JATIYA KABI KAZI NAZRUL ISLAM UNIVERSITY

RESEARCH REPORT

Title: Analysis and Detection of Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder Using Machine Learning Algorithm

Name Kazi Md Shahiduzzaman

Designation Associate Professor

Department Electrical and Electronic Engineering

Research and Extension Center

Jatiya Kabi Kazi Nazrul Islam University

Note

1. This report is applicable to researchers from Jatiya Kabi Kazi Nazrul Islam University, who have been granted funds from the Research and Extension Center of the university.
2. This form should be filled with correct grammar, clear text, and neat handwriting.
3. This research project report may not be distributed without Mr. Kazi Md Shahiduzzaman's concern.

Abstract

The application of machine learning in the diagnosis of neurodevelopmental disorders presents a promising approach to improve diagnostic accuracy and efficiency. The primary objective of this study is to create a more comprehensive model that integrates multiple features and uses various algorithms to accurately diagnose Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD). This research aims to develop and evaluate machine learning models for the accurate detection and classification of ASD and ADHD, including their co-occurrence, through a multimodal data fusion approach. The comprehensive evaluation of various machine learning and deep learning models across three distinct datasets, questionnaires, electroencephalography (EEG), and a combined multimodal dataset integrating both revealed significant insights into their efficacy for detecting autism spectrum disorders. The findings provide a strong foundation for the development of more accurate and reliable diagnostic tools that can complement existing clinical practices, leading to earlier interventions and improved patient outcomes. The ability to identify and differentiate between neurotypical, ASD, ADHD, and comorbid presentations offers a pathway towards more personalized and effective interventions, ultimately contributing to improved outcomes for individuals affected by these complex neurological disorders.

Key words: Neuro-divergance disorder, ASD, ADHD, Machine Learninng, Multimodal Data, Data Fusion, EEG, Questionnaries

Contents

List of Figures

List of Tables

Chapter 1 Introduction

Machine learning techniques have become increasingly prevalent in biomedical fields, especially in psychiatry, as potential diagnostic tools for complex neurodevelopmental disorders such as Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder[1]. These techniques excel at identifying patterns, grouping similar objects, and modeling data, offering the ability to learn from given datasets and predict outcomes for new unseen data[2]. The traditional clinical diagnosis of these disorders is highly dependent on behavioral data and, in some cases, neuroimaging techniques[3]. Machine learning approaches offer a supplementary avenue that has the potential to improve the efficiency and accuracy of the diagnostic process[4]. Traditional methodologies for identifying autism spectrum disorders, such as screening tests, can be resource-intensive and time-consuming[5]. Machine learning provides the opportunity to automate the detection process[6]. Using machine learning in this area does more than just categorize; it helps us better understand how ADHD shows up in different people and uncovers new groups of participants that can improve the precision of diagnoses.e diagnosti Machine learning methodologies have significantly improved diagnostic accuracy, particularly in identifying Autism Spectrum Disorder through mobile applications, surpassing conventional screening techniques in terms of sensitivity and specificity[7,8].

1.1 Background

Difficulties in social interactions with others characterize ASD. This disorder affects the way people interact with others, such as impairments in social behavior, atypical communication, and restricted interests and repetitive behaviors. Early diagnosis of autism spectrum disorder can lead to more effective interventions and improved outcomes for affected individuals[9]. Asperger and Kanner independently described patients who showed socioemotional limitations, emphasizing that these patients remained the same and were disturbed when routines changed[10]. The symptoms of Autism Spectrum Disorder can be identified in children as young as 12 months, but may also manifest later in life[11]. It is described as a 'developmental disorder' because symptoms generally appear in the first two years of life[12]. Early detection of the disorder can significantly improve the quality of life of the individual, primarily through prompt access to proper treatments[3]. The diagnosis of autism spectrum disorder is usually made based on clinical observation, interviews, and standardized diagnostic instruments.

ADHD is gaining awareness because it affects paying attention and controlling impulsive behaviors. ADHD is usually first diagnosed in childhood and often lasts until adulthood. Probably 5–10 percent of children and at least 3–5 percent of adults suffer from attention deficit hyperactivity disorder (ADHD)[13]. So, it is a public concern to detect ADHD early. The gold standard for diagnosing most mental disorders is based primarily on information collected from various informants on the onset, course, and duration of various behavioral descriptors that providers consider when making a diagnostic decision based on the criteria of DSM-5 / International Classification of Diseases 10th Edition[14] specific behavioral descriptors, which providers then consider when making a diagnostic decision based on DSM-5. Some of the standard characteristics include inattention, hyperactivity, and impulsivity[13].

As machine learning algorithms advance, they're becoming more helpful in diagnosing neurodevelopmental disorders like Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder. Both ASD and ADHD share similar symptoms to other neurodevelopmental disorders (NDD)[15]. They suffer from social communication impairments, language skills challenges, difficulty focusing, executive function issues, sensory overload, and hyperactivity/impulsivity. Numerous scientific studies have shown that the two conditions often coexist[16]. It is crucial to differentiate between these conditions for an accurate diagnosis, treatment, and support. The application of machine learning in the diagnosis of neurodevelopmental disorders presents a promising approach to improve diagnostic accuracy and efficiency.

However, the cause of these neurodevelopmental disorders remains unclear. To diagnose ASD and ADHD, several techniques were developed. A survey has been conducted that proves that AI-based techniques are more effective than conventional techniques in detecting NDD[17]. Another study demonstrated that machine learning and deep learning methods can help to accurately detect ASD and ADHD[14].

Therefore, there is an increasing need to develop an effective tool to identify ASD and ADHD in children and assist physicians in making accurate diagnoses. The primary objective of this study is to create a more comprehensive model that integrates multiple features and uses various algorithms to accurately diagnose ASD and ADHD.

1.2 Problem Statement

The increasing prevalence of Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder requires an accurate and timely identification and diagnosis of these disorders[14]. The insufficient number of trained clinicians coupled with limited access to quick and accu-

rate diagnostic tools caused the overlooking of early symptoms of ASD in children around the world[18]. Because ASD and ADHD share similar symptoms with other neurodevelopmental disorders, it is difficult to differentiate between these conditions.

The problem addressed in this research involves exploring the application of machine learning algorithms to enhance the accuracy and efficiency of diagnosing autism spectrum disorder and attention deficit hyperactivity disorder, using diverse data sets and feature engineering approaches to improve early detection and intervention strategies[19]. Current diagnostic methods are based on behavioral observation of symptoms and can be prone to misdiagnosis[20]. Early detection of ADHD can be facilitated through AI-powered screening tools that offer accessible and cost-effective means of identification, which, when combined with additional data, allow for timely intervention and support[21].

1.2.1 Objectives

This research aims to develop and evaluate machine learning models for the accurate detection and classification of Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder. The objectives include

1. Leveraging advanced techniques to process complex datasets and create predictive models that can aid clinicians in early diagnosis and intervention.
2. Exploring various machine learning algorithms to identify the most effective methods for distinguishing between ASD and ADHD based on behavioral and neuroimaging data.
3. Improving the dependability, precision, and speed of diagnosis and classification systems.
4. Developing a synthetic dataset combining questionnaire and EEG data to help improve algorithm accuracy and precision.

1.3 Contribution

This study contributes to the existing body of knowledge by comprehensively analyzing machine learning methods for diagnosing ASD and ADHD, utilizing the fusion of questionnaire data with electroencephalography data provides a more comprehensive feature set for machine learning models, potentially improving diagnostic accuracy. The expected outcomes of this research include the development of novel diagnostic models with improved accuracy and efficiency, which can be integrated into clinical practice to facilitate earlier and more accurate diagnoses of ASD and ADHD, leading to improved patient outcomes and quality of life.

Chapter 2 Literature Review

The use of machine learning methodologies has garnered substantial attention as a promising avenue to enhance the diagnosis and comprehension of autism spectrum disorder and attention deficit hyperactivity disorder[14]. Traditional diagnostic approaches for ADHD are based heavily on subjective assessments, such as questionnaires and interviews, which are susceptible to biases and inaccuracies[22]. Artificial intelligence techniques, including machine learning algorithms, are increasingly being used to analyze large datasets, identify patterns, and make predictions, offering the potential to create more objective and accurate diagnostic and classification models for ADHD[23].

Despite the growing application of ML and deep learning (DL) in healthcare care, a comprehensive and systematic understanding of their specific utility, effective algorithms, and optimal data modalities for the analysis and detection of ASD and ADHD remains fragmented. Existing research often focuses on single data modalities or specific algorithms, leading to a lack of consolidated perspectives on best practices and overarching trends. This fragmented landscape underscores the necessity for a structured approach to synthesize the available knowledge.

Therefore, this review of the literature is warranted to systematically synthesize the current state of research. Its purpose is to identify the most promising computational approaches, delineate the existing challenges that hinder progress, and pinpoint critical gaps that require further investigation. By consolidating these insights, the review aims to provide a clearer roadmap for future research and clinical translation in the field of ML / DL for the detection of ASD and ADHD.

2.1 Literature Review Strategy

A detailed literature review was performed to gather relevant research articles, conference papers, and books related to the application of machine learning in the diagnosis of Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder. The following figure 2-1 shows the overall literature review strategy.

2.1.1 Search Strategy

A robust and systematic search strategy is essential to ensure the comprehensiveness and relevance of this review of the literature. This involves careful selection of databases, precise

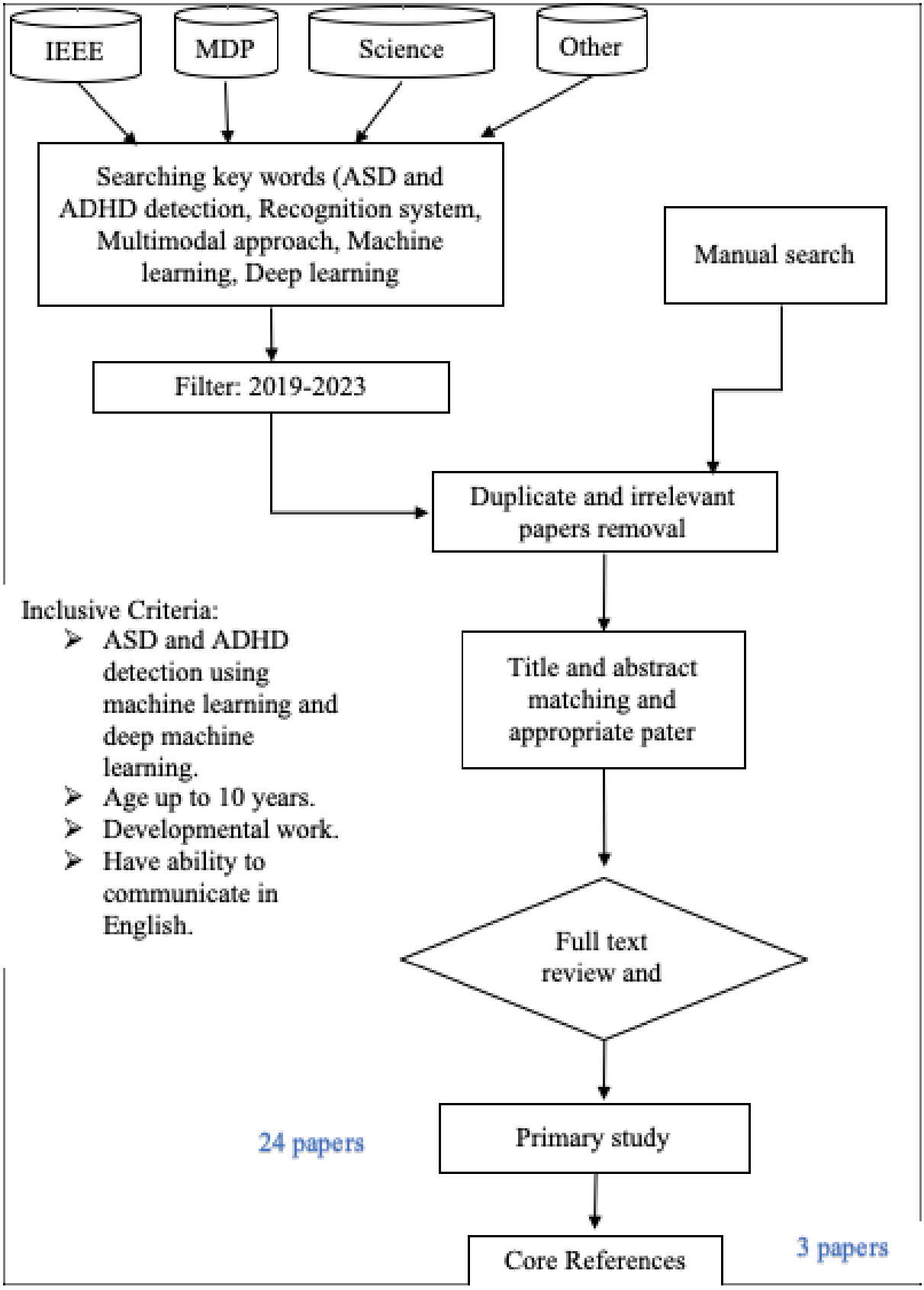

Figure 2-1 Literature Review Strategy

keyword formulation, and appropriate temporal filtering.

