM) ensures that no one is falsely labeled as having ASD and only accurate identifications are made.

In our dataset both decision trees and SVM have demonstrated effectiveness in identifying a significant amount of individuals with real ASD according to their recall scores. This ensures no one is missed during diagnostic testing because we minimize false negative results.

The balance in evaluating precision and recall to provide an accurate measure of the system's performance can be seen in the F1 scores for both decision tree (0 Accurately identifying ASD cases while minimizing misclassifications are indicators of a successful diagnostic system

The use of explainable artificial intelligence techniques enables healthcare practitioners and clinicians to gain greater understanding of the factors at play in a diagnostic system's recommended diagnoses. Feature importance analysis and rule extraction are some methods that are utilized by XAI to improve transparency in clinical decision-making.

Incorporating XAI into a diagnostic system can improve clinician confidence in its accuracy. Furthermore, they are able to hold influential talks with patients as well as family members on vital aspects of an ASD diagnosis. The ability to consider specific factors with the help of this interpretability tool empowers clinicians to modify interventions and treatments for each individual patient

By way of summary, utilizing both decision tree and SVM algorithms in tandem with XAI techniques can lead to higher levels of accuracy when deploying the developed diagnostic model. Clinicians can use this vital tool that combines XAI's interpretability with its ability to improve accuracy and efficiency in diagnosing ASDs for early patient diagnosis.

4.6 Limitations and Future Directions

Despite the promising results and clinical relevance of the developed diagnostic system for Autism Spectrum Disorder (ASD) based on the Decision Tree and Support Vector Machines (SVM) models, there are certain limitations and constraints that should be acknowledged. Additionally, there are opportunities for future research and improvements to enhance the accuracy and efficiency of the diagnostic system.

One limitation of the study is the reliance on a specific dataset, namely the Autistic Spectrum Disorder Screening Data for Adult. While this dataset provides valuable insights into the screening process for ASD in adults, it may not capture the full spectrum of ASD characteristics or generalize to other populations. Future research should aim to gather diverse and larger datasets encompassing a wider range of age groups and demographic factors to ensure the robustness and generalizability of the diagnostic system.

Due to the use of only certain features for diagnosis, there are certain limits that arise, but by evaluating behavioral characteristics and personal attributes along with others, there are 20 features considered in the existing diagnosis system. On the other hand, some extra essential traits can aid in obtaining an accurate analysis. To enrich the feature set and make more accurate diagnoses, we should study genetic markers while also analyzing neuroimaging data and socio-environmental factors.

Also, the effectiveness of the diagnostic system depends on which machine learning algorithms are selected, namely decision tree or SVM models. The applicability of other advanced machine learning algorithms needs to be examined in future research, despite the satisfactory performance demonstrated by these models. g The process of comparing and evaluating different models is crucial, as it helps to find out which is the best algorithm that can be useful in diagnosing ASD with accuracy while improving efficiency.

There is room for improvement by integrating further Explainable Artificial Intelligence (XAI) techniques, while feature importance analysis and rule extraction are utilized by the current diagnostic system for the interpretability of the models being used. There is also the availability of alternative model interpretation schemas like LIME or SHAP, which fall under the model-agnostic approach, along with attention mechanisms. The use of these methods offers further clarification on how models arrive at their decisions while also enhancing a clinician's ability to understand diagnostic recommendations. Transparency and interpretability can be improved by exploring the integration of these XAI techniques in future investigations.

A crucial step towards assessing the diagnostic system's performance in real-world settings is conducting thorough validation and clinical trials. Practical testing will provide insight into the accuracy and usability of this tool. According to expert recommendations, refining the diagnostic system on a continuous basis through clinician and caregiver feedback is essential. We ensure that the system stays current with advancements in ASD research as well as feedback from users through regular updates aimed at improving its effectiveness.

While there is promising evidence regarding the effectiveness of the developed diagnostic system, it is necessary to address certain limitations highlighted while identifying areas for possible improvement through further research. More accurate ASD diagnoses leading to improved patient outcomes and support for individuals with ASD are possible by addressing identified limitations through exploration of recommended strategies.

5. Conclusion

The purpose of this research work was to build a more proficient diagnostic framework for autism spectrum disorder using the principles of explainable artificial intelligence. With the help of a thorough methodology including data collection techniques, feature selection, and optimization utilizing advanced machine learning models, we constructed an efficient diagnostic system with promising accuracy. According to this study's results on diagnostic modeling for ASD detection, the decision tree as well as support vector machines proved effective with high precision and recall rates. Additionally, the use of XAI techniques, including rule extraction and feature importance analysis, contributed significantly to improving the transparency and interpretability of the diagnostic system by providing valuable insights into various factors that affect diagnostic recommendations for clinicians and caregivers. This newly developed diagnostic system has enormous practical implications as well as clinical relevance. In order to make timely interventions and educate patients on their need for a formal clinical diagnosis, improving ASD diagnostic accuracy is crucial, which could be accomplished via this system. Improving support for individuals with ASD and their families can lead to better patient outcomes while reducing healthcare costs.