2.1.1.1 Database Selection

The primary scientific databases for this systematic review will include IEEE Xplore, MDPI, and Science Direct, as depicted in the conceptual flow chart provided. These platforms are selected for their strong coverage of engineering, computer science, and interdisciplinary research, including machine learning applications in healthcare. To ensure a comprehensive and medically relevant search, additional databases will be incorporated: PubMed/PMC (PubMed Central), which is crucial for the biomedical and medical literature , and Frontiers in Artificial Intelligence, a specialized collection of journals highly relevant to the intersection of AI and medical informatics. Although not a primary source for peer-reviewed articles, the UCI Machine Learning Repository will be briefly consulted to understand the landscape of publicly available datasets in machine learning, particularly to gauge the presence and types of datasets relevant to neurodevelopmental disorders. This exploration can inform discussions on data availability and sharing challenges within the field.

2.1.1.2 Keyword Formulation and Boolean Operators

A robust and systematic keyword strategy will be developed to maximize the retrieval of relevant studies. The core search terms, derived from the user query and the conceptual flow chart provided, will include the following.

- ”ASD detection” OR ”Autism Spectrum Disorder detection” OR ”Autism diagnosis” OR ”Autism screening”
- ”ADHD detection” OR ”Attention Deficit Hyperactivity Disorder detection” OR ”ADHD diagnosis” OR ”ADHD screening”
- ”Recognition system” OR ”Prediction model” OR ”Diagnostic tool”
- ”Multimodal approach” OR ”Multimodal data integration”
- ”Machine learning” OR ”Deep learning” OR ”Artificial intelligence” OR ”ML” OR ”DL” OR ”AI”

These terms will be combined using Boolean operators (AND, OR) to construct precise search strings. For instance, a core search string might be: ((”ASD detection” OR ”ADHD detection”) AND (”Machine learning” OR ”Deep learning”)). Furthermore, to capture studies using specific data modalities identified in the research material, additional terms will be integrated: ”motion tracking” OR ”kinematic data” , ”electronic health records” OR ”EHR” OR

”population registers” , ”neuroimaging” OR ”EEG” OR ”MRI” OR ”cranial ultrasound” , and ”genetics” OR ”omics data”.

2.1.1.3 Publication Year Filter

To ensure that the review reflects the most current advances in this rapidly evolving field, the retrieved studies will be strictly limited to the publication period of **2019-2023**, as explicitly specified in the conceptual flow chart provided. This temporal constraint focuses the review on recent innovations and trends, capturing the latest developments in machine learning applications for the detection of neurodevelopmental disorders.

2.1.2 Study Selection Process

The study selection process will follow a rigorous multistage approach to ensure that only the most relevant and high-quality articles are included in the review. This systematic procedure is designed to enhance the transparency and reproducibility of the review's findings.

2.1.2.1 Initial Screening and Duplicate Removal

All articles retrieved from the selected databases will be systematically imported into a dedicated reference management software (e.g., EndNote, Zotero). Automated and manual checks will be performed to identify and meticulously remove all duplicate entries. An initial rapid scan of titles will then be conducted to discard any papers that are overtly irrelevant to the broad scope of machine learning in healthcare.

2.1.2.2 Title and Abstract Screening

The remaining unique articles will undergo a rigorous title and abstract screening phase. This process will be conducted by two independent reviewers (or by one reviewer following a predefined, strict protocol to minimize bias). The preliminary relevance of the articles will be assessed, ensuring that they explicitly mention the application of ML/DL for the detection, diagnosis, or prediction of ASD and/or ADHD. Studies clearly outside this scope will be excluded at this stage.

2.1.2.3 Full-Text Review for Eligibility

Articles that pass the title and abstract screening will be retrieved in full text for a detailed eligibility assessment. This stage is crucial for applying the precise inclusion and exclusion criteria, ensuring that each selected study directly contributes to the objectives of the review.

2.1.2.4 Inclusive Criteria

For a study to be included in this review, it must meet the following criteria:

- Studies must specifically focus on the analysis, detection, diagnosis, prediction, or detection of autism spectrum disorder (ASD) and / or attention deficit hyperactivity disorder (ADHD).
- Studies must explicitly utilize machine learning (ML) and / or deep learning (DL) algorithms as a core component of their methodology for the aforementioned purposes.
- Studies must involve developmental work, with a primary focus on children up to 10 years of age. This age criterion, derived from the provided conceptual flowchart, emphasizes the review's focus on early detection and intervention.
- Studies must be published in English.

2.1.2.5 Exclusion Criteria

Studies will be excluded if they meet any of the following conditions.

- Studies that do not employ ML/DL techniques for the detection of ASD / ADHD (e.g., purely clinical studies, genetic studies without ML).
- Studies focusing on other neurodevelopmental disorders without specific attention to ASD/ADHD.
- Review articles, editorials, opinion pieces, conference abstracts (unless a full peer-reviewed paper exists and meets criteria) or purely theoretical papers without empirical data.
- Studies published outside the specified 2019-2023 time frame.
- Studies not available in full text or not published in English.
- Studies focused primarily on treatment outcomes, interventions, or educational strategies without a direct diagnostic or screening component.

2.1.2.6 Flowchart Illustrating the Selection Process

The report will feature a comprehensive flow chart (figure 2-1) that visually represents the entire study selection process. This flowchart will directly incorporate and expand upon the structure provided in the user's conceptual image, detailing the number of papers identified at each stage: initial database search, after duplicate removal, after title and abstract screening, and finally, after full-text review, leading to the "Primary study" set (e.g., 24 papers) and "Core References" (e.g., 3 papers). This visual aid improves transparency and reproducibility. The process is summarized in Table 1, which provides a granular overview of the search and se-

lection, detailing the specific databases, search strings, and article counts at each stage. This table is indispensable for replicating the literature search and verifying the completeness of the review. By showing the reduction in article count at each stage, it highlights the systematic and rigorous application of inclusion/exclusion criteria, lending credibility to the findings. Furthermore, the numerical data within the table can itself reveal patterns; for instance, a large drop-off between abstract and full-text review might suggest a high rate of irrelevant studies despite keyword matching, which could prompt further discussion on search string refinement or the heterogeneity of the field. The following table 2.1 summarizes the main references.

Table 2.1 Summary of Studies on Biomarkers, Algorithms, and Accuracy for ASD and ADHD Detection

SL	References	Biomarker	Algorithm	Detect	Accuracy
1	[24]	fMRI	Generative Adversarial Network	ASD and ADHD	70-73.74% for ASD and 65.16-72.09% for ADHD
2	[25]	Questionnaires	Random Forest	ADHD	85.5%
3	[26]	EEG data	Convolutional Neural Network	ADHD	97.7%

2.1.3 Data Extraction and Synthesis

Following rigorous study selection, a systematic approach to data extraction and synthesis is crucial for drawing meaningful conclusions from the included literature.

2.1.3.1 Variables to be Extracted

For each primary study included in the review, a standardized data extraction form will be utilized to collect the following information systematically:

- **Bibliographic Details:** Author(s), Year of Publication, Journal/Conference Name.
- **Participant Demographics:** Total sample size (N), number of cases of ASD / ADHD versus controls, age range of participants (confirming ”up to 10 years”), gender distribution, specific diagnostic criteria used (e.g. DSM-5, ICD-10).
- **Data Modalities:** Detailed description of the type(s) of data collected (e.g. motion tracking data, specific sensor types, EHR fields, neuroimaging sequences such as fMRI, EEG

channels, genetic markers, omics data).

- **Machine Learning Methodology:** Specific ML/DL algorithms employed (e.g., LSTM, Random Forest, SVM, Naive Bayes, kNN), details of feature engineering / selection techniques, data preprocessing steps and cross-validation strategies.
- **Performance Metrics:** All reported evaluation metrics for ML models, including but not limited to accuracy, sensitivity (recall), precision (positive predictive value - PPV), F1 score, support and their confidence intervals if available.
- **Key Findings:** The main results of the study, particularly regarding model performance, insights into new or validated biomarkers (e.g. Fano factor, Shannon entropy of) and any clinical implications.
- **Limitations:** Any limitations explicitly acknowledged by the authors or identified during the review process (e.g., an AI tool struggling with comorbidity or models exhibiting low sensitivity).
- **Future Work:** Recommendations for future research as suggested by the authors.

2.1.3.2 Approach to Qualitative and Quantitative Synthesis

A thematic synthesis approach will be employed for the qualitative analysis of the extracted data. Studies will be grouped and discussed based on commonalities such as the type of ML algorithm used, the primary data modality, the diagnostic challenge addressed (e.g. early detection, differentiation of comorbidities) and the age group studied. This will enable a narrative discussion of the dominant trends, methodological strengths, and persistent weaknesses in the included literature.

Where studies utilize comparable methodologies and report similar performance metrics, a comparative analysis will be performed for quantitative synthesis. This will involve tabulating key performance indicators (e.g. accuracy, sensitivity, AUC) to identify patterns in model efficacy across different algorithms and data types. However, given the potential heterogeneity in study designs, datasets, and reported metrics, a formal meta-analysis may not be feasible and comparisons will be presented descriptively to highlight general trends and notable exceptions.

2.2 Machine Learning in ASD and ADHD Detection

Machine learning techniques have emerged as a powerful tool in the healthcare sector, offering the potential to revolutionize disease diagnosis and prediction [27]. Specifically, in the context of neurodevelopmental disorders such as Autism Spectrum Disorder and Attention Deficit

Hyperactivity Disorder, machine learning algorithms can play a crucial role in early and accurate detection, leading to timely interventions and improved outcomes[28]. The application of machine learning in this area involves training algorithms on large datasets comprising various features extracted from clinical assessments, behavioral observations, neuroimaging data, and genetic information.

ASD is diagnosed in 20 50% people with ADHD; therefore, the detection of ASD and ADHD is highly dependent on professional physicians. In recent years, different techniques such as questionnaires, functional magnetic resonance imaging (fMRI), structural magnetic resonance imaging (sMRI), electroencephalogram (EEG), eye tracking (ET), event-related potential (ERP), continuous performance test (CPT), and facial recognition have been used to identify ASD and ADHD[14]. Depending on the available data, Muhammad et al. proposed a model to estimate the prediction of autism spectrum disorder[29]. Yao Hu et al. developed a dual-stage pseudo-labeling framework, claiming that labeled data for fMRI samples with relatively small portion sizes can maintain reliability in estimating mental disorders[30]. On the other hand, Aarushi et al. showed the advantages of AI-based wearable smart sensors that detect psychological disorders[31]. Suman et al. [6] have studied a CNN-based model that achieved the highest precision compared to other techniques[32].

Recently, numerous studies have focused on the identification of ASD and ADHD, either separately or in combination. In the detection of ASD and ADHD, we sometimes find complex biomarkers such as epilepsy, Alzheimer's disorder, and schizophrenia. Maniruzzaman et al. have shown the need for a machine learning-based tool to detect ADHD because several other mental disorders overlap with each other[33]. Fahad et al. extracted useful features from GAF images using a CNN classification algorithm to detect ASD[34]. Dingfu et al. proposed a model that enables the recognition of ADHD using EEG data[35].

Most of the time, physicians focus on the patient's facial expression and behavioral history. Therefore, Alkahtani et al. [8] detected ASD based on facial landmarks using the MobileNetV2 model, as facial features can convey a person's expression and emotion[36]. Huang et al. detected ASD using a deep belief network and also provided some knowledge about the identification of subtypes with ASD[37]. Yan et al. offered a GAN-based technique that generates a realistic brain region network to achieve accuracy[38].

A single data model cannot achieve better performance. So, Junxia et al. proposed a multimodal approach to identify ASD[39]. To properly detect ASD, Bhuvaneshwari et al. designed a hybrid ensemble machine learning model that utilizes hard and soft vote classifiers[40].

2.3 Multimodal in ASD and ADHD Detection

Integrating multimodal data sources can significantly improve the precision and robustness of machine learning models for the detection of ASD and ADHD, as it allows for a more comprehensive understanding of the underlying neurobiological and behavioral characteristics of these disorders[41]. A multimodal approach, which integrates data from questionnaires and electroencephalography (EEG) signals, offers a promising avenue to improve the detection and classification of Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD) in children. Traditional diagnostic methods, often based on subjective behavioral observations and standardized questionnaires, can be time consuming and contribute to significant diagnostic delays[42]. Although these behavioral scales provide crucial clinical context, they inherently depend on the examiner's subjective evaluation. By merging these important but subjective observations with objective physiological signals, such as EEG, researchers aim to address the limitations of relying solely on one type of evaluation, thus achieving a more comprehensive and accurate understanding of these complex neurodevelopmental conditions[43]. EEG, a non-invasive method, captures real-time electrical activity in the brain, providing objective information on brain wave patterns and functional connectivity that can vary in individuals with ADHD and ASD[44]. These objective physiological data, when integrated with behavioral data, have the potential to enhance the reliability, validity, and efficiency of diagnostic processes.

This multimodal integration has been explored through various methodologies, demonstrating increased detection rates. For example, studies have combined data from behavioral tasks, such as handwriting patterns collected via pen tablets, with brain activity measured by EEG, to capture brain dynamics. The combination of behavioral data, typically collected from questionnaires, with real-time information from EEG signals provides a stronger and more detailed approach to identifying and classifying these disorders, thereby overcoming the limitations of relying solely on one type of data.

Chapter 3 Methodology

This section delineates the comprehensive methodology employed for the analysis and detection of autism spectrum disorder (ASD) and attention deficit / hyperactivity disorder (ADHD) using machine learning techniques. The approach integrates multimodal data, specifically electricalencephalography (EEG) signals and clinical questionnaire/interview data, to develop robust and interpretable diagnostic models. The overall framework follows a standard machine learning pipeline, encompassing data acquisition, pre-processing, feature engineering, model development, training, validation, and rigorous evaluation, while adhering to critical ethical guidelines. The comprehensive block diagram of the methodology is presented below with figure 3-1.

3.1 Data Acquisition

The foundation of this research relies on the meticulous acquisition of two distinct data modalities: Electroencephalography (EEG) data and clinical questionnaire/interview data. These modalities provide complementary information on the neurophysiological and behavioral manifestations of ASD and ADHD.

3.1.1 Electroencephalography (EEG) Data Collection

EEG data will be collected from a cohort of participants diagnosed with ADHD, ASD, and generally developing controls (TDC). The acquisition protocol is designed to capture brain activity under specific cognitive conditions relevant to the disorders.

The study will recruit participants, including individuals diagnosed with ADHD and healthy controls, aged 7-12 years, who adhere to established protocols. The diagnosis of ADHD will be confirmed by an experienced psychiatrist based on DSM-IV criteria. For ASD, similar diagnostic criteria, such as those described in DSM-5, will be applied to ensure a clear distinction between diagnostic groups and control subjects. Rigorous participant selection criteria are essential to ensure that any observed EEG abnormalities are indeed attributable to ADHD or ASD and not to other neurological or psychiatric conditions or medications. For example, the exclusion of control subjects with a history of psychiatric disorders or epilepsy, and the consideration of medication status (e.g., Ritalin use) in ADHD participants, are crucial steps. This careful control over participant characteristics directly impacts the validity of the machine learning model's

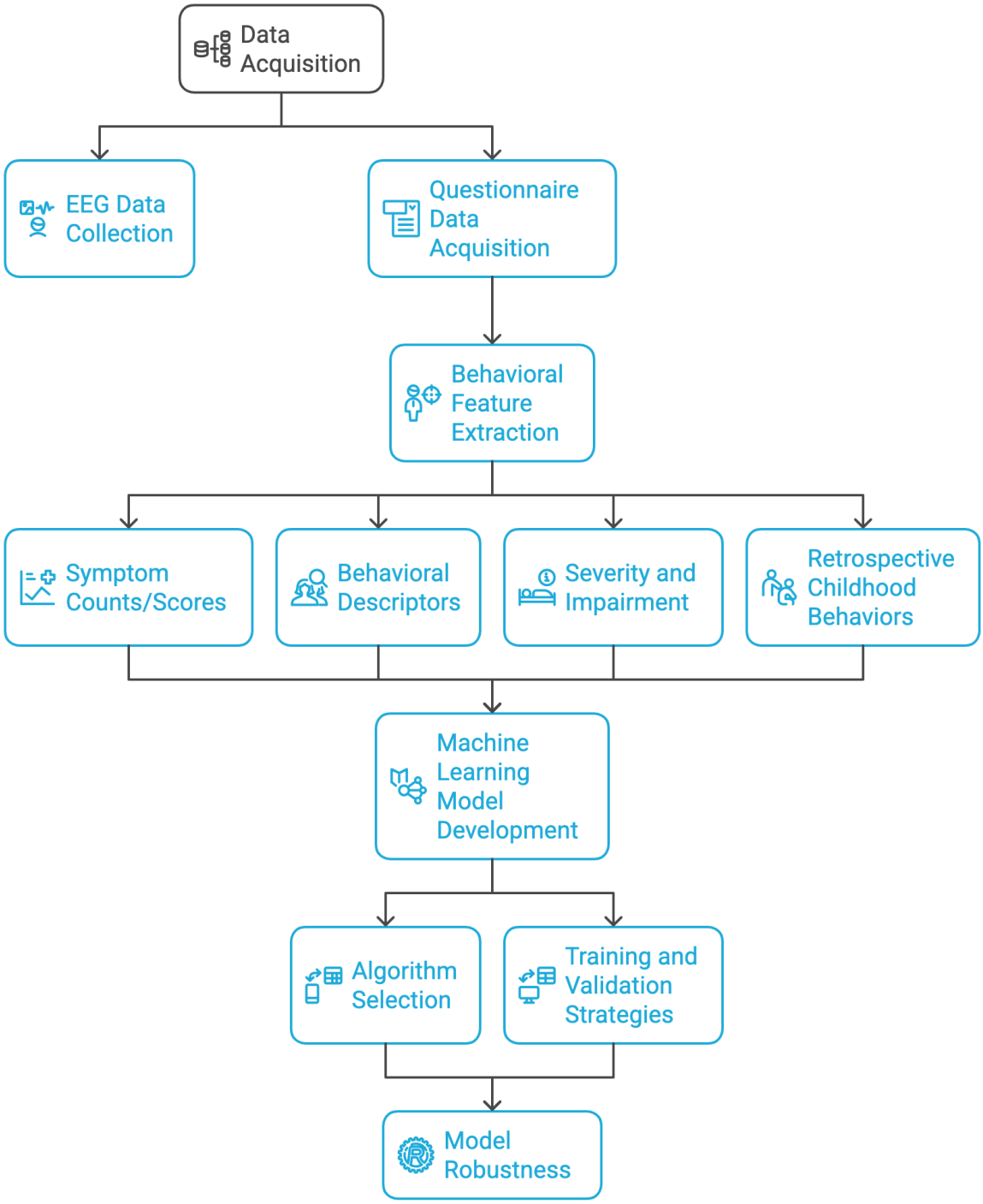

Figure 3-1 Block diagram of the proposed methodology

diagnostic capabilities, preventing the model from inadvertently learning to distinguish between medicated versus unmedicated individuals, or individuals with comorbidities, rather than the core disorder itself.

EEG recordings will be performed using a standard system of 10-20 with 19 channels (e.g. Fz, Cz, Pz, C3, T3, C4, T4, Fp1, Fp2, F3, F4, F7, F8, P3, P4, T5, T6, O1, O2) at a sampling frequency of 128 Hz. A1 and A2 electrodes will serve as earlobe references. The precise specification of these parameters is paramount for reproducibility. Without such detailed methodological descriptions, replicating the study or comparing the results in different research endeavors becomes challenging, if not impossible. In the context of developing medical diagnostic tools, where models must be generalizable across diverse clinical settings and patient populations, standardized acquisition protocols are fundamental for building trustworthy and widely applicable AI solutions.

To generate relevant brain responses, the EEG recording protocol will incorporate visual attention tasks. This is a strategic choice, given that visual attention deficits are a known characteristic in children with ADHD. Participants will be presented with a set of cartoon character pictures and asked to count the characters, with images displayed continuously after each response to ensure sustained stimulus. This task-based approach is crucial for capturing event-related potentials (ERP) and other dynamic brain responses, such as reduced P300 amplitudes, which are related to cognitive tasks and attention regulation in ADHD. The selection of a task designed to provoke specific neural responses that are known to be atypical in ADHD allows for the extraction of highly relevant biomarkers, directly impacting the quality and diagnostic utility of the features extracted for machine learning models.

3.1.2 Questionnaire Data Acquisition

Clinical information, including structured questionnaire responses, will be collected to provide a complete behavioral and historical context for each participant.

Standardized rating scales widely used in the diagnosis of ADHD and ASD will be used for structured questionnaire data. These include, but are not limited to, the Social Responsiveness Scale (SRS), the Child Behavior Checklist (CBCL), the Conners Rating Scale for Parents (CPRS), and the Conners Rating Scale for Teachers (CTRS). These questionnaires provide quantitative measures of symptoms and behaviors in various settings (e.g., home, school), which is crucial for diagnosis.

In addition to structured scales, qualitative data will be gathered through clinical interviews. These may include semi-structured interviews, such as the Structured Clinical Interview for

DSM Disorders (SCID). These interviews are vital to obtain detailed developmental histories, current presentation of symptoms, and examples of behaviors, often from multiple informants (patients, parents, teachers, and family members). The collection will emphasize retrospective accounts of childhood behaviors for the diagnosis of ADHD, as it is a lifelong condition. Clinical notes from expert clinicians will also be considered as a source of unstructured text.

The inclusion of unstructured clinical interview data and clinical notes presents both a significant challenge and a unique opportunity. This highlights a key area where machine learning can add considerable value beyond traditional methods, potentially revolutionizing screening and diagnostic processes by providing objective measures through the analysis of text and speech features, thereby mitigating the inherent subjectivity of traditional standardized rating scales.

Care will be taken to ensure that all questionnaire and interview data are collected and stored securely, in accordance with ethical guidelines and data protection regulations.

3.1.3 Behavioral Feature Extraction

Data preprocessing is an indispensable initial step in machine learning workflows that prepares raw data for subsequent analysis and modeling. Behavioral characteristics will be derived from both structured questionnaires and unstructured interviews.

- **Symptom Counts/Scores:** Direct scores from standardized questionnaires (SRS, CBCL, Conners) reflecting specific symptoms of the DSM criteria (e.g., inattention, hyperactivity, impulsivity, social responsiveness).
- **Behavioral Descriptors:** Extraction of explicit behavioral patterns and characteristics mentioned in unstructured interviews and clinical notes, such as difficulties with attention, planning, organization, impulse control, social interaction deficits, repetitive behaviors, irritability, and eye contact.
- **Severity and Impairment:** Features indicating the severity of symptoms and their impact on daily functioning in multiple settings (e.g., work, education, relationships).
- **Retrospective Childhood Behaviors:** Specific features related to the onset and persistence of childhood symptoms, as ADHD is a lifelong condition.

Questionnaires provide quantifiable and standardized data. The machine learning model can leverage the statistical power of structured data.

3.2 Machine Learning Model Development and Training

This section outlines the selection of machine learning algorithms, strategies for model training and validation, and techniques to ensure model robustness, particularly concerning overfitting and imbalanced datasets common in medical research.

3.2.1 Algorithm Selection

A diverse set of machine learning and deep learning algorithms will be explored to identify the most effective models to classify ASD and ADHD. The selection includes both traditional machine learning and deep learning models, reflecting a strategic decision based on the nature of the data and the diagnostic goal. Although deep learning excels with complex high-dimensional data, such as raw EEG, traditional machine learning models can offer better interpretability, which is highly valued in medical contexts. The choice of algorithm is a critical design decision in medical AI; high precision alone is often insufficient, as clinicians need to understand why a diagnosis is made. Therefore, the methodology considers the trade-offs between model complexity, performance, and the ability to provide transparent and understandable explanations for model decisions, potentially using techniques like SHAP. This implies that the ”best” algorithm is determined not only by its accuracy but also by its clinical actionability and trustworthiness.

Traditional machine learning algorithms such as Random Forest (RF), XGBoost, CatBoost, LightGBM, Support Vector Machines (SVM), Linear Discriminant Analysis (LDA), and Logistic Regression will be considered. These models have demonstrated effectiveness in various medical diagnostic tasks and offer a balance of performance and interpretability. Random Forest, for example, has shown good accuracy with questionnaire data and is robust to both precision and recall.