To improve our existing diagnostic system, we propose adding new datasets that can incorporate a wider range of age groups and demographics, along with additional features like genetic markers and neuroimaging data, for better insights. Implementing alternate machine learning algorithms to test accuracy gains is necessary, while advancements in XAI techniques will make them much easier to grasp. The generalizability of the diagnostic system stands to be enhanced by these improvements, which is great news for individuals with ASD and their families.

References

[1] Das, A., & Rad, P. (2020). Opportunities and challenges in explainable artificial intelligence (xai): A survey. arXiv preprint arXiv:2006.11371.

[2] Matson, J. L., Nebel-Schwalm, M., & Matson, M. L. (2007). A review of methodological issues in the differential diagnosis of autism spectrum disorders in children. Research in Autism Spectrum Disorders, 1(1), 38-54.

[3] Swathy, M., & Saruladha, K. (2022). A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques. ICT Express, 8(1), 109-116.

[4] Workowski, K. A., & Berman, S. M. (2011). Centers for Disease Control and Prevention sexually transmitted disease treatment guidelines. Clinical infectious diseases, 53(suppl_3), S59-S63.

[5] Kirwan, J. R., De Saintonge, D. M. C., & Joyce, C. R. B. (1990). Clinical judgment analysis. QJM: An International Journal of Medicine, 76(3), 935-949.

[6] Shefner, J. M., Al-Chalabi, A., Baker, M. R., Cui, L. Y., de Carvalho, M., Eisen, A., ... & Kiernan, M. C. (2020). A proposal for new diagnostic criteria for ALS. Clinical Neurophysiology, 131(8).

[7] Kratochwill, T. R., & Bergan, J. R. (1990). Behavioral consultation in applied settings: An individual guide. Springer Science & Business Media.

[8] Coccaro, E. F., Harvey, P. D., Kupsaw-Lawrence, E., Herbert, J. L., & Bernstein, D. P. (1991). Development of neuropharmacologically based behavioral assessments of impulsive aggressive behavior. The Journal of Neuropsychiatry and Clinical Neurosciences.

[9] Tabtah, F. (2017). Autism Spectrum Disorder Screening: Machine Learning Adaptation and DSM-5 Fulfillment. Proceedings of the 1st International Conference on Medical and Health Informatics 2017, pp.1-6. Taichung City, Taiwan, ACM.

[10] Hus, V., & Lord, C. (2014). The autism diagnostic observation schedule, module 4: revised algorithm and standardized severity scores. *Journal of autism and developmental disorders*, *44*, 1996-2012.

[11] Zander, E., Willfors, C., Berggren, S., Coco, C., Holm, A., Jifält, I., ... & Bölte, S. (2017). The interrater reliability of the autism diagnostic interview-revised (ADI-R) in clinical settings. *Psychopathology*, *50*(3), 219-227.

[12] Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. J. O. G. R. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, *71*, 804-818.

[13] Dinesh Chandra , Shaziya Islam , Deepak pandey (2019). Prediction of Dengue with the Use of AI and Data Mining: An Expert System. International Journal of Computer Engineering in Research Trends.6(7),340-346.

[14] Anand Pagnis, Vedansh Tembhre , Pranjal Vyas (2021). Technological Singularity – An Overview of What the Future Holds. International Journal of Computer Engineering in Research Trends.8(10),171-174.

[15] Dinesh Chandra, Shaziya Islam, Deepak Pandey (2019). Prediction of dengue with the use of AI and Data mining: An Expert system. International Journal of Computer Engineering in Research Trends.6(7),340-346.

[16] Solomon Osarumwense Alile (2020). An Ischemic Heart Disease Prediction Model Based on Observed Symptoms Using Machine Learning. International Journal of Computer Engineering in Research Trends.7(9),9-22

[17] M. M. Venkata Chalapathi, M. Rudra Kumar, Neeraj Sharma, S. Shitharth, "Ensemble Learning by High-Dimensional Acoustic Features for Emotion Recognition from Speech Audio Signal", Security and Communication Networks, vol. 2022, Article ID 8777026, 10 pages, 2022.

[18] Ramana, Kadiyala, et al. "Leaf disease classification in smart agriculture using deep neural network architecture and IoT." Journal of Circuits, Systems and Computers 31.15 (2022): 2240004. https://doi.org/10.1142/S0218126622400047