Given the complexity and high dimension of EEG and advanced linguistic features, deep learning (DL) models, including Convolutional Neural Networks (CNNs) and ResNet-based architectures, will be investigated. Deep learning models have shown superior performance over traditional machine learning methods for the detection of ASD in some studies and are adept at capturing complex patterns in EEG data. The field of machine learning in mental health diagnosis is dynamic, with newer and more complex models showing promise for improving diagnostic accuracy and efficacy by capturing complex patterns. However, the continued utility of traditional machine learning suggests that, for certain data types (e.g. structured questionnaire data) or for specific interpretability needs, simpler models remain valuable. This requires a flexible and comparative approach in model selection.

The potential for combining multiple models (ensemble techniques) will be explored to leverage the strengths of different algorithms and potentially improve overall accuracy and robustness.

3.2.2 Training and Validation Strategies

Robust training and validation strategies are crucial to ensure the generalizability and reliability of the developed machine learning models. The emphasis on techniques such as k-fold cross-validation and train-test splits directly addresses the problem of generalizability in medical AI. A model that performs well only on the training data (overfitting) is clinically unreliable. In medical diagnosis, where models must perform reliably on unseen patient data, rigorous validation is not just good practice; it is a requirement of safety and efficacy. The choice of validation strategy directly impacts the confidence clinicians can place in the model's predictions, ensuring that the model is not merely memorizing the training data but learning true underlying patterns of the disorders.

The data set collected will be divided into training and testing sets, typically using an 80:20 or 90:10 ratio for training and testing, respectively. To mitigate bias and improve model generalization, k-fold cross-validation (e.g. 5-fold or 10-fold) will be extensively used. This involves randomly splitting the data into sub-disks.

k subsets are trained in subsets k-1, and the remaining subset is validated, all repeated k times. Leave-one-out cross-validation (LOOCV) may also be considered for smaller datasets. The size of the available data set influences the choice of validation strategy; for smaller datasets, more exhaustive methods like LOOCV might be preferred to maximize data utilization for training, while for larger datasets, k-fold cross-validation or simple splits of tests between tests are more computationally feasible. This reflects a pragmatic approach to methodology that adapts to real-world data limitations while striving for statistical rigor.

Optimal hyperparameters for each selected algorithm will be determined using techniques such as grid search or random search combined with cross-validation. For deep learning models, suitable loss functions (like cross-entropy loss for classification) will be selected, and optimization methods such as stochastic gradient descent (SGD) with backpropagation will be used to adjust the model's parameters.

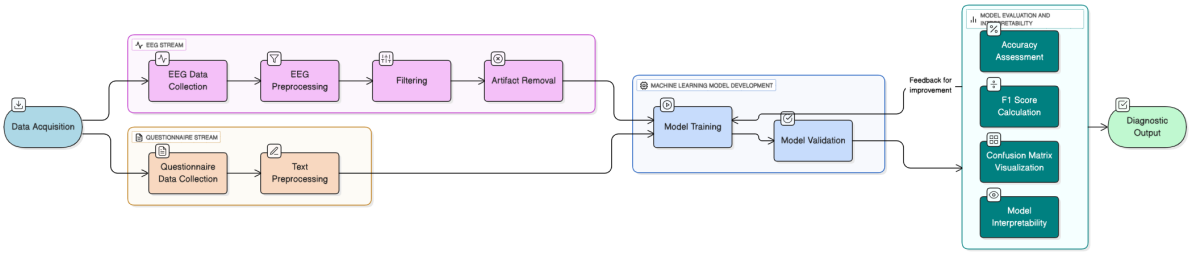

Figure 3-2 Experimental diagram of the proposed framework

3.3 Experimental Setup

This research outlines a comprehensive experimental framework designed for the analysis and detection of ASD and ADHD through a multimodal machine learning approach. The framework strategically integrates data derived from electroencephalography (EEG) signals and clinical text/questionnaires, leveraging their distinct but complementary information content to construct robust diagnostic models. The overall pipeline, as conceptually illustrated in Figure 1, is structured into several interconnected stages: Data Acquisition, Data Preprocessing, Machine Learning Model Development and Model Evaluation & Interpretability. Each stage is meticulously designed to ensure data quality, extract salient information, and develop predictive models that are both accurate and clinically relevant.

A visual representation of the proposed experimental pipeline shows in figure 3-2 would illustrate the sequential flow of data and processing steps. The diagram would commence with Data Acquisition, branching into two primary streams: EEG Data Collection and Questionnaire Data Collection. The raw data would then be fed into the machine learning model development stage, which encompasses model training and model validation. The output of this stage would lead to Model Evaluation & Interpretability, where performance metrics are assessed, and model decisions are explained. Finally, validated and interpreted models would produce diagnostic output for ASD and ADHD. This structured flow ensures a systematic approach from raw data to actionable diagnostic insights.

3.3.1 Evaluation Metrics

For medical diagnostic machine learning models, a comprehensive suite of evaluation metrics is necessary to provide a nuanced understanding of model effectiveness, especially given the potential for imbalanced data sets and the high stakes associated with diagnostic errors.

- **Accuracy:** This metric represents the overall proportion of correctly predicted instances

(both true positives and true negatives) out of the total instances. Although intuitive, precision can be misleading in unbalanced datasets. For example, if 90% of a data set consists of healthy controls, a model that simply predicts everyone as healthy would achieve 90% precision, despite not identifying any individuals with the disorder. Therefore, accuracy is most reliable when class distributions are roughly equal.

- **Precision (Positive Predictive Value - PPV):** Precision measures the proportion of correctly predicted positive cases (True Positives) out of all instances predicted as positive (True Positives + False Positives). It quantifies the model's ability to avoid false alarms. Precision is particularly important when False Positives are costly, such as in medical diagnosis where an incorrect positive diagnosis could lead to unnecessary and potentially harmful interventions or emotional distress for the patient and family.
- **Recall (Sensitivity):** Recall measures the proportion of actual positive cases that were correctly identified by the model (True Positives) out of all actual positive cases (True Positives + False Negatives). Quantifies the model's ability to find all relevant instances. Recall is critical in medical diagnosis because missing a disease (a False Negative) can have severe consequences, delaying necessary treatment or intervention.
- **Specificity:** Specificity measures the proportion of actual negative cases that were correctly identified by the model (True Negatives) out of all actual negative cases (True Negatives + False Positives). Indicates how well the model identifies people who do not have the condition.
- **F1-Score:** The F1-score is the harmonic mean of precision and recall, providing a balanced measure that considers both false positives and false negatives. It is particularly useful when both precision and recall are important, as is often the case in medical screening and diagnosis, where a balance between correctly identifying cases and avoiding false alarms is desired.

Chapter 4 Dataset Development

4.1 Overview of the Necessity for Robust and Relevant Datasets in Neurodevelopmental Disorder Research

The foundation of any robust machine learning application, particularly in the realm of medical diagnostics, rests critically on the quality and structure of its underlying dataset. For complex neurodevelopmental disorders (NDDs), such as Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD), the development of well-structured and high-quality datasets is not merely a technical prerequisite but a fundamental scientific endeavor. The inherent challenges in NDD diagnosis, which often depend on subjective clinical evaluations and behavioral observations, underscore the significant potential of objective data-driven approaches to enhance diagnostic precision and consistency. The generalizability and clinical utility of any subsequent machine learning model are directly dictated by the representativeness and integrity of the data used for its training and validation. Therefore, the meticulous development of such datasets forms a cornerstone in advancing diagnostic capabilities in this field.

4.2 Brief Introduction to the Multi-Modal Approach Combining Questionnaire and Neurophysiological Data

Recognizing the multifaceted nature of neurodevelopmental disorders, this research adopts a multimodal approach to dataset construction. This strategy involves integrating both behavioral and symptomatic data, derived from questionnaires, with neurophysiological data obtained through Electroencephalography (EEG). The rationale behind this fusion is to capture a more comprehensive profile of NDDs, leveraging the distinct strengths of each data type to overcome the inherent limitations of single-modality assessments. The design choice to combine these diverse data streams anticipates that a richer, combined feature set will offer superior discriminative power, ultimately leading to more accurate and nuanced diagnostic insights compared to approaches relying solely on one type of information. This integrative methodology reflects an understanding that NDDs manifest in various domains, from observable behaviors to underlying neural atypicalities, and a comprehensive data set should reflect this complexity.

4.3 Questionnaires Dataset: Design and Generation

4.3.1 Rationale for Developing a Synthetic Dataset for ASD and ADHD Detection

A significant challenge encountered in this research was the explicit absence of a comprehensive, pre-existing dataset or survey specifically designed for the detection of both ASD and ADHD. This data scarcity necessitated an innovative approach: the creation of a synthetic questionnaire dataset. This synthetic methodology allowed for the precise tailoring of questions to target specific diagnostic criteria and symptoms relevant to ASD, ADHD, and their differentiation from neurotypical development. While the creation of a synthetic dataset addresses the immediate need for data availability, it also introduces considerations regarding ecological validity, as the data does not originate from real-world patient populations. However, by meticulously designing questions based on established clinical sources, the aim was to mitigate this concern and ensure the dataset's relevance to clinical practice.

Characteristic	Description
Initial Dataset Shape	(100, 40) (100 samples, 40 features)
Post-SMOTE Dataset Shape	(1000, 40) (1000 samples, 40 features)
Number of Features	40 (questions)
Initial Labeling Scheme	0 = Normal, 1 = ASD, 2 = ADHD
Initial Class Assignment Logic	Class determined by the maximum value (0, 1, or 2) appearing among the 40 features.

4.3.2 Detailed Description of the Sources and Methodologies Used to Inform Question Design

To ensure clinical relevance and alignment of the synthetic questionnaire with recognized assessment tools, the design of the 40 questions was rigorously informed by established clinical and diagnostic resources. These included Mindmetric.ai (which provides ASD and ADHD tests), adhdcare.co.uk (specifically their Child Screening Test for ADHD Symptoms) and autismspeaks.org (referencing the revised Checklist for Autism in Toddlers, Revised (M-CHAT-RTM)). These reputable sources provided the foundational knowledge and established diagnostic criteria, ensuring that the questions developed were pertinent to the symptomatic domains charac-

teristic of ASD and ADHD.

4.3.3 Elaboration on the 40 Designed Questions and Their Relevance to NDD Symptoms

The synthetic dataset comprises 40 carefully designed questions, each targeting specific symptomatic domains crucial for the detection and differentiation of ASD and ADHD. Examples of these questions include: ”Having any problem to attention in regular activities?”, ”Having any problem to make eye-contact?”, ”Does he/she always become restless?”, ”Does he/she do some unusual posture or unusual facial expression?”, and ”Does he/she interpret others emotions well?”. Collectively, these questions cover a broad spectrum of symptomatic areas, encompassing attention deficits, social interaction difficulties, repetitive behaviors, hyperactivity, emotional dysregulation, and communication challenges—all of which are characteristic features of ASD, ADHD, or their co-occurrence.

4.3.4 Initial Dataset Characteristics, Including Shape and Labeling Scheme

The initial synthetic questionnaire data set was constructed with a shape of (100, 40), which signifies 100 individual samples and 40 distinct characteristics (questions). The labeling scheme for this initial data set was defined as follows: a value of ’0’ was assigned to a normal patient, ’1’ to an ASD patient and ’2’ to an ADHD patient. The classification logic for assigning these labels was based on the maximum value appearing among the 40 features, where the maximum of 0, 1, or 2 indicated the respective class of normal, ASD, or ADHD patient. This rule-based assignment provided the initial ground truth for the synthetic data.

4.3.5 Application of Synthetic Minority Over-sampling Technique (SMOTE) for Dataset Expansion and its Impact on Dataset Shape

To enhance the dataset’s utility for machine learning model training and to address potential limitations associated with small datasets, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. The original dataset of 100 samples was expanded tenfold, resulting in a new dataset shape of (1000, 40). The application of SMOTE serves to synthetically generate new samples for minority classes, thereby creating a more balanced and sufficiently sized dataset. This approach is critical for preventing machine learning models from being biased towards majority classes and for improving their generalization performance, particularly when real-world data collection is challenging or limited.

4.3.6 Identification of Dominating Questions for Classification

Within the questionnaires dataset, specific questions were identified as particularly influential in determining whether a patient was classified as normal, having ASD, ADHD, or both. These dominating questions include: ”Has he/she followed any special routines?”, ”Does he/she interpret others' emotions well?”, ”Can he/she eat all kinds of food?” and ”Does he/she have any mood swings?”. Identification of these questions provides early indications of the importance of features, suggesting that these particular symptomatic areas capture core diagnostic criteria or highly indicative behaviors for neurodevelopmental disorders. Such findings can inform the future questionnaire design, potentially leading to more efficient and targeted diagnostic tools.

4.4 Electroencephalography (EEG) Dataset: Acquisition and Preprocessing

4.4.1 Source of the EEG Data and Initial Channel Configurations

The electroencephalography (EEG) data used in this study for the detection of ASD and ADHD were sourced from Kaggle. This external acquisition means that the original data collection protocols and pre-processing steps were outside the direct control of the present study. The data initially comprised two distinct sets: one with 8 channels (C3, Cz, C4, CPz, P3, Pz, P4 and POz) and another with 19 channels (Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, Pz, Cz and Pz). The use of preexisting, potentially heterogeneous datasets from external sources requires careful documentation of their origin and any subsequent harmonization efforts to ensure consistency.

4.4.2 Methodology for Selecting Common EEG Channels for the Study

To create a unified and consistent EEG data set for the study, a critical data harmonization step was carried out. From the original 8- and 19-channel datasets, six common channels were meticulously selected. C3, Cz, C4, P3, Pz, and P4. This selection process was crucial to ensuring comparability among samples that may have come from different acquisition setups or studies. Although this approach ensures consistency, it also implies a potential reduction in the total information available by discarding channels that are not common in both original datasets. The choice to retain these specific common channels was probably driven by their established relevance in neurodevelopmental research and their robust signal quality.

Table 4.1 EEG Channel Selection and Dominance

Characteristic	Description
Original 8-Channel Configuration	C3, Cz, C4, CPz, P3, Pz, P4, POz
Original 19-Channel Configuration	Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, Pz, Cz, Pz
Selected Common Channels	C3, Cz, C4, P3, Pz, P4 (6 channels)
Final Processed EEG Dataset Shape	(1000, 6)
Dominating Channels for ASD Detection	C3, Cz, CPz, P3, Pz, POz
Dominating Channels for ADHD Detection	F3, C3, C4, P3, O2, F8, T7, T8, P7, P8

4.4.3 Description of the Processed EEG Dataset's Shape and Labeling Scheme

Following the selection of common channels, a new and unified EEG dataset was created. The final shape of this processed EEG dataset is (1000, 6), indicating 1000 samples and 6 features (EEG channels). Similarly to the questionnaire data set, the labeling scheme for this EEG data set was defined as: '0' for normal, '1' for ASD (Autism Spectrum Disorder) and '2' for ADHD (Attention-Deficit / Hyperactivity Disorder).

4.4.4 Discussion of EEG Channels Identified as Dominating for ASD and ADHD Detection

Analysis of EEG data revealed specific channels that exhibit greater discriminative power to detect ASD and ADHD. For ASD detection, the dominant channels identified include C3, Cz, CPz, P3, Pz, and POz. In contrast, for the detection of ADHD, a distinct set of channels was found to dominate. F3, C3, C4, P3, O2, F8, T7, T8, P7, and P8. The identification of these distinct, but sometimes overlapping (e.g., C3, P3) dominant channels for ASD versus ADHD suggests that these disorders may exhibit unique neurophysiological signatures. This finding, depicted in table 4.1 is significant in advancing the understanding of the neural underpinnings of these conditions and holds promise for the development of objective, targeted diagnostic biomarkers.

4.4.5 Methodology for Combining the Questionnaires and EEG Datasets

To take advantage of the complementary information from both behavioral and neurophysiological evaluations, the questionnaire dataset and the EEG dataset were combined to create

a new integrated dataset. This integration was performed at the feature level, where the 40 features derived from the questionnaires were concatenated with the 6 common EEG channel data points. This process effectively creates a richer, combined feature vector for each sample, aiming to provide a more holistic representation of an individual's neurodevelopmental profile.

4.4.6 Resulting Structure and Dimensions of the Combined Dataset

The outcome of this feature-level fusion is a new combined dataset with a shape of (1000, 46). This dimension clearly reflects the sum of the features from the two original datasets: 40 features from the questionnaires and 6 features from the EEG channels, totaling 46 features per sample.

4.4.7 Approach to Identifying and Handling Conflicting Rows Indicating Co-occurrence of ASD and ADHD

A crucial aspect of developing this integrated data set involved addressing the complex reality of comorbidity between ASD and ADHD. Upon combining the two datasets, 108 "conflicting rows" were specifically identified. These rows represent instances where a sample exhibited characteristics indicative of both autism spectrum disorder and attention deficit / hyperactivity disorder based on the combined data. The explicit identification and handling of these conflicting rows to form a distinct "Both" class is a sophisticated methodological and clinical refinement. This approach acknowledges the high rates of comorbidity commonly observed in clinical practice, making the data set more representative of real-world clinical presentations and significantly improving the diagnostic utility of the model for complex cases.

4.4.8 Establishment of the Final Four-Class Labeling Scheme

To accurately reflect nuanced diagnostic possibilities, including comorbidity, the combined data set was structured with a refined four-class labeling scheme. The four distinct classes present in this dataset are: '0' for Normal, '1' for ASD, '2' for ADHD, and '3' for 'Both' (indicating patients exhibiting characteristics of both ASD and ADHD). This multiclass labeling scheme allows for more granular and clinically relevant diagnostic outcomes, moving beyond simple binary classifications. The following table 4.2 summarized the final dataset.

Table 4.2 Combined Dataset Structure and Class Distribution

Characteristic	Description
Integration Method	Feature concatenation (40 questionnaire features + 6 EEG features)
Resulting Dataset Shape	(1000, 46)
Number of Questionnaire Features	40
Final Processed EEG Dataset Shape	(1000, 6)
Number of Questionnaire Features	40
Number of EEG Features	6
Total Features	46
Conflicting Rows Identified	108 (indicating both ASD and ADHD)
Final Class Distribution	0 = Normal, 1 = ASD, 2 = ADHD, 3 = Both

4.4.9 Rationale for Expecting Enhanced Detection Accuracy from the Combined Features

The core rationale underpinning the creation of this combined dataset is the expectation that integrating both questionnaire responses and EEG channel data will lead to more accurate detection of Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder. This assumption is rooted in the understanding that the information contained in the questionnaire and the EEG data is complementary rather than redundant. The fusion of these diverse data modalities provides a more holistic view of the patient's condition, potentially capturing subtle indicators that might be missed by single-modality approaches. This synergistic effect is hypothesized to create a more powerful predictive signal, ultimately leading to improved diagnostic performance and a more comprehensive understanding of these complex neurodevelopmental conditions.

4.5 Dataset Utility

4.5.1 Summary of the Comprehensive Dataset Development Process

The development of the data set for the analysis and detection of autism spectrum disorder and attention deficit hyperactivity disorder utilizing the research of machine learning algorithms involved a comprehensive and multifaceted process. This began with the independent creation

of a synthetic questionnaire data set, driven by the scarcity of complete existing data, and its expansion through SMOTE to enhance its robustness. During the same time, an EEG dataset was acquired from external sources and harmonized through meticulous selection of common channels. The culmination of this process was the intelligent integration of these two distinct modalities into a single, comprehensive data set. A critical methodological decision was the explicit identification and handling of samples that exhibited characteristics of both ASD and ADHD, leading to the refinement of the labeling scheme to include a distinct ”Both” class. These methodological decisions were crucial in establishing a robust foundation for subsequent machine learning analyzes.

4.5.2 Discussion on the Potential Utility and Implications of this Multi-Modal Dataset for Future Research in NDD Detection

This meticulously developed multimodal dataset, characterized by its detailed features and nuanced four-class labeling, represents a valuable resource for advancing research in the detection of neurodevelopmental disorders. Its design, particularly explicit accounting for ASD-ADHD comorbidity through the ”Both” class, is a significant advancement. This data set has substantial potential to serve as a benchmark for future research focusing on differential diagnosis and understanding the shared and distinct neural and behavioral markers of these often co-occurring conditions. By providing a rich, integrated data source, it can facilitate the development and testing of advanced machine learning models aimed at achieving more accurate, objective, and potentially earlier diagnostic tools. The ability to identify and differentiate between neurotypical, ASD, ADHD, and comorbid presentations offers a pathway toward more personalized and effective interventions, ultimately contributing to improved outcomes for individuals affected by these complex neurodevelopmental disorders.

Chapter 5 Model Evaluation

This chapter presents a comprehensive evaluation of the machine learning and deep learning models developed for the detection and classification of Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD), including their co-occurrence. The assessment encompasses models trained on three distinct datasets: a synthetic questionnaires dataset, an electroencephalography (EEG) dataset, and a combined multi-modal dataset integrating both. The primary objective is to analyze how effectively each model's predictions align with actual diagnostic labels, providing insights into their strengths, limitations, and overall performance in distinguishing between Normal, ASD, ADHD, and "Both" (ASD+ADHD) classifications. Visualizations and comparative analyses are included to illustrate the relative performance of the models across different data modalities.

5.1 Performance Evaluation Metrics

To rigorously evaluate the performance of the classification models, several key metrics were employed :

- **Accuracy**: This metric represents the ratio of correctly predicted instances to the total number of predictions made. Although providing a general measure of overall performance, it is often considered in conjunction with other metrics for a more complete assessment.

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

Precision: Precision quantifies the proportion of correctly predicted positive instances out of all instances predicted as positive. A high precision value indicates the model's effectiveness in accurately identifying the positive class with minimal false positives.

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

Recall (Sensitivity): Recall measures the proportion of correctly predicted positive instances out of the total number of actual positive instances. High recall signifies the model's ability to identify most true positive cases, minimizing false negatives.

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

F1-score: The F1-score is the harmonic mean of precision and recall, providing a balanced measure, particularly useful when there is an uneven class distribution. A higher F1-score reflects a better trade-off between precision and recall.

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

False Acceptance Rate (FAR): This metric indicates the proportion of incorrect acceptances, essentially the rate at which the model incorrectly classifies a negative instance as positive.

Where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

5.2 Model Performance on Questionnaires Dataset

Various machine learning and deep learning models, including Support Vector Machine (SVM), Naïve Bayes, K-Nearest Neighbors (KNN), Random Forest (RF), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), CNN + LSTM, Multi-Layer Perceptron (MLP) and Gated Recurrent Unit (GRU), were applied to the synthetic questionnaires dataset to classify subjects into four categories: Normal (0), ASD (1), ADHD (2), and Both (3).

The results are summarized in table 5.1. Among traditional machine learning models, Random Forest (RF) and Support Vector Machine (SVM) showed relatively strong and consistent performance, both achieving an overall accuracy of 67.5%. However, SVM notably did not detect cases of ADHD, resulting in zero precision, recall, and F1 score for that class. Naïve Bayes exhibited the lowest accuracy at 57.5% and a high FAR of 14.17%.

Deep learning models like CNN and CNN+LSTM performed similarly, each achieving 67% accuracy, but also struggled with ADHD detection, showing zero precision and recall for that class. LSTM alone performed less favorably, with 62% precision and a complete failure to classify ADHD and "Both" categories.

The multilayer perceptron (MLP) emerged as the dominant model for the questionnaire data set, achieving the highest overall precision of 68%. MLP maintained a good balance across all metrics and demonstrated relatively strong detection performance for the ADHD and 'Both' classes, which were generally challenging for other models. Furthermore, MLP had the low-

est false acceptance rate (10. 67%), indicating fewer false positives and higher classification reliability.

5.3 Model Performance on EEG Dataset

EEG data provide insight into brain electrical activity, making it a valuable tool for studying neurodevelopmental disorders. Various machine learning and deep learning models were applied to the EEG dataset to classify individuals into normal (0), ASD (1), ADHD (2) and Both (3) categories.

The performance of EEG data with machine learning techniques considered is shown in table 5.2. Most models demonstrated their strongest performance when classifying the normal (0) and ADHD (2) categories. For example, SVM achieved a precision of 0.58 and recall of 0.82 for the Normal class, and a precision of 0.44 and recall of 0.92 for ADHD. CNN showed comparable results for these classes. These trends suggest that Normal and ADHD classes possess more distinguishable EEG patterns that models are better able to learn.

In contrast, models struggled significantly with the ASD (1) and "Both" (3) classes. Several models, including CNN, LSTM, GRU, and CNN+LSTM, achieved zero recall and F1 scores for both ASD and "Both". This indicates that while these models might occasionally have predicted ASD, those predictions were incorrect. The MLP model was the only one to show some improvement in detecting the "Both" class, with a precision of 0.67, recall of 0.08, and F1-score of 0.15, though the recall still reflects a very low detection rate.

The overall classification accuracy across all models remained modest. The MLP model achieved the highest overall accuracy of 53%, closely followed by SVM, CNN, and CNN+LSTM, all at 52.5%. The False Acceptance Rate (FAR) was generally high across the board, with Naïve Bayes and KNN exhibiting the worst FARs.

5.4 Model Performance on Combined (Questionnaire + EEG) Dataset

To take advantage of the complementary information from both behavioral and neurophysiological evaluations, the questionnaire dataset and the EEG dataset were combined. This integration resulted in a new dataset with a shape of (1000, 46), combining 40 questionnaire features and 6 EEG channel data points. This combined dataset also explicitly accounts for 108 "conflicting rows" indicating patients with both ASD and ADHD, leading to a four-class labeling scheme: Normal (0), ASD (1), ADHD (2), and Both (3). The complete results are in table 5.3.

When applied to the combined dataset, the Random Forest (RF) model achieved the highest

accuracy of 81%. It demonstrated strong F1 scores across all four classes: Normal (0.85), ASD (0.83), ADHD (0.76) and Both (0.66), along with the lowest False Acceptance Rates (FAR), particularly for the ”Both” class (0.0225). Deep learning models like CNN and LSTM also performed well, achieving accuracies of 77% and 71. 5%, respectively, with CNN showing a high recall for the ADHD (0.91) and the ”Both” (0.90) classes. In contrast, models like SVM and MLP significantly underperformed, failing to identify ASD and ”Both” classes (F1-scores of 0.00) and showing low overall accuracy.

These results highlight that combining EEG with questionnaire data significantly improves classification performance, especially when using ensemble or deep learning approaches. The superior performance of Random Forest in the combined case is likely attributable to its ensemble nature, which allows it to better capture patterns from diverse feature types (structured questionnaire input and time series EEG data). RF robustly handles feature interactions and noise, and by averaging over many decision trees, it reduces overfitting and leverages the complementary strengths of both data sources, an advantage that simpler models like MLP or SVM may struggle with.

5.5 Confusion Matrix Analysis: Random Forest on Combined Dataset

Given its superior performance, the confusion matrix in figure 5-1 for the Random Forest classifier in the combined data set provides further insight into its classification behavior. The confusion matrix reveals that the Random Forest classifier generally performs well in detecting Normal, ASD, ADHD, and ”Both” classes.

- **Normal (Class 0)**: Out of 88 actual Normal cases, 73 were correctly predicted as Normal. However, 15 Normal cases were wrongly classified as ADHD.
- **ASD (Class 1)**: Out of 45 actual ASD cases, 38 were correctly predicted as ASD. Some confusion is present with the Normal class (3 cases) and the ”Both” class (4 cases).
- **ADHD (Class 2)**: Out of 45 actual ADHD cases, 38 were correctly predicted as ADHD. There was some confusion with the Normal class (7 cases).
- **Both (Class 3)**: Out of 22 actual ”Both” cases, 13 were correctly predicted. However, 8 ”Both” cases were misclassified as ASD, and 1 as ADHD.

The most challenging aspect was the detection of the ”Both” (ASD+ADHD) class. The confusion matrix indicates some overlap between Normal and ADHD, and between ”Both” and ASD. This suggests a need for more refined features to better distinguish these overlapping conditions.

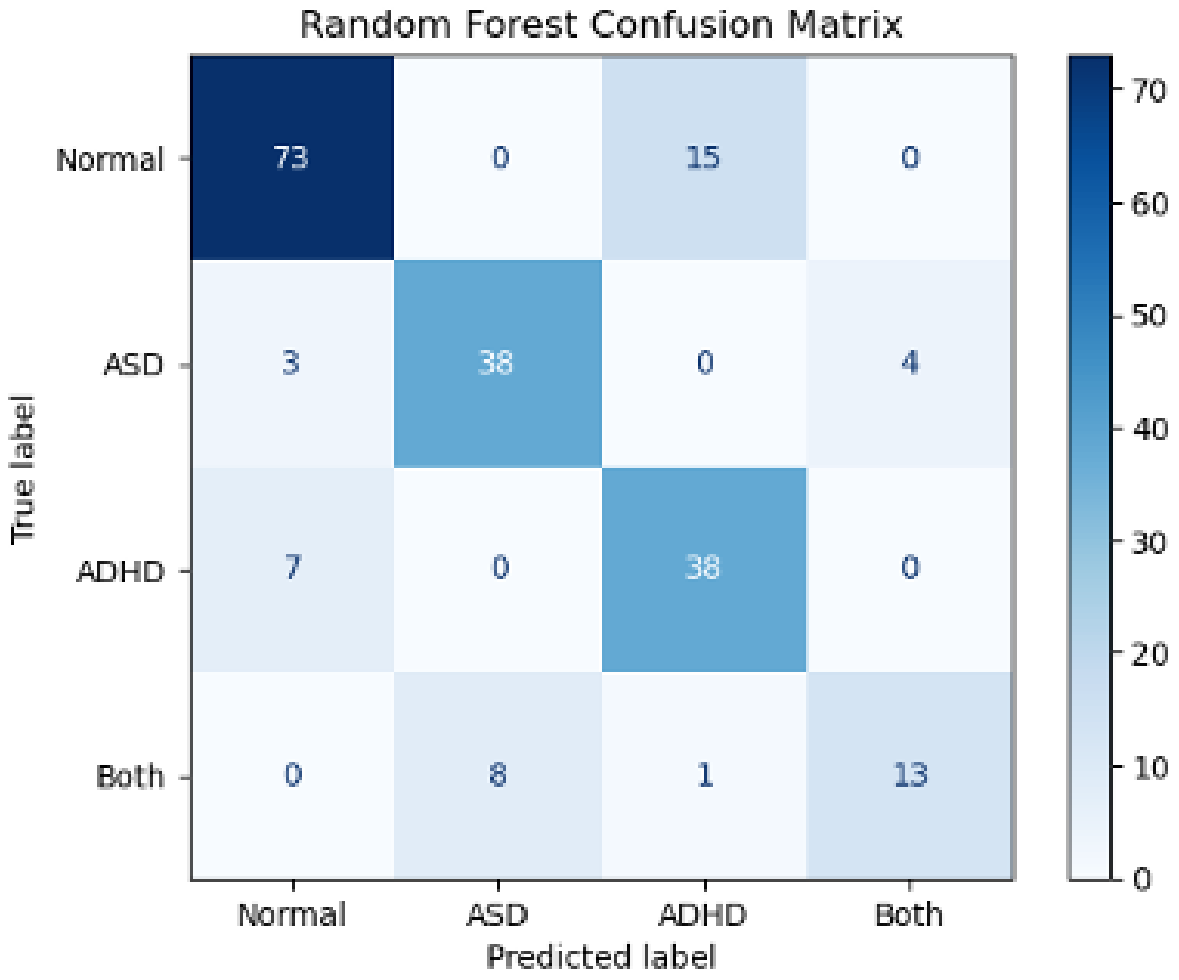

Figure 5-1 Confusion Matrix of Random Forest (Combined Dataset)

5.6 Overall Summary of Model Performance

The overall summary of the classification performance is shown in table 5.4.

When using only questionnaire data, the MLP model achieved the highest precision of 68%, performing particularly well to detect normal classes (F1 = 0.79) and ASD (F1 = 0.79). Similarly, with only EEG data, MLP again showed the best performance with accuracy 53%, demonstrating moderate F1 scores for Normal (0.70) and ADHD (0.57), although it did not detect ASD (F1 = 0.00).

However, when combining both the EEG and the questionnaire data, the Random Forest (RF) model significantly outperformed all others, achieving precision 81% and high F1 scores across all classes, especially Normal (0.85) and ASD (0.83). This improvement with RF in the combined case is probably due to its ensemble nature, which allows it to better capture patterns from the diverse types of characteristics (structured questionnaire inputs and time series EEG data). RF handles feature interactions and noise robustly, and by averaging over many decision trees, it reduces overfitting and leverages the complementary strengths of both data sources, something simpler models like MLP or SVM may struggle with. Hence, while MLP performs well individually on each data type, RF excels when the richness of both modalities is available.

5.7 Discussion of Findings

5.7.1 Overview of Comparative Performance Across Data Modalities

The comprehensive evaluation of various machine learning and deep learning models across three distinct datasets, questionnaires, electroencephalography (EEG), and a combined multimodal dataset, revealed significant insights into their efficacy for detecting Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD), including their co-occurrence. A clear trend emerged: the integration of diverse data modalities substantially improved classification performance, underscoring the complementary nature of behavioral and neurophysiological information in diagnosing complex neurodevelopmental disorders.

5.7.2 Performance Insights from Questionnaire Data

In the synthetic questionnaire data set, the multilayer perceptron (MLP) model demonstrated the most robust performance, achieving an overall precision of 68%. MLP exhibited balanced performance across various metrics, showing particular strength in detecting ”Normal” and ”ASD” classes, with F1 scores of 0.79 for both. Although other models such as Random

Forest (RF) and Support Vector Machine (SVM) also achieved respectable precision (67. 5%), a notable limitation was observed in their ability to accurately identify ADHD and the categories "Both" (ASD + ADHD). For example, SVM completely failed to detect ADHD cases, resulting in zero precision, recall, and F1 score for that class. This suggests that while questionnaire data can effectively capture distinct symptomatic profiles for "Normal" and "ASD" individuals, differentiating ADHD and especially comorbid presentations based solely on these self-report or observational features proved challenging for many models. The consistent difficulty in classifying ADHD and "Both" categories across most models points to either subtle symptom overlap or potential class imbalance issues within the synthetic questionnaire data.

5.7.3 Performance Insights from EEG Data

The analysis of EEG data presented a more challenging classification landscape. The overall precision was considerably lower compared to questionnaire-based models, with the MLP model again leading, although with a modest precision 53%. Although models showed reasonable performance in classifying "normal" and "ADHD" categories (e.g., MLP's F1-scores of 0.70 for Normal and 0.57 for ADHD), a critical finding was the widespread failure to detect "ASD" and "Both" classes. Several models, including CNN, LSTM, GRU, and CNN+LSTM, recorded zero recall and F1 scores for these categories, indicating that while they could have made predictions, none were correct. This suggests that the specific EEG features extracted or the inherent complexity of ASD and comorbid presentations within this dataset made their neural signatures less distinguishable for the models. The higher False Acceptance Rates (FARs) across the board for EEG-based classification further underscore the increased difficulty and potential for misclassification when relying solely on neurophysiological data for these specific diagnostic distinctions.

5.7.4 The Synergistic Power of Multi-Modal Data Fusion

The most compelling finding of this research lies in the significant performance uplift achieved by integrating both questionnaire and EEG data into a combined multi-modal dataset. The Random Forest (RF) model emerged as the main performer in this integrated data set, achieving impressive precision 81%. This represents a substantial improvement over the best individual modality performance (68% for questionnaires and 53% for EEG).

The superior performance of RF on the combined dataset can be attributed to its ensemble nature. Using the complementary strengths of both data sources, structured questionnaire responses that capture behavioral and symptomatic information, and time series EEG data that

reflect the underlying neural activity, RF was able to build a more robust and comprehensive understanding of neurodevelopmental conditions. Its ability to handle various feature types, manage feature interactions, and reduce overfitting by averaging across multiple decision trees allowed it to effectively capture complex patterns that single-modality approaches struggled with. Deep learning models like CNN also performed commendably on the combined dataset, achieving 77% accuracy and high recall for ADHD (0.91) and "Both" (0.90) classes, further supporting the value of multimodal integration.

Despite the overall improvement, the confusion matrix for the Random Forest model on the combined dataset revealed that the "Both" (ASD+ADHD) class remained the most challenging to detect, with some misclassifications occurring between the "Both" and "ASD" categories. This indicates that while the combined characteristics significantly improved the detection of comorbidity, more refined characteristics or more sophisticated modeling approaches might be necessary to fully unravel the overlap characteristics of ASD and ADHD in comorbid presentations.

5.7.5 Clinical Implications and Future Directions

The findings strongly advocate for the adoption of multimodal approaches in the machine learning-driven detection of neurodevelopmental disorders. Integration of questionnaire-based symptomatic data with objective neurophysiological markers from EEG provides a more holistic and accurate diagnostic framework than either modality alone. This has significant clinical implications, suggesting a pathway toward developing more objective, data-driven tools that can complement traditional subjective clinical assessments, potentially leading to earlier and more precise diagnoses of ASD and ADHD, including their often complex comorbid presentations.

Although the use of a synthetic questionnaire dataset addressed data scarcity, future research should prioritize the collection and validation of real-world, clinically diverse multimodal datasets to further enhance the generalizability and clinical utility of these models. Additionally, continued efforts to refine feature engineering, particularly for distinguishing comorbid conditions, and exploring advanced deep learning architectures capable of better handling the intricacies of multimodal data, will be crucial for advancing the field. The identified "dominating questions" in the questionnaires and specific EEG channels also provide valuable insights for future targeted data collection and feature selection.

Table 5.1 Classification Report for Questionnaires Data

Models	Class	Precision	Recall	F1-Score	Accuracy	FAR
SVM	Normal-0	0.66	0.99	0.79	67.5%	10.83%
	ASD-1	0.71	0.96	0.82		
	ADHD-2	0.00	0.00	0.00		
	Both-3	0.50	0.04	0.08		
Naïve Bayes	Normal-0	0.61	0.89	0.72	57.5%	14.17%
	ASD-1	0.61	0.66	0.64		
	ADHD-2	0.00	0.00	0.00		
	Both-3	0.36	0.21	0.26		
KNN	Normal-0	0.70	0.89	0.79	65%	11.67%
	ASD-1	0.67	0.89	0.76		
	ADHD-2	0.44	0.18	0.25		
	Both-3	0.14	0.04	0.06		
RF	Normal-0	0.71	0.86	0.77	67.5%	10.83%
	ASD-1	0.73	0.91	0.81		
	ADHD-2	0.43	0.23	0.30		
	Both-3	0.55	0.25	0.34		
CNN	Normal-0	0.68	0.99	0.81	67%	11%
	ASD-1	0.68	0.94	0.79		
	ADHD-2	0.00	0.00	0.00		
	Both-3	0.25	0.04	0.07		
LSTM	Normal-0	0.61	0.95	0.74	62%	12.67%
	ASD-1	0.64	0.83	0.72		
	ADHD-2	0.00	0.00	0.00		
	Both-3	0.00	0.00	0.00		
CNN+LSTM	Normal-0	0.68	0.99	0.81	67%	11%
	ASD-1	0.68	0.94	0.79		
	ADHD-2	0.00	0.00	0.00		
	Both-3	0.25	0.04	0.07		
MLP	Normal-0	0.69	0.93	0.79	68%	10.67%
	ASD-1	0.71	0.91	0.79		
	ADHD-2	0.58	0.18	0.27		
	Both-3	0.43	0.12	0.19		
GRU	Normal-0	0.66	0.96	0.79	66.5%	11.7%
	ASD-1	0.71	0.92	0.80		
	ADHD-2	0.40	0.05	0.09		
	Both-3	0.25	0.04	0.07		

Table 5.2 Classification Report for EEG Data

Models	Class	Precision	Recall	F1-Score	Accuracy	FAR
SVM	Normal-0	0.58	0.82	0.68	52.5%	15.83%
	ASD-1	0.00	0.00	0.00		
	ADHD-2	0.44	0.92	0.60		
	Both-3	0.00	0.00	0.00		
Naïve Bayes	Normal-0	0.50	0.92	0.65	48.5%	17.17%
	ASD-1	0.50	0.02	0.04		
	ADHD-2	0.49	0.44	0.46		
	Both-3	0.20	0.08	0.12		
KNN	Normal-0	0.50	0.77	0.61	47.5%	17.50%
	ASD-1	0.27	0.08	0.12		
	ADHD-2	0.49	0.56	0.52		
	Both-3	0.40	0.17	0.24		
RF	Normal-0	0.59	0.73	0.65	51%	16.17%
	ASD-1	0.44	0.13	0.20		
	ADHD-2	0.44	0.82	0.57		
	Both-3	0.38	0.12	0.19		
CNN	Normal-0	0.62	0.80	0.70	52.5%	15.83%
	ASD-1	0.00	0.00	0.00		
	ADHD-2	0.41	0.97	0.58		
	Both-3	0.00	0.00	0.00		
LSTM	Normal-0	0.61	0.80	0.69	51.5%	16.17%
	ASD-1	0.00	0.00	0.00		
	ADHD-2	0.40	0.92	0.55		
	Both-3	0.00	0.00	0.00		
CNN+LSTM	Normal-0	0.63	0.80	0.71	52%	15.83%
	ASD-1	0.00	0.00	0.00		
	ADHD-2	0.40	0.97	0.57		
	Both-3	0.00	0.00	0.00		
MLP	Normal-0	0.58	0.88	0.70	53%	15.67%
	ASD-1	0.00	0.00	0.00		
	ADHD-2	0.45	0.77	0.57		
	Both-3	0.67	0.08	0.15		
GRU	Normal-0	0.61	0.80	0.69	52%	16%
	ASD-1	0.00	0.00	0.00		
	ADHD-2	0.41	0.95	0.57		
	Both-3	0.00	0.00	0.00		

Table 5.3 Classification Report for Combined Data

Models	Class	Precision	Recall	F1-score	Accuracy	FAR
SVM	Normal-0	0.60	0.76	0.67	54.5%	0.3928
	ASD-1	0.00	0.00	0.00		
	ADHD-2	0.47	0.93	0.63		
	Both-3	0.00	0.00	0.00		
Naïve Bayes	Normal-0	0.72	0.83	0.77	66%	0.2500
	ASD-1	0.72	0.62	0.67		
	ADHD-2	0.55	0.51	0.53		
	Both-3	0.44	0.36	0.40		
KNN	Normal-0	0.70	0.74	0.72	62.5%	0.2500
	ASD-1	0.77	0.60	0.68		
	ADHD-2	0.48	0.62	0.54		
	Both-3	0.36	0.23	0.28		
RF	Normal-0	0.87	0.82	0.85	81%	0.0893
	ASD-1	0.82	0.84	0.83		
	ADHD-2	0.70	0.84	0.76		
	Both-3	0.76	0.59	0.66		
CNN	Normal-0	0.87	0.69	0.77	77%	0.0804
	ASD-1	0.94	0.71	0.81		
	ADHD-2	0.60	0.91	0.72		
	Both-3	0.71	0.90	0.80		
LSTM	Normal-0	0.87	0.71	0.78	71.5%	0.0804
	ASD-1	0.87	0.75	0.80		
	ADHD-2	0.51	0.73	0.60		
	Both-3	0.52	0.59	0.55		
CNN+LSTM	Normal-0	0.80	0.63	0.70	69.50%	0.1250
	ASD-1	0.83	0.57	0.68		
	ADHD-2	0.53	0.93	0.68		
	Both-3	0.71	0.68	0.69		
MLP	Normal-0	0.56	0.79	0.66	45.5%	0.4732
	ASD-1	0.00	0.00	0.00		
	ADHD-2	0.41	0.33	0.37		
	Both-3	0.31	0.59	0.41		
GRU	Normal-0	0.61	0.69	0.64	63%	0.3482
	ASD-1	0.00	0.00	0.00		
	ADHD-2	0.43	0.93	0.59		
	Both-3	0.25	0.04	0.07		

Table 5.4 Final Classification Report (Best Models per Data Type)

Data type	Class	Precision	Recall	F1-score	Accuracy	FAR
Questionnaire	Normal-0	0.69	0.93	0.79	68%	10.67%
	ASD-1	0.71	0.91	0.79		
	ADHD-2	0.58	0.18	0.27		
	Both-3	0.43	0.12	0.19		
EEG	Normal-0	0.58	0.88	0.70	53%	15.67%
	ASD-1	0.00	0.00	0.00		
	ADHD-2	0.45	0.77	0.57		
	Both-3	0.67	0.08	0.15		
Questionnaire + EEG	Normal-0	0.87	0.82	0.85	81%	0.0893
	ASD-1	0.82	0.84	0.83		
	ADHD-2	0.70	0.84	0.76		
	Both-3	0.76	0.59	0.66		

Chapter 6 Conclusion

This research successfully demonstrated the enhanced capability of machine learning models in detecting Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD), including their co-occurrence, through a multi-modal data fusion approach. By meticulously developing a synthetic questionnaire dataset and integrating it with an externally available electroencephalography (EEG) dataset, a comprehensive and nuanced data set was created, explicitly accounting for the comorbidity of ASD and ADHD.

The comparative analysis unequivocally showed that the models trained in the combined questionnaire and the EEG data significantly outperformed those trained in single modalities. Specifically, the Random Forest (RF) model achieved the highest precision of 81% in the integrated data set, demonstrating its robustness in handling different types of features and capturing complex patterns indicative of these neurodevelopmental conditions. Although individual modalities, particularly questionnaire data, showed promise with multi-layer perceptron (MLP) achieving 68% precision and EEG data with MLP achieving 53% accuracy, the synergistic effect of combining data streams proved crucial for improving overall diagnostic accuracy and, critically, for better identifying the challenging "Both" class (ASD + ADHD). The ensemble nature of Random Forest allowed it to effectively leverage the complementary strengths of structured questionnaire input and time series EEG data, leading to superior performance by robustly handling feature interactions and reducing overfitting.

This study underscores the immense potential of multi-modal machine learning in advancing the objective diagnosis of neurodevelopmental disorders. The findings provide a strong foundation for the development of more accurate and reliable diagnostic tools that can complement existing clinical practices, leading to earlier interventions and improved patient outcomes. The ability to identify and differentiate between neurotypical, ASD, ADHD, and comorbid presentations offers a pathway towards more personalized and effective interventions, ultimately contributing to improved outcomes for individuals affected by these complex neurodevelopmental disorders.

6.1 Future Directions

Building upon the insights gained from this research, several avenues for future work are identified.

- **Real-World Data Validation:** While the synthetic questionnaire data set addressed data scarcity, future efforts should prioritize the collection and validation of multimodal, clinically diverse, real-world datasets. This will be crucial in assessing the generalizability and clinical utility of the models developed in authentic diagnostic settings.
- **Advanced Feature Engineering:** Further research into more sophisticated feature engineering techniques, particularly to distinguish comorbid conditions such as ASD + ADHD. Exploring novel ways to extract discriminative patterns from both questionnaire responses and raw EEG signals could further refine classification accuracy.
- **Sophisticated Deep Learning Architectures:** Investigating more advanced deep learning architectures, such as attention mechanisms or graph neural networks, specifically designed to handle multimodal data fusion and capture intricate relationships between different data types, could yield further improvements.
- **Longitudinal Studies:** Incorporating longitudinal data to track developmental trajectories and observe changes in symptoms and neural markers over time could provide invaluable information for early detection and intervention strategies.
- **Explainable AI (XAI):** Future work should also focus on developing more interpretable models to understand which specific characteristics or combinations of characteristics contribute the most to a diagnosis. This would improve trust in AI-driven diagnostic tools and provide clinicians with actionable insights.
- **Larger and More Diverse Datasets:** Expanding the size of the data set and ensuring greater diversity in terms of age, sex, socioeconomic background, and severity of symptoms will be essential to create truly robust and universally applicable diagnostic models.

Bibliography

[1] S. J. Moon, J. Hwang, R. K. Kana, J. Torous, and J. W. Kim, "Accuracy of machine learning algorithms for the diagnosis of autism spectrum disorder: Systematic review and meta-analysis of brain magnetic resonance imaging studies," 12 2019. [Online]. Available: https://doi.org/10.2196/14108

[2] M. Kundu and M. S. Das, "Predicting autism spectrum disorder in infants using machine learning," *Journal of Physics Conference Series*, vol. 1362, pp. 12 018–12 018, 11 2019. [Online]. Available: https://doi.org/10.1088/1742-6596/1362/1/012018

[3] G. Ismail, K. Chesoli, G. Moni, and K. Gikunda, "Comparison of probabilistic deep learning methods for autism detection," *arXiv (Cornell University)*, 01 2023. [Online]. Available: https://arxiv.org/abs/2303.12707

[4] R. A. Rasul, P. Saha, D. Bala, S. Karim, I. Abdullah, and B. Saha, "An evaluation of machine learning approaches for early diagnosis of autism spectrum disorder," *arXiv (Cornell University)*, 01 2023. [Online]. Available: https://arxiv.org/abs/2309.11646

[5] N. Pavlidis, V. Perifanis, H. Briola, C. Nikolaidis, E. Katsiri, P. S. Efraimidis, and D. E. Filippidou, "Federated anomaly detection for early-stage diagnosis of autism spectrum disorders using serious game data," *arXiv (Cornell University)*, 10 2024. [Online]. Available: http://arxiv.org/abs/2410.20003

[6] H. Sharif and R. A. Khan, "A novel machine learning based framework for detection of autism spectrum disorder (asd)," *Applied Artificial Intelligence*, vol. 36, 11 2021. [Online]. Available: https://doi.org/10.1080/08839514.2021.2004655

[7] H. Abbas, F. Garberson, E. Glover, and D. P. Wall, "Machine learning approach for early detection of autism by combining questionnaire and home video screening," *arXiv (Cornell University)*, 01 2017. [Online]. Available: https://arxiv.org/abs/1703.06076

[8] ——, "Machine learning approach for early detection of autism by combining questionnaire and home video screening," *Journal of the American Medical Informatics Association*, vol. 25, pp. 1000–1007, 04 2018. [Online]. Available: https://doi.org/10.1093/jamia/ocy039

[9] C. Lord, M. Elsabbagh, G. Baird, and J. Veenstra-VanderWeele, "Autism spectrum disorder," pp. 508–520, 08 2018. [Online]. Available: https://doi.org/10.1016/s0140-6736(18)31129-2

[10] S. Jacob, J. J. Wolff, M. Steinbach, C. B. Doyle, V. Kumar, and J. T. Elison, "Neurodevelopmental heterogeneity and computational approaches for understanding autism," 02 2019. [Online]. Available: https://doi.org/10.1038/s41398-019-0390-0

[11] V. Ramesh and R. Assaf, "Detecting autism spectrum disorders with machine learning models using speech transcripts," *arXiv (Cornell University)*, 01 2021. [Online]. Available: https://arxiv.org/abs/2110.03281

[12] A. P. da Silva, Ítalla Maria Pinheiro Bezerra, T. P. C. Antunes, M. P. E. Cavalcanti, and L. C. de Abreu, "Applied behavioral analysis for the skill performance of children with autism spectrum disorder," *Frontiers in Psychiatry*, vol. 14, 04 2023. [Online]. Available: https://doi.org/10.3389/fpsyt.2023.1093252

[13] M. Ceroni, S. Rossi, G. Zerboni, E. Biglia, E. Soldini, A. Izzo, L. Morellini, and L. Sacco, "Attentive-executive functioning and compensatory strategies in adult adhd: A retrospective case series study," *Frontiers in Psychology*, vol. 13, 10 2022. [Online]. Available: https://doi.org/10.3389/fpsyg.2022.1015102

[14] T. Eslami, F. Almuqhim, J. S. Raiker, and F. Saeed, "Machine learning methods for diagnosing autism spectrum disorder and attention- deficit/hyperactivity disorder using functional and structural mri: A survey," 01 2021. [Online]. Available: https://doi.org/10.3389/fninf.2020.575999

[15] A. Dentz, C. Martin-Soelch, C. Fahim, A. Torsello, V. Parent, A. Ponsioen, M. Guay, S. Bioulac-Rogier, C. Clément, M. Bader, and L. Romo, "Non-pharmacological treatment of attention deficit disorder with or without hyperactivity (adhd). overview and report of the first international symposium on the non-pharmacological management of adhd," pp. 309–328, 02 2024. [Online]. Available: https://doi.org/10.1016/j.encep.2023.04.010

[16] S. S. Martínez, K. Stoyanov, and L. Carcache, "Unraveling the spectrum: overlap, distinctions, and nuances of adhd and asd in children," 09 2024. [Online]. Available: https://doi.org/10.3389/fpsyt.2024.1387179

[17] L. Reinhart, A. Bischops, J.-L. Kerth, M. Hagemeister, B. Heinrichs, S. B. Eickhoff, J. Dukart, K. Konrad, E. Mayatepek, and T. Meißner, "Artificial intelligence in child development monitoring: A systematic review on usage, outcomes and acceptance," pp. 100 134–100 134, 01 2024. [Online]. Available: https://doi.org/10.1016/j.ibmed.2024.100134

[18] N. Çavuş, A. A. Lawan, Z. Ibrahim, A. Dahiru, S. Tahir, U. I. Abdulrazak, and A. Hussaini, "A systematic literature review on the application of machine-learning models in behavioral assessment of autism spectrum disorder," pp. 299–299, 04 2021. [Online]. Available: https://doi.org/10.3390/jpm11040299

[19] M. D. Hossain, M. A. Kabir, A. Anwar, and M. Z. Islam, "Detecting autism spectrum disorder using machine learning techniques," *Health Information Science and Systems*, vol. 9, 04 2021. [Online]. Available: https://doi.org/10.1007/s13755-021-00145-9

[20] T. Eslami, V. Mirjalili, A. Fong, A. R. Laird, and F. Saeed, "Asd-diagnet: A hybrid learning approach for detection of autism spectrum disorder using fmri data," *Frontiers in Neuroinformatics*, vol. 13, 11 2019. [Online]. Available: https://doi.org/10.3389/fninf.2019.00070

[21] A. R. Bongurala, D. Save, and A. Virmani, "Revolutionizing attention deficit hyperactivity disorder with artificial intelligence," *PLOS mental health.*, vol. 1, 11 2024. [Online]. Available: https://doi.org/10.1371/journal.pmen.0000181

[22] A. Alhussen, A. I. Alutaibi, S. K. Sharma, A. R. Khan, F. Ahmad, and G. G. Tejani, "Early attention-deficit/hyperactivity disorder (adhd) with neurodct-ica and rhinofish optimization (rfo) algorithm based optimized adhd-attentionnet," *Scientific Reports*, vol. 15, 02 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-90649-1

[23] X. Wang, Q. Jia, L. Liang, W. Zhou, W. Yang, and J. Mu, "Artificial intelligence in adhd: a global perspective on research hotspots, trends and clinical applications," 04 2025. [Online]. Available: https://doi.org/10.3389/fnhum.2025.1577585

[24] D. Yan, S. Wu, M. T. Sami, A. Almudaifer, Z. Jiang, H. Chen, D. Rangaprakash, G. Deshpande, and Y. Ma, "Improving brain dysfunction prediction by gan: A functional-connectivity generator approach," in *2021 IEEE International Conference on Big Data (Big Data)*, 2021, pp. 1514–1522.

[25] M. Maniruzzaman, J. Shin, and M. A. M. Hasan, "Predicting children with adhd using behavioral activity: A machine learning analysis," *Applied Sciences*, vol. 12, no. 5, 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/5/2737

[26] D. Zhou, Z. Liao, and R. Chen, "[retracted] deep learning enabled diagnosis of children's adhd based on the big data of video screen long-range eeg," *Journal of Healthcare Engineering*, vol. 2022, no. 1, p. 5222136, 2022. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1155/2022/5222136

[27] H. Sadr, M. Nazari, Z. Khodaverdian, R. Farzan, S. Yousefzadeh-Chabok, M. T. Ashoobi,

H. Hemmati, A. Hendi, A. Ashraf, M. M. Pedram, M. Hasannejad-Bibalan, and M. R. Yamaghani, "Unveiling the potential of artificial intelligence in revolutionizing disease diagnosis and prediction: a comprehensive review of machine learning and deep learning approaches," 05 2025. [Online]. Available: https://doi.org/10.1186/s40001-025-02680-7

[28] A. N. Nielsen, D. M. Barch, S. E. Petersen, B. L. Schlaggar, and D. J. Greene, "Machine learning with neuroimaging: Evaluating its applications in psychiatry," pp. 791–798, 11 2019. [Online]. Available: https://doi.org/10.1016/j.bpsc.2019.11.007

[29] M. S. Qureshi, M. B. Qureshi, J. Asghar, F. Alam, and A. Aljarbouh, "[retracted] prediction and analysis of autism spectrum disorder using machine learning techniques," *Journal of Healthcare Engineering*, vol. 2023, 01 2023. [Online]. Available: https://doi.org/10.1155/2023/4853800

[30] Y. Hu, Z.-A. Huang, R. Liu, X. Xue, L. Song, and K. C. Tan, "A dual-stage pseudo-labeling method for the diagnosis of mental disorder on mri scans," 07 2022, pp. 1–8. [Online]. Available: https://doi.org/10.1109/ijcnn55064.2022.9892792

[31] A. Sharma, S. Arora, S. Kumar, and A. Bansal, "Ai ml enabled wearable smart sensors detecting psychological disorders," 01 2023, pp. 651–656. [Online]. Available: https://doi.org/10.1109/confluence56041.2023.10048832

[32] S. Raj and S. Masood, "Analysis and detection of autism spectrum disorder using machine learning techniques," *Procedia Computer Science*, vol. 167, pp. 994–1004, 01 2020. [Online]. Available: https://doi.org/10.1016/j.procs.2020.03.399

[33] M. Maniruzzaman, J. Shin, and M. A. M. Hasan, "Predicting children with adhd using behavioral activity: A machine learning analysis," *Applied Sciences*, vol. 12, pp. 2737–2737, 03 2022. [Online]. Available: https://doi.org/10.3390/app12052737

[34] F. Almuqhim and F. Saeed, "Asd-grestm: Deep learning framework for asd classification using gramian angular field," vol. 284, 12 2023, pp. 2837–2843. [Online]. Available: https://doi.org/10.1109/bibm58861.2023.10385743

[35] D. Zhou, Z. Liao, and R. Chen, "Deep learning enabled diagnosis of children's adhd based on the big data of video screen long-range eeg," *Journal of Healthcare Engineering*, vol. 2022, pp. 1–9, 04 2022. [Online]. Available: https://doi.org/10.1155/2022/5222136

[36] H. Alkahtani, T. H. H. Aldhyani, and M. Y. Alzahrani, "Deep learning algorithms to identify autism spectrum disorder in children-based facial landmarks," *Applied Sciences*, vol. 13, pp. 4855–4855, 04 2023. [Online]. Available: https://doi.org/10.3390/app13084855

[37] Z.-A. Huang, Z. Zhu, C. H. Yau, and K. C. Tan, "Identifying autism spectrum disorder from resting-state fmri using deep belief network," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, pp. 2847–2861, 07 2020. [Online]. Available: https://doi.org/10.1109/tnnls.2020.3007943

[38] D. Yan, W. Shengbin, M. T. Sami, A. I. Almudaifer, Z. Jiang, H. Chen, D. Rangaprakash, G. Deshpande, and Y. Ma, "Improving brain dysfunction prediction by gan: A functional-connectivity generator approach," 12 2021, pp. 1514–1522. [Online]. Available: https://doi.org/10.1109/bigdata52589.2021.9671402

[39] J. Han, G. Jiang, G. Ouyang, and X. Li, "A multimodal approach for identifying autism spectrum disorders in children," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 2003–2011, 01 2022. [Online]. Available: https://doi.org/10.1109/tnsre.2022.3192431

[40] M. K. S. Rani, S. T. Fathima, S. L. Gayathri, M. Pavani, and Y. Swathi, "Detection of autism spectrum disorder using machine learning," *International Journal of Advanced Research in Science Communication and Technology*, pp. 71–74, 11 2022. [Online]. Available: https:

//doi.org/10.48175/ijarsct-7600

[41] K. Niu, J. Guo, Y. Pan, X. Gao, X. Peng, N. Li, and H. Li, "Multichannel deep attention neural networks for the classification of autism spectrum disorder using neuroimaging and personal characteristic data," *Complexity*, vol. 2020, pp. 1–9, 01 2020. [Online]. Available: https://doi.org/10.1155/2020/1357853

[42] C. Küpper, S. Stroth, N. Wolff, F. Hauck, N. Kliewer, T. Schad-Hansjosten, I. Kamp-Becker, L. Poustka, V. Roessner, K. Schultebraucks, and S. Roepke, "Identifying predictive features of autism spectrum disorders in a clinical sample of adolescents and adults using machine learning," *Scientific Reports*, vol. 10, 03 2020. [Online]. Available: https://doi.org/10.1038/s41598-020-61607-w

[43] Y. B. Mao, X. Qi, L. He, S. Wang, Z. Wang, and F. Wang, "Advanced machine learning techniques reveal multidimensional eeg abnormalities in children with adhd: a framework for automatic diagnosis," *Frontiers in Psychiatry*, vol. 16, 02 2025. [Online]. Available: https://doi.org/10.3389/fpsyt.2025.1475936

[44] S. Alhassan, A. Soudani, and M. AlMusallam, "Energy-efficient eeg-based scheme for autism spectrum disorder detection using wearable sensors," *Sensors*, vol. 23, pp. 2228–2228, 02 2023. [Online]. Available: https://doi.org/10.3390/s23042228

Researcher Seal and Signature _____

Date: