STATE-OF-THE-ART OF MACHINE LEARNING IN NEURO DEVELOPMENT DISORDER: A SYSTEMATIC REVIEW

Lilian Lee Yen Wei [1,2], Ag Asri Ag Ibrahim[1] and Rayner Alfred[1]

(Received: 15-Jul.-2024, Revised: 18-Sep.-2024, Accepted: 2-Oct.-2024)

ABSTRACT

This paper presents a comprehensive literature review focusing on the utilization of machine-learning (ML) and deep-learning (DL) methods for predicting and detecting Neurodevelopmental Disorders (NDDs), such as Intellectual Disability (ID), Autism Spectrum Disorder (ASD), Attention Deficit Hyperactivity Disorder (ADHD), Dyslexia, among others. While existing reviews often lack detailed discussions on the specific ML algorithms, datasets and performance metrics employed in NDD prediction and detection, this study aims to address this gap by examining two primary aspects: prediction and detection. Objective: The objective of this study is to investigate the current state-of-the-art methodologies, challenges and future directions in leveraging ML and DL techniques for the prediction and detection of NDDs. It aims to categorize the literature based on these two major aspects and provide insights into the various approaches, datasets, parameters and performance measures used in previous research. Methodology: This review encompasses articles published in journals and conference proceedings indexed in Scopus from 2013 to 2023. The search employed terms such as "Predicting Neurodevelopmental Disorder" and/or "Detection of Disorder Using Machine Learning." The analysis focuses on identifying common ML and DL approaches, ensemble models, types of datasets utilized, as well as the parameters and performance metrics employed in NDD-prediction and detection studies. Results: The findings of this review shed light on prevalent ML and DL methodologies, the challenges encountered and potential avenues for future research aimed at enhancing services for the NDD community through improved prediction and detection techniques.

KEYWORDS

Detection, Prediction, Classification, Deep learning, Machine learning, Mental health, Neurodevelopment disorders.

1. INTRODUCTION

Neurodevelopmental Disorders (NDDs), as outlined in the DSM V Diagnostic and Statistical Manual by the American Psychiatric Association, encompass a range of conditions affecting the development of the central nervous system [1]. These conditions manifest in difficulties in behaviours, cognition, social interaction and emotional functioning. Included within NDDs are intellectual disability (ID), communication disorders, Autism Spectrum Disorder (ASD), Attention-Deficit/Hyperactivity Disorder (ADHD), neurodevelopmental motor disorders such as Tic Disorders and Specific Learning Disorders [2]. Despite the prevalence of NDDs, which affects roughly 17% of the general population, many individuals may remain undiagnosed. Factors contributing to NDDs include maternal and fetal genotype, early environmental influences and some causes that are still not fully understood. Particularly concerning is the rising prevalence of NDDs, with autism rates reported by the Centre for Disease Control and Prevalence (CDC) increasing from 1 in 150 children in 2000 to 1 in 36 presently, with around 40% of affected individuals also experiencing ADHD and other comorbidities [3]. NDDs represent a significant mental-health category with profound impacts on daily functioning, potentially jeopardizing the physical and mental well-being of affected individuals as they transition from childhood to adulthood. Given the increasing frequency of NDDs and their substantial impact, it is imperative to address the challenges associated with early identification and intervention. Developing a rapid, reliable and automated method for identifying early signs of mental-health issues is critical in this rapidly evolving world.

Hence, we conduct a systematic review encompassing medical and computer-science literature on the

1. Lilian L.Y. W., A. A. A. Ibrahim and R. Alfred are with Faculty of Computing and Informatics, Universiti Malaysia Sabah, 88400, Kota Kinabalu, Sabah, Malaysia. Emails: lilian_lee_di22@iluv.ums.edu.my, awgasri@ums.edu.my, ralfred@ums.edu.my
2. Lilian L.Y.W. is with Malaysian Administration Modernization and Management Planning Unit, Prime Minister's Department, Malaysia. Email: lilian.phoebe@gmail.com

detection of NDD issues using machine-learning (AI) methodologies. This review spans articles published between 2013 and 2023 sourced from databases such as Scopus, IEEE Explore, PubMed and Web of Science (WOS). Utilizing the Preferred Reporting Items for Systematic Reviews and Meta- Analysis (PRISMA) methodology, we meticulously selected 81 articles from an initial pool of 811.

Our review highlights a notable research gap concerning the utilization of machine learning in interventions for neurodevelopmental disorders (NDDs), particularly in the domain of automated neuro-feedback. We also explore the machine learning techniques utilized in developing EEG-based detection methods for NDDs. Furthermore, we conduct a thorough examination of challenges outlined in existing literature and provide forward-looking recommendations. These recommendations encompass various facets, such as data-fusion techniques, integrating hybrid classification models, emphasizing the importance of publicly available datasets, addressing uncertainties in model predictions, enhancing model interpretability and devising strategies for hardware implementation. Essentially, our systematic review illuminates the current landscape of machine-learning applications in NDD detection and intervention, while also charting a course for future research aimed at bridging existing gaps and overcoming challenges. The objective of this review is to guide future researchers in adopting potential trends or models that can significantly enhance the diagnosis and detection of NDDs.

Responding promptly to a diagnosis is crucial for minimizing the time required for intervention once a diagnosis is detected. Machine-learning techniques can assimilate and analyze integrated data from multiple sources, including population statistics, lifestyle factors and medical records to predict the occurrence and distribution of diagnoses within a specific area. Medical practitioners can utilize machine-learning methods to enhance the implementation of existing interventions and speed up in developing new interventions. For example, deep-learning algorithms can be employed to analyze extensive datasets comprising medical information gathered from hospitals. For example, clinical test data from patients diagnosed with mental health can be utilized as input for machine-learning models, enabling doctors to expedite diagnoses. This research endeavours to explore the current advancements, obstacles and prospective directions in leveraging machine-learning techniques for managing neurodevelopmental disorders, as outlined in the two previously mentioned categories. The study uses the same method of systematic review conducted by Rayner & Obit, 2021, the roles of machine learning methods in limiting the spread of deadly diseases. Thus, the work here is to conduct a comprehensive review of different methodologies, dataset types, parameters or variables, individual and ensemble models, performance metrics and approaches employed in prior studies [4].

We categorized all articles and conference papers based on Scopus Indexed – whether pertaining to prediction or detection strategies. This review's results center on frequently employed machine learning methods, obstacles encountered and future directions aimed at supporting intervention and therapy for neurodevelopmental disorders through both detection and prediction. The trend and distribution of objectives for machine learning and recent works for NDD detection are described in Figure 1.

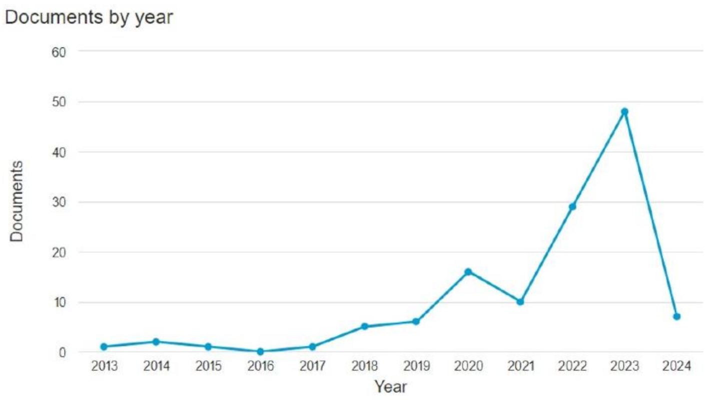

Figure 1. The trend and distribution of objectives for machine learning and recent works for NDD detection.

2. Methodology

The purpose of this Systematic Literature Review (SLR) is to conduct a sequential process of PRISMA to make available research applicable to machine-learning approaches in assisting medical health diagnosing Neurodevelopment Disorders. Four primary stages of PRISMA are identified to be included in this SLR as shown in Figure 2. They are called: Identification, Screening, Eligibility and Inclusion [5].

2.1 Content Retrieval

Apart from adhering to the PRISMA stages, this literature review underwent two distinct phases: planning and conducting. The initial phase is geared towards defining the prerequisites for a systematic review while mitigating potential researcher biases. It involves crafting a comprehensive review protocol, acting as a blueprint for conducting an unbiased review process. Key elements of this proposed review protocol in our study include delineating research questions, formulating a search strategy to pinpoint relevant studies, specifying inclusion and exclusion criteria, establishing a method for assessing study quality and extracting and synthesizing data, all of which will be detailed in the subsequent section. The planning phase involves crafting research questions centered on employing machine-learning techniques for predicting and detecting neurodevelopmental disorders, followed by setting up suitable search procedures to efficiently execute the research activities. During the conducting phase, several actions are taken, including setting predetermined selection criteria to pinpoint relevant studies and assessing their quality using the predefined quality-assessment procedure outlined in this study. This phase involves extracting information from the selected studies and conducting data synthesis to provide a succinct summary of the reviews. These processes are visually depicted in Figure 2, facilitating the incorporation of new information into the report in the future.

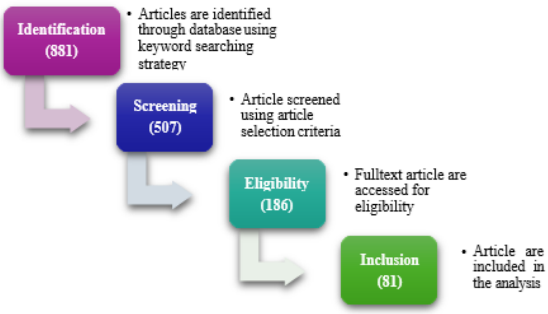

Figure 2. PRISMA method.

2.1.1 Formulating Research Questions

The research questions (RQs) were structured to define the study's boundaries from three distinct angles: population, intervention and outcomes [5]. From the population perspective, the focus is on the domains or functions affected by the intervention, such as detection, prediction and classification. These populations may relate to various aspects, including specific machine-learning methods or types of machine learning models and their applications. The intervention viewpoint centers on machine-learning approaches addressing specific challenges, such as diagnosing, detecting and predicting Neurodevelopment Disorders. Lastly, the outcomes perspective concerns factors significant to practitioners, such as improved prediction accuracy, reduced diagnostic costs for specific disorders and shortened response time in detecting potentially severe disorders. All relevant outcomes must be explicitly stated. For instance, interventions may aim to enhance one aspect of NDD prediction without affecting another, such as improving reliability without increasing costs. The primary goal of this Systematic Literature Review (SLR) is to gather and scrutinize relevant evidence to tackle the defined research questions (RQs). Our motivation for undertaking this endeavor is to obtain responses to a series of seven RQs, aiming to gain deeper insights into key aspects of our research focus. This entails enhancing our comprehension of the roles played by machine-learning technologies in facilitating the prediction and detection of Neurodevelopment Disorders, as well as identifying research constraints to guide future-research directions. The RQs and their rationale are thoroughly elaborated in Table 1.

Table 1. Research questions.

ID	Research Question
RQ1	What are the roles of machine-learning models in assisting in screening neurodevelopment disorders?
RQ2	What types of NDD datasets in previous works have been used to build the models? What types of parameters or variables have been used?
RQ3	What types of problems addressed using these models?
RQ4	Which individual models achieved the highest performance?
RQ5	What evaluation metrics and methods are employed to measure the performance of the machine-learning models?
RQ6	What types of ensemble methods are used in machine-learning models?
RQ7	What types of deep-learning approaches used in NDD Detection?

2.1.2 Search Process

In the identification stage, all publications up to Dec. 2023 were compiled from searches made in Scopus, IEEE Explore, PubMed and Web of Science (WOS) databases. The retrieval was performed for articles from journals and conference proceedings published from 2010 to 2023 using the following Boolean search expression: "Prediction" OR "Detection" OR "Classification" OR "Diagnosing" OR "Identification" AND "ADHD" OR "AUTISM" OR "DYSLEXIA" OR "Neuro Development Disorder" AND "Artificial Intelligence" OR "Machine Learning" OR "Deep Learning".

Only final papers were considered in this review. The inclusion and exclusion criteria are shown in Table 2. The search process is designed to thoroughly address all predefined research questions. This involves selecting appropriate digital libraries, setting a time frame for the published articles and defining the search keywords. We will explore six of the most popular and largest online digital libraries in computer science, along with the Medline digital library, which publishes peer-reviewed articles. These digital libraries are listed in Table 3.

Table 2. Criteria of inclusion and exclusion.

Criteria	Inclusion	Exclusion
Type of Article	Journal articles	Others (Thesis, Handbook, Literature Review and Survey Paper)
Language	English	Non-English
Subjects Covered	Computer Science, Neuroscience, Health Professional and Psychology.	Multidisciplinary
Year of Publications	2013-2023	< 2013
Domain	Mental Health	Other Disorder or Comorbidities
Mental Health	Neuro Development Disorder (ASD, ADHD, Dyslexia)	Other neurodisorders Listed in DSM-V (Schizophrenia, psychotic, bipolar, depression, …etc.)
ML Models	Traditional, Deep Learning and Ensemble Model	Transfer Learning
Dataset Type	• Demographic, Medical, Observation & Behavioural Data • Facial Image Data • Eye-Tracking Data • EEG-based Data • Functional Magnetic Resonance Imaging (BMRI) and Functional Magnetic Resonance Imaging (FMRI) Data	Genetic Data Heart-rate Data Handwriting Data Speech Data

Table 3. Online digital libraries and number of studies screened and reviewed.

No.	Database	URL	Screened	Eligible	Inclusion
1	Elsevier	https://www.sciencedirect.com/	68	14	7
2	Springer	https://link.springer.com/	56	28	2
3	IEEE eXplore	https://ieeexplore.ieee.org/	197	73	35
4	MDPI	https://www.mdpi.com/	92	29	9
5	Wiley	https://onlinelibrary.wiley.com/	12	5	2
6	Medline (PubMed)	https://pubmed.ncbi.nlm.nih.gov/	82	37	26
	TOTAL		507	186	81

Furthermore, we reviewed various independent relevant journals and conference proceedings in the field of artificial intelligence, as detailed in Table 3. The search is limited to articles published between 2013 and 2023. This time frame was chosen, because machine learning has been extensively applied to problems related to Neurodevelopment Disorder (NDD) since the 2010s. Table 4 lists the number of studies reviewed based on year (2013 - 2023). Therefore, this paper aims to systematically summarize artificial-intelligence methodologies, encompassing both machine-learning and deep-learning techniques, applied in the prediction and detection in-response to neurodevelopmental disorders (NDDs).

Table 4. Number of studies reviewed based on year (2013 - 2023).

Year	Studies
2013 - 2017	6
2018	3
2019	4
2020	9
2021	14
2022	29
2023	16
Total	81

Table 5. Type of machine-learning problems and related studies.

Problems	*Roles*	*Related Studies*
Regression	Predict Neurodevelopment Disorder	[6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23]
Classification	Detect Neurodevelopment Disorder	[24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86]

2.2 Content Summarization

Quantitative data was extracted from the chosen studies focusing on the research questions outlined and the findings are presented in Tables 5-12 and Figure 3.

2.2.1 Roles of Machine-learning Model

Predicting and detecting NDDs plays crucial roles in enhancing healthcare systems. The tasks primarily involve predicting NDDs or modeling disorder frequencies using regression methods. Conversely, machine learning models predominantly address classification problems in detecting NDD. Table 5 summarizes all studies focused on NDD prediction and detection.

2.2.2 Type of Datasets and Parameters Used

Table 6 summarizes structured data utilized for predicting and detecting the diagnosis of NDDs, as well as the number of studies conducted. From the collective findings of these studies, six sets of structured datasets were the most employed. Structured databases encompass Demographic Data, Medical Data, Observation & Behavioral Data, Visual Video Data, Meteorological Data, MRI (including BMRI and FMRI), face-expression data, eye-tracking data and EEG-based data. Demographic data includes information on Age, Gender, Race and Ethnicity. Medical data involves a systematic analysis of a child's conditions, incorporating parameters, such as head measurements, weight, height, signs and symptoms of the disorder and treatment information. Observation & Behavioral Data entail numerical representations obtained from responses, speech, cognitive abilities, quotient scores and questionnaire assessments, like M-CHART, Q-Chat, AQ-10, ADI-R and ADOS Screening, UCI repository, IQ score, NCHS survey data, SDS ASDTest, OBTest, UK’s National Health Service (NHS), PAAS India and Scale data Questionnaires from Germany Clinic. Some of the observation and behavioral subjects were captured in video for further investigation. Visual Video Data captures activities involving a child during intervention or therapy sessions, focusing on parameters, such as movement, behavior, sensory perception, angle, direction and speed. Magnetic Resonance Imaging Data (BMRI and FMRI) aids in detecting and monitoring brain characteristics, particularly changes in blood flow. Some BMRI/FMRI datasets are publicly available through ABIDE (ADHD-200, Craddock 400 (CC400), …etc.). Common parameters for detection include normalized region volume, reduced corpus callosum volume and increased amygdala volume. Facial Image Data analyzes emotions through facial expressions, categorizing them as Happy, Sad, Angry or Neutral. Facial analysis employs an arousal- valence model to assess parameters, such as positive-active and negative-passive readings. These datasets can be collected through public databases, like Kaggle (KDEF dataset, …etc.). Finally, EEG-based Data, recorded *via* devices like BCI, captures spectral power of EEG signals, including Beta, Alpha, Theta, Gamma and Delta waves. Analysis often involves spectral, temporal, spatial or time-frequency features, revealing specific brain activities, ERPs, recurrences and transitional states. The EEG dataset is available in Dataport IEEE, …etc. Eye-tracking Data collected through EyeGaze apps, Eye movement, Automatic retinal-image analysis (ARIA) monitors parameters, like retina movement and pupil size.

2.3 Reporting of Review Findings

The summary of findings in the review was derived from the selected studies, focusing on the defined research questions. The overall State-of-the-Art of Neurodevelopment Disorder Prediction and Detection is described in Figure 3.

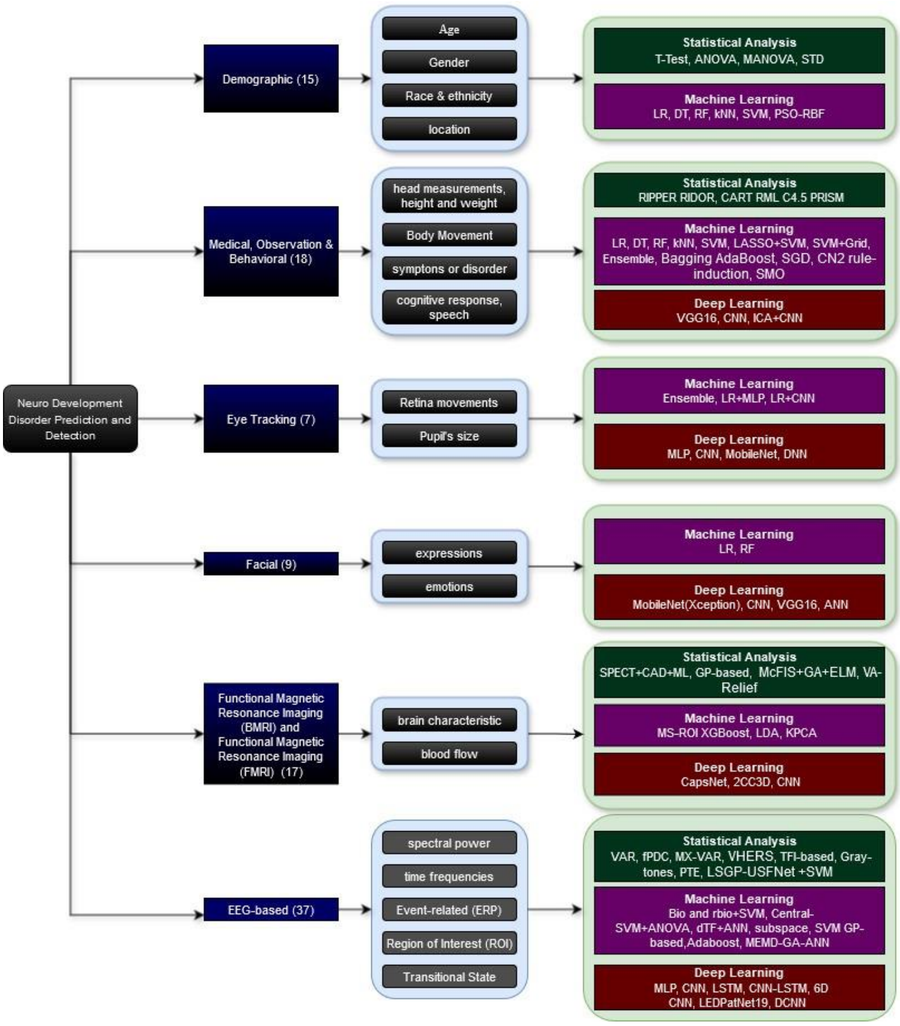

Figure 3. State-of-the-art of neurodevelopment disorder prediction and detection.

2.3.1 Roles of Machine-learning Models

This sub-section summarizes and discusses the findings of RQ1: What are the roles of machine-learning models in assisting in screening neurodevelopment disorders? The roles of machine-learning models can be categorized into regression and classification.

Table 6. Datasets and parameters used.

Databases	*(Frequency) Features*
Demographic Data (18)	Age, Gender, Race, Ethnicity, Location
Medical Observation & Behavioral Data (16)	Head measurements, Weight, Height, Response, Speech, Cognitive Ability, Quotients scores, Movement, Behavior, Sensory, Angle, Direction, Speed, …etc.
BMRI/FMRI Data (13)	Brain and Functional Magnetic Resonance Imaging: Changing in the blood flow within brain
Facial Image Data (7)	Facial features and expression that can be detect through facial emotion: Happy, Sad, Angry, Null
EEG-based Data (26)	EEG reading: Beta, Alpha, Theta, Gamma, Delta Other features (spectral, time, spatial or time-frequency features, activity, Event Related Potential (ERP), Recurrences, transition state)
Eye Tracking Data (5)	retina movement, pupil size

2.3.2 Regression Problems for Detecting NDDs

While logistic regression is commonly used for binary target variable datasets, it may not be accurate when the sample size is small. The connection between the predictors and a categorical response variable is modeled using logistic regression [58]. Regression problems are commonly addressed in the task of predicting or modeling the disorders as shown in Table 7.

Table 7. Regression: Types of machine-learning approaches and individual models used.

Study	Model	Best Model with Accuracy	Additional Performance Metrics
[6]	LDA, Ensemble (SVM-RBF, regression, Fuzzy sets, SVM+RBF	Cubic Regression method+SVM+RBF 98%	_
[7]	LR, SVM Polynomial, SVM RBF, NB, DT, RF, XGB, RF	LR 90.27%(ASD), RF 80.89% (Dyslexia)	PCS 92.30%, REC 90%, AUC 91.60%, F1-S 91.13%, CV 99.77% (ASD) PCS 83.56%, REC 77.21%, AUC 79.78%, F1-S 80.26%, CV 95.05% (Dyslexia)
[8]	DT, LR, RF	LR +CNN 81%	-
[9]	NB, LR, SVM	LR 95.87% (Adolescent), LR 99.82% (Adult), SVM 97.82% (Toddler), SVM 99.61% (Child)	KP 91.74%, F1-S 95.90%, AUROC 99.00% (Adolescent) KP 99.59%, F1-S 99.90%, AUROC 99.80% (Adult) KP 94.87%, F1-S 97.80%, AUROC 99.70% (Toddler) KP 99.21%, F1-S 99.60%, AUROC 99.60% (Child)
[10]	LR, MLP, CNN	∗MLP+LR 81%	PCS 77%, REC 78%, F1-S 76%
[11]	LR, SVM, polynomial regression, RF, MLP	Polynomial Regression 92.6%	PCS 91%, REC 89%, F1-S 92%
[12]	SVM, NB, DT, VGG16, DenseNet, AlexNet	LR 97.15%, RF 82%	-
[13]	LR, kNN, SVM, NB, DT, RF	LR 98.11%, SVM 98.11%, kNN 96.22%, NB 96.22%	F1-S, SP, SE
[14]	LR, MLP and SVC	LR 82.26%, SVC + PSO 93.55%	F1-S, SP, SE
[16]	NB, LR	LR 94.23% (Adults), LR 99.85% (Adolescent), LR 97.94% (Child)	SE 99.90%, SP 99.70% (Adults), SE 92.20%, SP 92.68% (Adolescent), SE 98%, SP 97.35% (Child)
[15]	RIPPER RIDOR Bagging AdaBoost CART RML C4.5 PRISM	RML LR-based 94.0% (Adult), 88% (Adolescent), 92% (Child)	SE 94.00%, SP 97.00% (Adults), SE 87.00%, SP 80.00% (Adolescent), SE 91.00%, SP 91.00% (Child)
[23]	PSO -RBF, PSO-BPNN	∗PSO+RBF 97%	SE 90%, SP 89%

Models: Cubic Regression Model, Logistic Regression (LR), Polynomial Regression, Naïve-Bayes (NB), Random Forest (RF), K Nearest Neighbour (kNN), Convolution Neural Networks (CNN), Multi-Layer Perceptron (MLP), Scala Vector Classifier (SVC), RIPPER RIDOR Bagging AdaBoost CART, Rules Machine Learning (RML), C4.5 PRISM, Back Propaganda Neural Network (BPNN), Particle Swam Optimizer (PSO), Radial Basis Function (RBF). Additional Performance Matrix: Precision (PREC), Recall (REC), F1-Score(F1-S), Confusion Matrix(CM), Specificity (SP), Sensitivity (SE), Area under Curve (AUC), Cohen Kappa (KP), Matthew Correlation Coefficient (MCC), False Discovery Rate (FDR), Cross Validation (CV), Negative Predictive Value (NPV), Positive Predictive Value (PPV), False Positive Value (FPV), Area under Receiver Operating Characteristic (AUROC)
Note: ∗Belongs to the Neural Network family.

For instance, Shilaskar et al. showed that logistic regression is the most accurate type of regression for autism, while the random forest is the most accurate type for dyslexia. However, given that the source data is highly biased and several performance indicators tend to zero, the results are not all that encouraging and their work needs to apply the Synthetic Minority Oversampling Technique (SMOTE) to handle imbalance data [7]. Besides, Thabtah et al. illuminated recent research that utilizes machine learning for ASD classification and investigated the utility of machine learning with Decision Tree and Random Forest for ASD prediction. They developed an ASDTest application to detect symptoms of ASD based on AQ-10 scoring data (cognitive, behavioral and social skill test) [15]-[16]. The results obtained using the Machine Learning with DT algorithm in WEKA were compared to the results obtained by other statistical models, such as Logistic Regression, showed superiority in detecting autistic traits over probabilistic classifiers derived by Naïve Bayes; however, the performance decreased when data collected through the application gets bigger [15]. Accordingly, the work on using regression models to detect NDDs seems to decrease in its suitability due to its inefficiency in handling big datasets.

2.3.2.1 Classification Problems for Detecting NDDs

Classification is a common task in supervised learning, utilized when predicting categorical outcomes and determining whether a given example belongs to a specific category or not. This is distinct from regression, which is employed for predicting continuous values. However, although it can help categorize results based on certain tasks, it may not be able to handle all complex tasks within a timeframe. Most classification problems address the task of detecting NDDs as shown in Table 6. For instance, Alice Jacob et al. focused on ADHD detection from Time-Frequency images (TFIs) to identify the frequency band with higher ADHD-related data resulting in the TFI-based CNN model and GLCM-based KNN classifier supported higher ADHD-related information at the theta band compared with the upper- beta and lower- gamma bands [24]. The results showed the best performance on specific features in a short timeframe, but not with bigger TFI. Wang et al. (2023) proposed a model of ICA-CNN and achieved an accuracy of 67%; however, the training dataset comprises only 168 observations, which is insufficient for thoroughly training the parameters [25]. On the other hand, Omar et al. (2023) in their work on LSTM-CNN model in detecting epilepsy which is common in ADHD’s children achieved a high accuracy of 97%. They also emphasized the significance of temporal dependencies in EEG signals, which reflect the connectivity and evolving state of the subject's cognition [26]. Due to the multi-dimensional and different datasets used in previous studies, there is an inconsistent affirmation stating which of the five bands has a significant effect on discriminating ADHD. In the U-Profile, resting-state power graph highlighted theta and beta best to detect ASD and ADHD [87], supported by Alim (2023) who indicated that signals less than 30Hz or the first four sub-bands are significant [80]. On the other hand, Parashar (2021) pointed out that all bands from each regional cortex are significant. Furthermore, different feature selections cause mixed statements about which nodes affect ADHD the most. Holker (2022) mentioned only six nodes; namely, FP2 (right pre-frontal), O2 (right occipital), F7 (left frontal), F8 (right frontal), T7 (left temporal) and P8 (right post-temporal) are the most important nodes [74]. Chen (2018) agreed on that all nodes are equally important [49]. Previously, a classical ML-based classifier was used to identify ADHD by extracting the features manually. Although the contributions have already been proved, they cannot achieve multiple-class classification with automated feature extraction. Meanwhile, the identifiable EEG segments of ADHD are too long to limit the real-time ADHD detection. Furthermore, methods of extraction that involved complex time-series features have not been extensively explored for ADHD [32]. The Deep Neural Network Framework has more layers (more depth) and each layer adds complexity to the architecture while enabling the framework to process the inputs concisely for outputting the ideal solution. LSTM can be applied when there is a long series with a sequence prediction that’s required and some long- term dependency of data to go parallel with it. The CNN-LSTM framework proposed by Wang et al. (2022) incorporated features extracted by the CNN across various frequency bands and intricate ERP waveforms. However, despite this comprehensive feature set, the framework struggled to identify the ultimate key activities due to spatial feature-extraction problems [34].

2.3.3 Type of Datasets and Parameters Used

This sub-section summarizes and discusses the findings of RQ2: What type of NDD datasets in previous works have been used to build the models? and RQ3: What type of parameters or variables have been used? Table 9 shows the type of neurodevelopment disorders, dataset sources and related studies working on the prediction and detection of NDD.

There are different diagnosing way being performed to collect the datasets. Existing diagnostic tools to detect NDD include Clinical Observation, Statistical Evaluation, ML Classification, IOT/Robotics …etc. For instance, for ASD, most studies have used the Demographic, Medical, Facial Image Data, BMRI, Eye Tracking data, BMRI/FMRI and EEG data to perform the predictions and detections based on their representation stated in DSM-V. For structured datasets, the most frequently used databases include Observation and Behaviorial, BMRI/FMRI and EEG-based data. This is due to their availability on published well-known websites, like IEEE and Kaggle, for advanced research [87].

Incorporating multiple sources of data can be useful if there is a lack of data availability to predict and detect NDD Disorder. For instance, the dynamics of certain disorders, (e.g., cerebral palsy, epilepsy, GDD) could be associated with other information (e.g. neuron-defect density, population density and mobility) and this information should be incorporated in the process of modeling the classification of NDDs to reduce the residual errors of the models [83].

2.3.4 Type of Problems Addressed by Machine-learning Models

This sub-section summarizes and discusses the findings of RQ4: What types of problems are addressed using these models? and RQ5: Which models achieved the highest performance? Table 8 tabulates and summarizes the regression problems and all the individual machine-learning models applied to achieve the objectives of each study. On the other hand, Table 9 tabulates and summarizes the classification problems and all the relevant individual machine-learning models applied to solve these classification problems. The best models and their performances for each study are also tabulated in these tables. The details of the findings are discussed in the next sub-section 2.3. Based on the results shown in Table 8 and Table 9, for time-series data, VAR and LSTM are the most common machine-learning algorithms used to perform detection and prediction [17], [18], [31], [32], [83], [85], [87]. On the other hand, the family of ANN, kkNN, MLP and CNN algorithms are widely used in solving classification tasks [24], [30], [33]-[34], [36], [39], [47], [84].

Table 8. Classification: Types of machine-learning approaches and individual models used.

Study	Model	Best Model with Accuracy	Additional Performance Metrics
[24]	CNN, GLCMbased, KNN Time-Frequency image (TFIs)	∗Deep-CNN 99.75%	PREC 96.33%, REC 96.74%, F1-S 96.54%, CM FN 3.26% FP 3.75%
[25]	LR, SVM, RF, CNN	∗ICA+CNN 67%	SP 89%, SE 42%, PREC 77%, AUC 0.65
[26]	EEGNet, DeepConvNet and ShallowConvNet	∗DeepConvNet (LSTM-CNN) 96%	PREC 96%, REC 96%, F1-S 96%, KP 95.20%, MCC 95.23%, Robustness Difference 39%
[27]	SVM, kNN, RF, DT, CNN	SVM 88%	-
[28]	GPC, RF, kNN, MLP, DT, LR	GP-based 97.53%	REC98.46%, PREC 96.92%, AUC 0.99
[29]	SVM, LR, NB, CNN, CNN+LSTM	∗CNN+LSTM 98.03%	SP 98.97%, SE 99.25%, F1-S 99.13%, FDR 71%
[30]	SVM+RBF, expEEGNetwork-LSTM	∗expEEGNetwork-LSTM 99.06% and 98.68%	F1-S 99.14% and 99.24%, MCC 98%
[31]	EEGNET, ConvNets, LSTM	∗LSTM 90.50%	-
[32]	Graph FMRI, FCNet, fusion FMRI, Deep FMRI, SVM RFE	SVM RFE 75%	-
[33]	NLSVM, LR, RF, GNB, kNN, CNN	∗CNN+LR 95.83%	PREC 100%, REC 92%, F1-S 96%, AUC 0.96
[34]	LSTM, LessCNN+LSTM, DeepCNN+LSTM,CNN+LSTM	∗CNN 84.44%	SE 85.39%, SP 80.57%
[35]	ANN, SVM, kNN, MPL, LR, RF, GPC	LASSO+SVM 94.2%	SE 93.3%, SP 90.2%, AUC 0.96
[36]	CNN, 4D CNN, 6D CNN	∗4D CNN 98.56%, *6D CNN 98.85%	PREC 98.69%, REC 98.81%, F1-S 98.75% (4D CNN) PREC 98.75%, REC 99.25%, F1-S 99% (6D CNN)
[37]	VGGFace, ResNet50, VGG19, MobileNet (Xception)	∗MobileNet(Xception) 91%	SP 94%, SE 88%, CM FN 26%, FP 14%
[38]	VGGFace, MobileNet, resNet50, VGG19	∗MobileNet 97.60%	PREC 97.50%, SE 97%, SP 97%, AUC 0.97
[39]	RF, LR, DNN, CNN resting state	∗CNN 97%	CV 93-96%
[40]	CNN, VGG16, VGG19, ResNet50, ResNet101, ResNet152, AutoML	∗AutoML 96%	CV 94%
[41]	DSVM, DT, BDT, DNN	∗DNN 93.3% (AUC)	AUC 0.97, SE 93.28%, SP 91.38%, CV NPV 94.46%, PPV 90.06%
[42]	BQC, FF-NN, IF-SVM, kNN, LDA, SCNN, MBCNN, SVM-MLP, IPSO-NN, RF, SVM-RBF	∗1DCNN 99.70%-99%	PREC 98-99%, REC 98-99%, F1-S 98-99%
[43]	LR, kNN, SVM, NB, ALexNet, GoogleNet, SqueezeNet	∗CNN +LASSO (SqueezeNet) 88.33%	PREC 83%, AUC 0.83, FPV 16%
[44]	NB, kNN, LR	kNN 86%	PREC 100%, REC 78%
[45]	kNN, SVM, MLPNN, LEDPatNet19	∗LEDPatNet19 99.29% (Arousal 94.58%, Dominance 92.86% and Valance 94.44%)	PREC 99.29%, REC 99.30, F1-S 99.29 (Arousal FC6) PREC 94.43%, REC 94.63%, F1-S 94.53% (Valance F7)
[46]	SVM, kNN, J48, Bagging, Stacking, AdaBoost, NB	kNN 99.1%	CV 98.6%-99.2%
[47]	MLP, RF, CNN	∗CNN 92.31%	AUC 0.96 F1-S 91.54%, PREC 89.72%, REC 93.45%
[48]	SVMLinear, SVM+RBF, SVM+Grid,RF, RF+Grid	SVM+Grid 97.42%	PREC 96%, REC 91.4%, F1-S 93.4%
[49]	SVM, LR, NB, RF, DT, kNN	DT and NB 79.71%	AUROC 0.83, SP 96.4%, PPV 20.5%, SE 40%
[50]	rbio1.1 +kNN	rbio1.1 +kNN 99.17%	CV
[51]	CNN, VGG16	∗VGG16 68.54%	CV
[52]	LR, SVM, NB, kNN, ANN, CNN	∗CNN 96.88%	SP 100%, SE 93.35%
[53]	SVM, kNN, RF, CNN	*CNN 70.20%	SP 61%, SE 77% CV
[54]	Stacked autoencoders, Stacked autoencoders+MLP	∗MLP 85.06%	SE 81%, SP 89%
[81]	RF, SVM, DNN, CapsNet	∗CapsNets 71%	SE 73%, SP 66%
[55]	MLPNN, DeepCNN	∗Deep CNN 98.48%	PREC 97.48%, REC 97.47%, F1-S 97.47%, CV 99.06%
[56]	SVM, LDA, DT, RF, kNN+RKF	kNN+RKF 88.37%	SP 91.3%, SE 85%, AUC 0.88
[57]	SVM, SVM+RBF	SVM+RBF 91.3%	CV
[58]	kNN, Efficient Net, LR	ANN 97%	CV
[59]	LR, SVM, SVM+RBF	SVM RBF 98.62%	PREC 89%, REC 89%, F1-S 89%, CV 59.78%
[82]	SVM, RF, LR, 2CC3D	∗2CC3D 89%	F-Score 89%
[60]	MLP + DISR, MLP + mRMR	∗MLP+DISR 93.65%, ∗MLP+ mRMR 92.28%	Variance 0.7%
[61]	SVM	SVM 59-66.3%	SP 68%-87.7% SE 22.9%-55.6%

Models: Gray level co-occurrence matrix (GLCM)-based, Long Short Term Memory (LSTM), Gaussian Processes (GP), Naïve-Bayes (NB), Locations of Sophie Germain's Primes on Ulam's Spiral-Based Features (LSGP-USFNet), Expert EEG Network (expEEGNetwork), Least Absolute Shrinkage and Selection Operator with Support Vector Machine (LASSO +SVM), One Dimension Convolutional Neural Network (1D+CNN), extreme inception (Xception), Artificial Neural Network (ANN), LED Pattern Feature Extraction (LEDPatNet19), Back Propagation Neural Network (BPNN), Decision Tree (DT), Linear Regression (LR), kk-Nearest Neighbour (kk-NN), Support Vector Machine RBF kernel (SVM+RBF), Support Vector Machine (SVM), GoogleNet, AlexNet, Residual Neural Network (RNN), 2 Channel Convolutional 3 Deep Neural Network (2CC3D), Double Input Symmetrical Relevance (DISR), minimum Redundancy Maximum Relevance (mRMR), BQC: Bayesian quadratic classifier, FF-NN: Feed forward neural network, IF-SVM: Immune feature weighted SVM, QDA: Quadratic discriminant analysis, KNN: K nearest neighbor, SVM-RBF: SVM-radial basis function, SVM-RFE: SVM-Recursive Feature Elimination, Deep Belief Network (DBN), Decision Tree (BDT), Deep Support Vector Machine (DSVM). Additional Performance Matrix: Precision (PREC), Recall (REC), F1-Score(F1-S), Confusion Matrix(CM), Specificity (SP), Sensitivity (SE), Area under Curve (AUC), Cohen Kappa (KP), Matthew Correlation Coefficient (MCC), False Discovery Rate (FDR), Cross Validation (CV),

Negative Predictive Value (NPV), Positive Predictive Value (PPV), False Positive Value (FPV), Area under Receiver Operating Characteristic (AUROC).
Note: ∗Belongs to the Neural Network family.

Table 9. Disorders, database sources and studies.

NDD Disorder	Database Sources or Parameters
ASD	Demographic Data [12], [13], [15], [16], [19], [30], [52], [56], [67], [84] Medical Observation and Behavioral Data [9], [13], [15], [16], [40], [46], [49], [52], [69], [75] Eye Tracking Data [8], [10], [38], [41], [76] Facial Images Data [12], [37], [40], [47], [51], [58], [70], [72], [85] BMRI/FMRI Data [34], [39], [53], [54], [81], [82] EEG Data [71]
ADHD	Demographic Data [25], [61], [62], [78], [79] Medical Observation & Behavioral Data [25], [61], [62], [28], [35] Eye Tracking Data [11] BMRI/ FMRI Data [21], [22], [25], [32], [61], [62], [68], [78], [79], [86] EEG Data [6], [17], [18], [20], [23], [24], [25], [27], [28], [30], [31], [33], [35], [36], [42], [43], [44], [45], [50], [55], [57], [59], [60], [63], [64], [66], [73], [74], [80], [83], [84]
Dyslexia	Medical Observation & Behavioral Data [7], [48], [65] Eye Tracking Data [77] BMRI/FMRI Data [65] EEG Data[14], [27]
Others	EEG Data [26], [29], [87]

(Please, refer to Table 6 and Figure 3)

2.3.4.1 Approaches to Solve Regression Problems

The approaches to solve regression problems in detecting and predicting the occurrence of NDDs can be divided into statistical and machine learning approaches. Based on the information tabulated in Table 9, for the statistical approaches, several models have been used to perform the detection and prediction of NDDs, including the Cubic Regression Model [6], LR [7], [8], [11]-[12], [14], [16], [45], [60], [64] and ANN [38]. In multivariate and time-series modeling, Cubic Regression combined with SVM-RBF by Delisle et al. (2023) outperformed the statistical approach MX-VAR model by Redondo. Based on the review, deep-learning algorithms have outperformed the statistical approaches in detecting and predicting the disorders with multi-vatraite approaches, such as, Locations of Sophie Germain's Primes on Ulam's Spiral-based (LSGP-USFNet) [17], mixed-effect functional-coefficient autoregressive (MX-FAR) [18], Single Photon Emission Computed Tomography (SPECT) [62], Variational Mode Decomposition and Hilbert Transform-based (VHERS) [63], Multi-layer Perceptron (MLP), Phase-transfer Entropy (PTE) [20], Deep Variational Autoencoder (DVAE), Attention Attribute-enhanced Network (AAEN), Metaheuristic Spatial Transformation (MST) , Graph Signal Processing (GSP), Graph Learning (GL), Meta-cognitive Neuro-fuzzy Inference System (McFIS) (International Conference on Cognitive Computing and Information Processings 1. 2015 Noida et al., n.d.), Local Binary Encoding Method (LBEM), Linear Discriminate Analysis (LDA) [51], Kernel Principal-component Analysis (KPCA) [68], [79].

2.3.4.2 Approaches to Solve Classification Problems

Based on the information tabulated in Table 8, neural network methods have been found to be very effective in detecting NDDs. This review reports that the neural network-based methods have achieved 27 best results out of 81 studies [24]-[26], [29]-[34], [36]-[43], [45], [47], [51]-[55], [58], [60], [64]. These classification approaches use different methods of extraction and selection depending on the type of datasets represented for the purposes of their studies. Few researchers have used T-test and LASSO [28], [73], while few others used ICA and PCA to select the most discriminate features to optimize the multi-dimensional features within their datasets before being fed into their proposed models [20], [51]. Some authors applied the grid method to improve accuracy performance, such as Pralhad et al. (2021) who compared SVM and RF models using the grid method in dyslexia detection through Video on Observation and Behavioral datasets, resulting SVM using grid achieved the highest accuracy of 97.42% [48].

Various studies have investigated autism classification using diverse methodologies and datasets. For instance, Alsaade et al. (2022) evaluated deep-learning models' performance in detecting ASD *via* facial features, highlighting Xception's effectiveness [37]. Elshoky (2022) employed deep learning (VGG16), achieving a remarkable accuracy of approximately 96% compared to other deep-learning models, such as VGG19 and ResNet [40]. Kanhirakadavath and Chandran (2022) utilized eye-tracking datasets along with deep-learning models, while Kanhiraka, Rashid and Lin (2022) employed machine-learning techniques on eye-tracking data for early autism screening in children [41], [68]. Studies like Shilaskar et al. (2023) and Delisle-Rodriguez et al. (2023) utilized observation and behavioral (AQ-10) datasets with supervised-learning models, noting SVM's superior performance [6]-[7]. In the area of medical research datasets, such as ABIDE and fMRI, were commonly used alongside machine-learning and deep-learning models, like MLP, NB, RF, CNN, ResNet and GoogleNet. Researchers like Attlah et al. and D.Wang et al. (2023) observed improved accuracies with their trained models compared to pre-trained ones [25]. Rabbi et al. (2021) compared various models, finding CNN to be highly accurate in detecting autism from facial images [47]. Ahmed et al. (2022) developed a web application using deep learning, achieving 95% accuracy with models like MobileNet [38]. Ahmed et al. (2022) adopted a deep transfer-learning approach, with MobileNet exhibiting the highest accuracy of 97% in detecting autism from children's facial images. Deep-learning algorithms offer significant benefits over statistical methods when it comes to uncovering inherent patterns for prognosis or diagnosis in neuropsychiatry [29]. In recent decades, research in neuropsychiatric diagnosis using EEG has primarily centered on addressing the "multi-dimensional problem" of localizing the complex brain-activity measurements. EEG-based models have seen extensive research in investigating dysfunctions across various neuropsychiatric disorders such as depression, Alzheimer's disease, epilepsy, phobias, conduct disorder, schizophrenia and NDD. Often, these methods are combined with artificial intelligence or machine-learning approaches, as shown in Table 10.

Table 10. Statistical methods used for classification and regression problems.

Study	Model	Best Model	Additional Performance Metrics
[17]	kNN, CNN, LSTM, SVM, NB, LSGP-USFNet	LSGP-USFNet +SVM 97.5%-98.9%	SE 90.57%-99.06%, PREC 91.45%-98.49%, F1-S 93.03%-98.77%
[18]	FAR, VAR	MX-VAR 95% (fPDC)	Mean fPDC 95%
[62]	SPECT, CAD, ML (Different Brain Region)	SPECT+CAD+ML 80% (frontal cortex)	F-Measure 79.95%
[63]	CNN, MLP, VHERS	ELM VHERS 99.95% (delta)	SE 100%, SP 99.89%, KP 99.9%, PREC 99.91%, F1-S 99.9%, MCC 99.9%
[64]	MLP	∗MLP 90.01% (Trend)	SE 90.55%, SP 89.84%
[20]	gECV+ANN+GA (PTE Brain Region)	*ANN+gECV 89.7%	PTE $p<0.01$ dPTE $0.5<dPTExy\leq 1$
[19]	AAEN	∗AAEN 86.22%	SE 44.45%-98.18% SP 66.66%-97.14%
[22]	McFIS+GA+ELM	McFIS+GA+ELM (63 voxels taken from Top- 50 best binary solutions)	PREC 92%, REC 90%, F1-S 90%
[21]	VA-Relief	VA-Relief 98.04%	-
[78]	Functional connectivity, resting state	LDA 80.08%	SE 80.7%, SP 79.47%
[79]	KPCA-SVM	KPCA-SVM 81%	-

Models: Locations of Sophie Germain's Primes on Ulam's Spiral-Based (LSGP-USFNet). mixed-effects functional-coefficient autoregressive (MX-FAR), functional Partial Directed Coherence (fPDC), Single Photon Emission Computed Tomography (SPECT), Variational Mode Decomposition and Hilbert Transform-Based (VHERS), extreme learning machine(ELM), Multi-Layer Perceptron (MLP), Phase Transfer Entropy (PTE), Genetic Algorithm (GA), Global Effective Connectivity Vector (gECV), Deep Variational Autoencoder (DVAE), attention attribute-enhanced network (AAEN), Graph Signal Processing (GSP), Graph Learning (GL), Meta- Cognitive Neuro-Fuzzy Inference System (McFIS), Extreme Learning Model (ELM), linear discriminate analysis (LDA), kernel principal component analysis (KPCA). Additional Performance Metrics: Sensitivity (SE), Specificity(SP), Precision (PREC), Kohen's kappa (KP), F1-Score (F1-S), Mathews Correlation Coefficient (MCC), Functional Partial Directed Coherence (f PDC), Multivariate Analysis Of Variance (MANOVA).
Note: ∗Belongs to the Neural Network family.

This line of inquiry offers considerable potential for revealing neural correlates of NDD, enhancing diagnostic methods and progressing treatment strategies. This entails employing sophisticated statistical techniques, like low-resolution electromagnetic tomography (LORETA), Phase Transfer Entropy (PTE), Variational Mode Decomposition and Hilbert Transform-Based (VHERS), optimization methods, among others, to overcome the inherent spatial resolution limitations of EEG [20], [63]. Furthermore, there was a drastic increasing amount of research conducted with EEG-based datasets. Studies and methods of feature extraction and selection are shown in Table 12.

2.3.5 Assessment Measures and Methods

This sub-section summarizes and discusses the findings of RQ6: What evaluation metrics and methods are employed to measure the performance of the machine-learning models? (e.g. Accuracy, Precision, Recall, F-Measure, ROC, AUC, Kappa) of the proposed machine learning algorithms for prediction and detection models? In most regression problems, all the proposed methods or algorithms are measured by using Autoregressive (VAR), mean, standard deviation (STD), mean functional partial directed coherence(fPDC), Root Mean Square Error (RMSE), t-test, two-way ANOVA analysis, average shortest path (d) and betweenness centrality (Cbetweenness), Friedman test, Nemenyi test, 10-fold metrics (Recurrence, Determinism, Entropy, Laminarity, Trapping Time and Trend), permutation statistical test, VOXELS'COUNTS and high testing efficiency (fitness value), nested cross-validated accuracy and kappa score. On the other hand, Accuracy and ROC are mostly used for evaluating the performance of the classifiers proposed in those studies. In this paper, 27 out of 81 (33%) studies found that the individual models that belong to the neural-network family performed better when compared to other linear and non-linear methods. Tables 9 and 10 show that machine-learning models achieved lower Mean Absolute Error (MAE) and Mean Squared Error (MSE) measurements compared to other statistical models (e.g. VAR and MX-VAR) [18]. However, for long-term trend, it can be observed from these tables that deep-learning approaches improve RMSE readings in non-linear models' classification which achieved above 98% of accuracy [36], [44], [55]. As we have noticed, based on summaries stated in previous sub-sections, machine-learning approaches performed better than statistical approaches. Deep-learning and ensemble algorithms consistently exhibit a trend of achieving higher accuracy measurements [24], [55], $FF1$ Score measurement [11]-[12], [28], [33], [53], [55], [59], [74], AUC [35], [41] and ROC measurement [67], [73], in comparison to other statistical and machine-learning models evaluated in this study.

2.3.6 Ensemble Method

This sub-section summarizes and discusses the findings of RQ7: What types of ensemble methods are used in machine-learning models?

Various ensemble approaches have been introduced for predicting NDDs. Table 11 provides an overview of these methods used for predicting and detecting disorder outcomes, along with summarizing the evaluation techniques and metrics employed in ensemble learning. Further exploration is warranted to assess the potential of ensemble or hybrid models based on deep-learning techniques utilizing multi-source data, as they have demonstrated enhancements in base-model performance. An ensemble method refers to a strategy that employs multiple independent models or weak learners, which may be similar or diverse, to generate an output. Ensemble methods are typically classified into boosted trees, bagged trees, subspace kNN and stacked approaches [73]. Bagging involves employing homogeneous weak learners arranged independently in parallel and aggregating their predictions to determine the final output.

Table 11. Ensemble methods used for classification problems.

Study	Dataset	Best Model with Accuracy	Additional Performance Metrics
[65]	MS-ROI XGBoost	MS-ROI XGBoost 99.87%	PREC 92.36%, REC 91.65%, SE 99.89%, SP 99.91%
[74]	ANN, RF, SVM	RF 81.82%	F1-S 81.79%, PREC 81.95%, REC 81.82%
[66]	SVM, RF, AdaBoost	Adaboost 82%	SE 75%, SP = 86%
[73]	Ensemble	Ensemble 98.33%	CF
[67]	Boosting, DT, NN, NB	RF+SMOTE 98%(ROC)	CF TPR 88% TNR 93%
[75]	SVM, RF, SMO	RF 87% (ROC)	TPR 88.5%
[68]	CDAE+AdaDT	CDAE+AdaDT 90% (AUC)	SE 76.92%, SP 73.08%, CF
[76]	DT, NB, kNN, SVM, Stacking	Ensemble(stacking) 89.82%	SE 89.21%, SP 90.31%, KP 0.33%
[69]	SVM, kNN, RF, NB, AdaBoost, SGD, CN2 rule inducer	SGD 99.6% (Adult), RF 97.2% (Adolescent) RF & SGD 99.7% (Toddler)	F1-S, PREC & REC (90%-100%)
[72]	SVM, RF, LR, kNN, SVM+PSO	SVM-PSO 95.6%, RF 90.45%	-
[70]	DT, CNN, AdaBoost	Adaboost 98.77% (Toddler), 97.20% (Child), 93.89% (Adolescent), 98.36% (Adult)	SE 99.39%, SP 99.39%, KP 97.10%, AUROC 99.98%, Logloss 3.01% (Toddler),SE 98.40%, SP 98.46%, KP 94.41%, AUROC 99.89%, Logloss 9.62% (Toddler) SE 97.50%, SP 98.33%, KP 89.37%, AUROC 98.61%, Logloss 15.80% (Toddler),SE 99.30%, SP 96.11%, KP 96.02%, AUROC 99.95%, Logloss 5.64% (Toddler)
[71]	RF, LR, Bagging, CNN	RF 97%	PREC 97%, REC 97%, F1-S 97%
[77]	kNN,	kNN 53.4%	MANOVA p-value<0.01

Models: MS-ROI XGBoost, AdaBoost, Ensemble, Random Forest (RF), Random Forest Based (RF-based), Stochastic Gradient Descent (SGD), Partical Swam Optimization (PSO), Gradient Boosting Machine (GBM), Sequential Minimal Optimization (SMO), Convolutional Denoising Autoencoder (CDAE), Adaptive Boosting Decision Trees (AdaDT). Additional Performance Metrics: Precision (PREC), Recall (REC), Sensitivity(SE), Specificity(SP), Confusion Matrix(CF), True Positive Rate(TPR), True Negative Rate(TNR).

For instance, in their study on classifying ASD *versus* control groups, M. Rakic and M. Cabezac combined data from functional and structural MRI and assessed it on a sizable multi-site dataset. Their quantitative analysis was conducted on 817 cases from the International Autism Brain Imaging Data Exchange I (ABIDE I) dataset, comprising 368 ASD patients and 449 control subjects. They achieved a classification accuracy of 85.06% with a standard deviation of 3.52% when employing an ensemble of classifiers. Combining information from both functional and structural sources resulted in significantly improved performance compared to using an individual pipeline [54]. Sangeetha et al. (2022) showed that ensemble methods, especially MS-ROI with XGBoost, are capable optimizing computational time in detecting dyslexia within smaller data sizes [65]. Hamedi et al. (2021) used the stacking method in detection for ASD with rs-MEG signals and achieved an accuracy of 89.82%, showing that the left central (LC) of the brain can discriminate the ADHD group [76]. Thus, ensemble methods have proven to improve predictive performance using an individual model and multiple learning algorithms although they are time and space-consuming compared to other machine-learning models [67], [74], [76]. Efforts need to be directed towards harnessing the potential of ensemble methods in future-research endeavors, in order to bolster their applications for addressing various disorders.

Table 12. Feature-extraction and machine-learning models from related EEG-based studies.

	Features Extraction										Features Selection								ACC	Model
Paper	3FD	PSO	ICA	SPM(MSP)	LASSO	T-Test	RFE	3EDAs	VMD-HT	DISRmRMR	Bandsδ θ/β	ROI	Recurrence	Resting State	Changing state	ERP	PCA	Time Series		
[60]	✓									✓	✓								93.65%	MLP
[23]		✓									✓								90%	SVM (RBF)
[57]											✓		✓	✓					91.3%	SVM (RBF)
[55]											✓					✓			98.48%	CNN (DCNN)
[73]	✓				✓		✓				✓								98.33	Ensemble
[66]											✓	✓							84%	Ensemble (Adaboost)
[20]			✓	✓													✓		98%	ANN (dTF+ANN)
[35]					✓	✓					✓								94.2	SVM (LASSO)
[36]				✓							✓								98.85%	CNN (6D CNN)
[33]											✓								95.83%	CNN (CNN+LR)
[74]				✓							✓								81.82%	RF
[28]					✓	✓					✓								97.53%	SVM (GP-based)
[30]								✓			✓								96.16%	ANN (MEMD-GA-ANN)
[80]											✓						✓		94%	SVM (Gaussian)
[63]									✓		✓								99.81%	DNN+ELM (VHERS)
[59]	✓										✓								98%	SVM (RBF)
[55]				✓							✓								98.48%	CNN
[42]				✓							✓								98%	DCNN
[6]				✓							✓								81.37%	SVM (RBF)
[51]			✓								✓			✓					85%	CNN
[50]											✓							✓	99.17	SVM (RBF) kNN Bio and rbio
[73]	✓										✓								98.33%	Ensemble (subspace)
[45]				✓							✓								99.29%	LEDNet (LEDPatNet19)
[83]				✓							✓							✓	97.75%	LSTM
[24]				✓							✓								99.75%	CNN (TFI-based)
[34]				✓							✓			✓				✓	98.23%	CNN-LSTM
[17]				✓							✓							✓	97.46%	Gray-tones (LSGP-USFNet)
[26]			✓								✓							✓	96%	CNN (DeepConvNet)
[31]	✓										✓							✓	90.50%	LSTM

Notes: 3FD: Higuchi, Katz and Petrosian fractal dimensions Largest Lyapunov Exponent (LLE),PSO: Partical Swam Optimization, ICA: Independent Component Analysis,SPM: Statistical Parametric Mapping applied multiple sparse priors (MSP) algorithm, LASSO: Least absolute shrinkage and selection operator,T-Test: T score = (difference between the group)/(difference within the groups),RFE: Recursive feature elimination,3EDAS: three multivariate EDAs (MEMD, MEWT and MVMD), VMD-HT: variational mode decomposition (VMD) and Hilbert transform (HT),DISR: Double Input Symmetrical Relevance (DISR),mRMR: minimum Redundancy Maximum Relevance, ROI: Region of Interest, ERP: Event Related Potential, PCA: Performance Component Analysis

2.3.7 Deep Learning Method

This sub-section summarizes and discusses the findings of RQ7: What types of deep-learning approaches are used in NDD detection?

Within the emergence of machine learning, the most effective methods identified for predicting neurodevelopment disorders are predominantly associated with the neural-network family. The experimental results showed consistent performance improvements by the proposed deep-learning approaches over other representative linear and non-linear methods on multiple real-world datasets. These algorithms include the Long Short-Term Memory (LSTM) [30]-[32], [83], Convolutional Neural Network (CNN)[24], [27], [33], [34], [36], [39], [47], [55], Multi-layer Perceptron (MLP) [11], [14], [45], [60], [64], Neural Network [31], [34], [42]-[43], [47], [53], [55], Hybrid Neural Network (HNN) [30], [45], [81], [82] and combinations of statistic and deep-learning approaches. LSTM algorithms were shown to be superior in detecting ADHD, which supports long sequential data, like EEG [29], [31], [34]. A feature selection-based time-series modeling has been proposed for predicting future disorders [24], [26], [87]. The work proposed a multi-objective evolutionary algorithm to find the best neural-network algorithm (deep learning) for detection differences. Although the Convolutional Neural Network (CNN) is the best model when it comes to process image data, as it is capable to excel local features and is good in pattern recognition [47], it has limited effectiveness for sequential data. For large datasets, training takes a long time to complete. In previous studies, Kaur et al. (2021), Moghaddari et al. (2020), Mafi & Radfar (2022), TaghiBeyglou et al. (2022) and Saini et al. (2022) have conducted their work on ADHD detection using CNN model [33], [36], [51], [55]. Moghaddari in his work to tackle the ERP fatigue problem using deep CNN achieved an accuracy of 98.48%. Mafi & Radfar (2022) used 4D and 6D connectivity tensors as a convolutional neural network input, achieving an accuracy of 98.85%. While Taghibeyglou et al. (2022) achieved an accuracy of 95.83% on their CNN+LR model, the framework suffers from a time-consuming training procedure, since the method focuses only on raw time series in both spatial and temporal domains. Furthermore, in the work of Saini et al. (2022) on the evaluation of their proposed architecture 1DCNN on three databases, the best accuracy was achieved on the database with few features compared to the database with more features [44]. Hence, an improved model is required to be able to overcome the limitation of CNN model in processing more features for ADHD detection.

Long Short-Term Memory (LSTM) is a deep recurrent neural-network architecture utilized for the classification of time-series data, a crucial aspect of time-series analysis focusing on comprehending and predicting sequential data points over time [68]. Within deep learning, LSTM models are applied to regression analysis, addressing issues of non-linearity and data interdependence to enhance traditional regression models. These networks are trained to classify sequence data, leveraging LSTM's capability to retain information from previous inputs over extended periods. This characteristic renders LSTM particularly effective for handling sequences with prolonged dependencies, where earlier time steps significantly influence subsequent ones. Sharma & Singh (2023) in their novel approach on expEEGNetwork-LSTM achieved an accuracy of about 98.02% [30]. In other works, Huang et al. (2022) with their objectives to solve time window issues in deep learning, they achieved an accuracy of 90.50% with their LSTM model. One drawback of LSTM models is their computational intensity, requiring more processing time compared to alternative methods [31]. While LSTM models can achieve high accuracy, there remains room for improvement with certain datasets. Notably, LSTM overcomes the limitations of traditional RNNs by employing separate memory cells capable of storing long-term information independently of current inputs or outputs [30], [31], [34], [83]. This property enables LSTM to learn and retain long-term dependencies while mitigating issues like the vanishing or exploding-gradient problem. Another way to optimize the LSTM model is to use hyper-parameter optimization, which is a process that involves searching for the best combination of values for the parameters that control the behavior and performance of the model, such as the number of layers, units, epochs, learning rate or activation function like sigmoid, hyperbolic tangent and rectifier.

A CNN-LSTM network on the other hand uses convolutional and LSTM layers to learn from the training data. Huang et al. (2022) and Zhang et al. (2023) showed that the proposed EEG-based LSTM networks can extract the varied temporal characteristics of high-resolution electrophysiological signals to differentiate between ADHD and NT children and bring new insights to facilitate the diagnosis of ADHD [26], [31]-[32] by leveraging LSTM's ability to capture temporal dynamics and Convolutional Neural Network (CNN) capability to detect spatial patterns. The proposed method proved successful in enhancing EEG classification by outperforming existing models developed for similar EEG-based classification tasks. Wang et al. (2022) in their work with the LSTM-CNN model to process multiple frequency bands and complex ERP waveforms achieved an accuracy of 98.23% [25]. Somehow, this did not help the network find the final key activities. An improved deep-learning model that can extract more spatial feature information from multi-channel EEG signals could be employed to identify commonalities and sub-types [34]. Omar et al., 2022 in their work on detecting epilepsy applied Convolutional Neural Networks (CNNs) for extracting spatial features and Long Short-Term Memory (LSTM) for identifying temporal dependencies, achieving an accuracy of 96% focusing on scalability and efficiency. However, their result suggests that even models with fewer trainable parameters may still require many epochs or batch sizes to achieve optimal performance, highlighting the importance of careful model selection and hyper-parameter tuning [26].

Table 13 illustrates brief description of methods and techniques: their principles, advantages and limitations, in terms of each machine-learning model and technique used in this study.

Table 13. Description methods of each machine-learning model and technique used.

Methods/Techniques	Principles	Advantages	Limitations
Logistic Regression	• Linear model used for binary classification. • Predicts the probability of a binary outcome by applying the logistic (sigmoid) function to a linear combination of input features.	• Coefficients can be interpreted to understand the relationship between features and the probability of the outcome • Computationally efficient with a closed-form solution. • Provides probability estimates for classification	• Assumes a linear relationship between features and the log-odds of the outcome, which may not capture complex patterns • Requires proper feature scaling for optimal performance. • Best suited for binary outcomes, though variations exist for multiclass classification
Decision Trees	• Non-linear model that splits data into subsets based on feature values • Creates a tree-like structure where each internal node represents a feature (or attribute), each branch represents a decision rule and each leaf node represents an outcome. • Constructs the tree using criteria such as Gini impurity or information gain to make splits.	• Easy to visualize and interpret decision rules • Handles both numerical and categorical data without scaling • Can model complex relationships through hierarchical splits	• Prone to overfitting, especially with deep trees. • Small changes in data can lead to different tree structures • May create biased trees if some classes dominate
Random Forest	• Ensemble method using multiple decision trees • Aggregates predictions from multiple decision trees to improve accuracy and robustness • Builds a multitude of trees using bootstrapped samples and random feature subsets, then averages (regression) or votes (classification) to make the final prediction	• Typically, more accurate than a single decision tree due to averaging and reducing variance • Less prone to overfitting compared to individual decision trees • Can provide insights into the importance of different features	• Less interpretable compared to single decision trees • More computationally intensive, requiring more memory and processing power • Making predictions can be slower due to the need to aggregate results from multiple trees
Support Vector Machine (SVM)	• Supervised learning algorithm for classification and regression • Finds the hyperplane that best separates classes in a high-dimensional space. For regression, it finds the hyperplane that best fits the data within a specified margin of tolerance. Allows the algorithm to operate in higher-dimensional spaces using kernel functions (e.g., polynomial, RBF)	• Works well in high-dimensional spaces and with a clear margin of separation • Especially effective in cases with a clear margin of separation • Can use different kernels for non-linear classification	• Training can be time-consuming, especially with large datasets • Performance heavily depends on the choice of kernel and hyperparameters • May not perform well with very large datasets compared to other methods
Multi-Layer Perceptron (MLP)	• Type of artificial neural network with multiple layers of neurons. • Consists of an input layer, one or more hidden layers and an output layer. Uses non-linear activation functions (e.g., ReLU, sigmoid) to model complex relationships • Trained using backpropagation and gradient descent to minimize a loss function	• Capable of modelling complex non-linear relationships • Can be used for various types of tasks, including classification, regression and more • Automatically learns features from raw data	• Can be slow to train, especially with large networks and datasets • Prone to overfitting, especially with a large number of parameters • Performance can be sensitive to hyperparameters and network architecture
Convolutional Neural Networks (CNN)	• Specialized neural network for processing grid-like data (e.g., images). • Uses convolutional layers to automatically learn spatial hierarchies of features (edges, textures, etc.) and pooling layers to reduce dimensionality • Comprises convolutional layers, activation functions, pooling layers and fully connected layers	• Automatically learns and extracts features from images or spatial data • Reduces the number of parameters and computational load through convolutional filters • Performs exceptionally well in tasks like image classification and object detection	• Requires significant computational power and memory • Can be slow to train, especially with large networks • Typically needs large amounts of labelled data for effective training
Recurrent Neural Network (RNN)	• Neural network designed for processing sequential data • Uses loops to maintain a state across sequences, allowing it to handle temporal dependencies • Contains recurrent connections that process sequences one element at a time and update the internal state	• Suitable for tasks involving sequential data, such as time series or text • Can maintain context over sequences to some extent	• Struggles with long-term dependencies due to vanishing gradient issues • Difficult to train on long sequences; often requires more sophisticated architectures like LSTMs or GRUs • Can be computationally demanding, especially for long sequences
Long Short-Term Memory (LSTM)	• A type of Recurrent Neural Network (RNN) designed to handle long-term dependencies and sequential data • Uses gates (input, forget and output) to control the flow of information and manage long-term dependencies in sequences • Comprises LSTM cells that maintain a memory cell to remember information over long periods	• Effectively captures long-term dependencies in sequential data • Mitigates the vanishing gradient problem common in traditional RNNs • Used in various applications like time series forecasting, language modelling and sequence prediction	• Training can be resource-intensive due to the complexity of the model • More complex to understand and tune compared to simpler mod

3. CONCLUSIONS

Based on previous studies on Neurodevelopment Disorder, a summarization included in this review shows strengths, limitations and future directions for research on this domain.

This literature review endeavors to identify and examine various methodologies, datasets, parameters, individual models, ensemble models, performance metrics and approaches utilized in prior research on employing machine-learning techniques to mitigate the escalation of Neurodevelopment Disorder. Six online digital libraries were utilized to retrieve pertinent peer-reviewed articles, resulting in the selection of 81 studies published between 2013 and 2023. The primary objective of this systematic literature review (SLR) was to assess and curate all pertinent research studies concerning the detection and prediction of Neurodevelopment Disorder using machine learning, guided by the mentioned seven questions. The contributions of this paper can be summarized as follows:

- Recognition of the improvement in predicting NDDs by leveraging diverse data sources.
- Acknowledgement of the superior efficacy of neural-network algorithms over alternative linear and non-linear machine-learning approaches.
- Validation of the efficacy of deep learning and hybrid methodologies, showcasing their superior performance and appropriateness in predicting and detecting ND Disorders.

3.1 Significance, Limitations and Future Directions

From this review, we have identified limitations that affect previous works on detection and prediction of Neurodevelopment Disorder using machine learning. Autism Spectrum Disorder (ASD) detection is well studied and achieved maximum performance and highlighted the strengths of signal fusion utilizing Signal-processing and Decision-making techniques. The review cautions that the focus on detecting ASD may overshadow research into other diseases, despite the promising results achieved in ASD studies. While signal-fusion techniques have been extensively explored, other Neurodevelopment Disorders (NDDs) have not received as much research attention. Sustainable ML models are suggested for future work to provide models with feature fusion able to merge different extracted features from various sources and compressed to a single layer before being fed into ML models. Therefore, fusing only important features and suppressing the others will reduce time complexity, thus improving the model's performance. The limited research on signal fusion for NDDs is due in part to challenges in information technology and computer science. A new approach is needed to manage and integrate signals from multiple sensors using artificial intelligence to create a single, optimized feature for meaningful analysis. Although current cloud technologies, such as Google Colab and Kaggle, enable researchers to upload and test datasets, collaboration is often hindered by issues related to credentials and copyrights. Additionally, the limited number of investigations conducted on NDD prediction based on multi-source data underscores the potential for obtaining a more comprehensive understanding of the disorder by integrating such data sources. Analyzing the complex relationships among multi-source data can yield more robust modeling outcomes. To address this issue, researchers need to collaborate openly and be properly credited for their contributions. This would allow signal fusion for NDDs to receive the attention that it deserves and facilitate more effective investigations.

These studies also explored multiple validations that prove the accuracy of each prediction. However, due to the limitations of public datasets, average testing can be performed to varying performances of ML and DL techniques. The analysis highlights that the limitations of publicly available datasets often undermine the effectiveness of machine-learning and deep-learning techniques. This restriction hampers thorough testing and leads to inconsistent performance results across different research studies. Some professionals may face challenges in sharing datasets online due to limited access to technology or varying levels of expertise. To bridge this gap and enhance research outcomes, greater collaboration between medical professionals and data scientists is essential.

Future directions should be ready for the paradigm shift through the emerging technologies which require models in handling big datasets that allow fusion of features to be processed simultaneously. This study underscores the need for future research to embrace new technologies that can manage vast amounts of data. Current algorithms may struggle to handle the simultaneous integration of multiple features, which is crucial for enhancing detection and prediction accuracy. As the volume of data continues to grow, it is essential to develop technologies that can process large datasets while effectively merging various data characteristics. Advanced cloud solutions capable of intelligently integrating these features are needed. Approaches such as genetic algorithms, sentiment analysis and Large Language Models (LLMs) have made strides in this area, but further innovation is required to address the challenges of data integration. On the other hand, studies on ADHD detection have increased the research exposure, especially research related to neurons which acquired deeper feature explorations and sustainable approaches.

Despite the increasing number of studies on Attention Deficit Hyperactivity Disorder (ADHD), more research into its characteristics and the development of sustainable strategies is still needed. This suggests that, while some progress has been made, there remains a significant gap in understanding and managing ADHD through machine learning. To address this gap, clear guidelines on the application of machine learning and deep learning for ADHD are essential to enhance researchers' knowledge and comprehension. Researchers can access handbooks and other resources online, including detailed explanations on platforms like YouTube. Forums on GitHub, Kaggle and Ubuntu also provide opportunities for discussion. Additionally, platforms such as Medium.org and blogs featuring data science can help bridge the gap in understanding and treating ADHD using machine learning.

There is also a need for further exploration of the capacity of deep-learning models or hybrid models in leveraging multi-source data, given their demonstrated ability to enhance the performance of base models. Although several studies have applied cross-fold validations and proven models to be powerful, models are tested on single datasets and are non-data driven. This research also highlights the fact that existing literature frequently lacks in-depth descriptions of specific machine-learning algorithms, datasets and performance indicators. When comparing the predictive and identification efficacy of various approaches, this discrepancy creates a challenge. Some approaches involved data augmentation or ablation approach to train the models. A new performance matrix is required to complement the current evaluation metrics, like accuracy, RMSE, Confusion Matrix and k-fold validation. The new performance metrics should be able to encompass the differences between models which applied different machine-learning algorithms, signal fusions and overfitting/underfitting regardless small/large capacity of data.

Furthermore, to improve the uncertainty and explainability of proposed models, it is essential to explore publicly available datasets with diverse modalities. Enhancing model interpretability is crucial for industry professionals, as understanding how models generate predictions is vital for trust and effective use of these technologies. Approaches such as Explainable AI, Interpretable AI, Responsible AI and Generative AI provide valuable tools and frameworks to facilitate the understanding and interpretation of machine-learning predictions. Integrated with various Google products and services, these approaches help in troubleshooting and refining model performance while also aiding in comprehending how models function. Applying these methods to each testing model can address the challenge of model interpretability.

ACKNOWLEDGEMENTS

The first author would like to thank and acknowledge the financial support from the Public Service Department (PSD) for their generosity in giving the sponsorship of “Hadiah Latihan Persekutuan”. The authors take responsibility for the integrity of the data and the accuracy of the data analysis. The authors also thank the anonymous reviewers and the editor for their helpful comments.

REFERENCES

[1] Diagnostic and Statistical Manual of Mental Disorders (DMS-5-TR), 5th edn. Arlington, American Psychiatric Association, 2013.

[2] D. J. Morris-Rosendahl and M. A. Crocq, "Neurodevelopmental Disorders-the History and Future of a Diagnostic Concept," Dialogues in Clinical Neuroscience, vol. 22, no. 1, pp. 65–72, Mar. 2020.

[3] M. J. Maenner et al., "Prevalence and Characteristics of Autism Spectrum Disorder among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2020," MMWR, Surveillance Summaries, vol. 72, no. 2, pp. 1–14, Mar. 2023.

[4] R. Alfred and J. H. Obit, "The Roles of Machine Learning Methods in Limiting the Spread of Deadly Diseases: A Systematic Review," Heliyon, vol. 7, no. 6, Elsevier Ltd., Jun. 01, 2021.

[5] B. Kitchenham et al., "Systematic Literature Reviews in Software Engineering: A Systematic Literature Review," Information and Software Technology, vol. 51, no. 1. pp. 7–15, Jan. 2009.

[6] D. Delisle-Rodriguez et al., "Multi-channel EEG-based BCI Using Regression and Classification Methods for Attention Training by Serious Game," Biomedical Signal Processing and Control, vol. 85, DOI: 10.1016/j.bspc.2023.104937, Aug. 2023.

[7] S. Shilaskar et al., "Prediction of Autism and Dyslexia Using Machine Learning and Clinical Data Balancing," Proc. of the 2023 IEEE Int. Conf. on Advances in Intelligent Computing and Applications (AICAPS 2023), DOI: 10.1109/AICAPS57044.2023.10074161, 2023.

[8] K. F. Kollias et al., "Autism Detection in High- functioning Adults with the Application of Eye-tracking Technology and Machine Learning," Proc. of the 2022 11th IEEE Int. Conf. on Modern Circuits and Systems Technologies (MOCAST 2022), DOI: 10.1109/MOCAST54814.2022.9837653, 2022.

[9] M. Bala, M. H. Ali, M. S. Satu, K. F. Hasan and M. A. Moni, "Efficient Machine Learning Models for Early Stage Detection of Autism Spectrum Disorder," Algorithms, vol. 15, no. 5, DOI: 10.3390/a15050166, May 2022.

[10] R. M. Kannan and R. Sasikala, "Predicting Autism in Children at an Early Stage Using Eye Tracking," Proc. of the 2nd IEEE Int. Conf. on Vision towards Emerging Trends in Communication and Networking Technologies (ViTECoN 2023), DOI: 10.1109/ViTECoN58111.2023.10157663, 2023.

[11] A. R. Khan et al., "EXECUTE: Exploring Eye Tracking to Support E-learning," Proc. of the IEEE Global Engineering Education Conf. (EDUCON), pp. 670–676, Tunis, Tunisia, 2022.

[12] Ramanjot, D. Singh, M. Rakhra and S. Aggarwal, "Autism Spectrum Disorder Detection Using the Deep Learning Approaches," Proc. of IEEE Int. Conf. on Technological Advancements in Computational Sci. (ICTACS 2022), pp. 761– 766, Tashkent, Uzbekistan, 2022.

[13] N. Zaman, J. Ferdus and A. Sattar, "Autism Spectrum Disorder Detection Using Machine Learning Approach," Proc. of the 2021 IEEE 12th Int. Conf. on Computing Communication and Networking Technologies (ICCCNT 2021), DOI: 10.1109/ICCCNT51525.2021.9579522, 2021.

[14] V. Chakraborty et al., "Intelligent Dyslexia Prediction Empowered with Optimization Techniques," Proc. of the 2021 6th IEEE Int. Conf. on Recent Trends on Electronics, Inf., Comm. and Tech. (RTEICT 2021), pp. 412–415, DOI: 10.1109/RTEICT52294.2021.9573823, 2021.

[15] F. Thabtah and D. Peebles, "A New Machine Learning Model Based on Induction of Rules for Autism Detection," Health Informatics J., vol. 26, no. 1, pp. 264–286, Mar. 2020.

[16] F. Thabtah, "An Accessible and Efficient Autism Screening Method for Behavioural Data and Predictive Analyses," Health Informatics J., vol. 25, no. 4, pp. 1739–1755, Dec. 2019.

[17] O. Atila et al., "LSGP-USFNet: Automated Attention Deficit Hyperactivity Disorder Detection Using Locations of Sophie Germain's Primes on Ulam's Spiral-based Features with Electroencephalogram Signals," Sensors, vol. 23, no. 16, DOI: 10.3390/s23167032, Aug. 2023.

[18] P. V. Redondo, R. Huser and H. Ombao, "Functional-coefficient Models for Multi-variate Time Series in Designed Experiments: With Applications to Brain Signals," arXiv: 2208.00292, [Online], Available: http://arxiv.org/abs/2208.00292, 2022.

[19] S. Subramaniam, "Enabling Innovative Technologies for Global Healthcare," 42nd Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society Proc., Institute of Electrical and Electronics Engineers and Canadian Medical and Biological Engineering Society, Montreal, Canada, July 2020.

[20] A. Ekhlasi, A. M. Nasrabadi and M. R. Mohammadi, "Analysis of Effective Connectivity Strength in Children with Attention Deficit Hyperactivity Disorder Using Phase Transfer Entropy," Iran J. Psychiatry, vol. 16, no. 4, pp. 374–382, DOI: 10.18502/ijps.v16i4.7224, Oct. 2021.

[21] D. Hai Shi Da Xue, J. Da Xue, M. IEEE Systems, M. IEEE Systems, B. International Federation of Automatic Control. Technical Committee on Economic and Institute of Electrical and Electronics Engineers, 2017 Int. Conf. on Information, Cybernetics and Computational Social Systems (IEEE ICCSS 2017), Dalian, Liaoning, China, July 24-26, 2017.

[22] Galgotias University, Universitatea "Aurel Vlaicu" din Arad, IEEE Industry Applications Society and Institute of Electrical and Electronics Engineers, 2020 IEEE 5th Int. Conf. on Computing Communication and Automation (ICCCA), Galgotias University, Noida, India, Oct. 30-31, 2020.

[23] S. S. Beriha, "Computer Aided Diagnosis System to Distinguish ADHD from Similar Behavioral Disorders," Biomedical and Pharmacology Journal, vol. 11, no. 2, pp. 1135–1141, Jun. 2018.

[24] L. Alice, B. K. S Jacob and S. Ramachandran, "ADHD Detection Based on Time Frequency Image and Deep- learning CNN from Event-related EEG," DOI: 10.21203/rs.3.rs-2986442/v1, 2023.

[25] D. Wang, D. Hong and Q. Wu, "Attention Deficit Hyperactivity Disorder Classification Based on Deep Learning," IEEE/ACM Trans. Comput. Biol. Bioinform., vol. 20, no. 2, pp. 1581–1586, Mar. 2023.

[26] S. M. Omar, M. Kimwele, A. Olowolayemo and D. M. Kaburu, "Enhancing EEG Signals Classification Using LSTM-CNN Architecture," Engineering Reports, DOI: 10.1002/eng2.12827, 2023.

[27] N. Gupte, M. Patel, T. Pen and S. Kurhade, "Early Detection of ADHD and Dyslexia from EEG Signals," Proc. of the 2023 IEEE 8th Int. Conf. for Convergence in Technology (I2CT 2023), DOI: 10.1109/I2CT57861.2023.10126272, 2023.

[28] M. Maniruzzaman et al., "Optimal Channels and Features Selection Based ADHD Detection from EEG Signal Using Statistical and Machine Learning Techniques," IEEE Access, vol. 11, pp. 33570–33583, DOI: 10.1109/ACCESS.2023.3264266, 2023.

[29] A. Afzali et al., "Automated Major Depressive Disorder Diagnosis Using a Dual-input Deep Learning Model and Image Generation from EEG Signals," Waves in Random and Complex Media, pp. 1–16, DOI: 10.1080/17455030.2023.2187237, Mar. 2023.

[30] Y. Sharma and B. K. Singh, "Classification of Children with Attention-Deficit Hyperactivity Disorder Using Wigner-Ville Time-Frequency and Deep expEEGNetwork Feature-based Computational

Models," IEEE Trans. Med. Robot Bionics, vol. 5, no. 4, pp. 890–902, Nov. 2023.
[31] I. W. Huang et al., "Optimal EEG Data Segmentation in LSTM Networks for Learning Neural Dynamics of ADHD," Proc. of the IEEE 2022 Int. Conf. on System Science and Engineering (ICSSE 2022), pp. 33–38, DOI: 10.1109/ICSSE55923.2022.9948260, 2022.
[32] Y. Zhang et al., "ADHD Classification by Feature Space Separation with Sparse Representation," Proc. of the 2018 IEEE 23rd Int. Conf. on Digital Signal Processing (DSP), DOI: 10.1109/ICDSP.2018.8631658, Shanghai, China, 2019.
[33] B. Taghi Beyglou et al., "Detection of ADHD cases using CNN and Classical Classifiers of Raw EEG," Computer Methods and Programs in Biomedicine Update, vol. 2, p. 100080, Jan. 2022.
[34] C. Wang, "Indentification of Autism Spectrum Disorder Based on an Improved Convolutional Neural Networks," Proc. of the 2021 IEEE 3rd Int. Conf. on Machine Learning, Big Data and Business Intelligence (MLBDBI 2021), pp. 235–239, DOI: 10.1109/MLBDBI54094.2021.00051, 2021.
[35] M. Maniruzzaman et al., "Efficient Feature Selection and Machine Learning Based ADHD Detection Using EEG Signal," Computers, Materials and Continua, vol. 72, no. 3, pp. 5179–5195, 2022.
[36] M. Mafi and S. Radfar, "High Dimensional Convolutional Neural Network for EEG Connectivity-based Diagnosis of ADHD," J. of Biomedical Physics and Engineering, vol. 12, no. 6, pp. 645–654, 2022.
[37] F. W. Alsaade and M. S. Alzahrani, "Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms," Computational Intelligence and Neuroscience, vol. 2022, DOI: 10.1155/2022/8709145, 2022.
[38] I. A. Ahmed et al., "Eye Tracking-based Diagnosis and Early Detection of Autism Spectrum Disorder Using Machine Learning and Deep Learning Techniques," Electronics (Switzerland), vol. 11, no. 4, DOI: 10.3390/electronics11040530, Feb. 2022.
[39] U. B. Mahadevaswamy and C. Manjunath, "f-MRI Based Detection of Autism Using CNN Algorithm," Proc. of the 2022 IEEE 2nd Mysore Sub Section Int. Conf. (MysuruCon 2022), DOI: 10.1109/MysuruCon55714.2022.9972394, 2022.
[40] B. R. G. Elshoky et al., "Comparing Automated and Non-automated Machine Learning for Autism Spectrum Disorders Classification Using Facial Images," ETRI J., vol. 44, no. 4, pp. 613–623, 2022.
[41] M. R. Kanhirakadavath et al., "Investigation of Eye-tracking Scan Path As a Biomarker for Autism Screening Using Machine Learning Algorithms," Diagnostics, vol. 12, no. 2, 2022.
[42] M. Saini, U. Satija and M. D. Upadhayay, "One-dimensional Convolutional Neural Network Architecture for Classification of Mental Tasks from Electroencephalogram," Biomed Signal Process Control, vol. 74, DOI: 10.1016/j.bspc.2022.103494, Apr. 2022.
[43] S. Altun, A. Alkan and H. Altun, "Automatic Diagnosis of Attention Deficit Hyperactivity Disorder with Continuous Wavelet Transform and Convolutional Neural Network," Clinical Psychopharmacology and Neuroscience, vol. 20, no. 4, pp. 715–724, 2022.
[44] S. Saini, R. Rani and N. Kalra, "Prediction of Attention Deficit Hyperactivity Disorder (ADHD) Using Machine Learning Techniques Based on Classification of EEG Signal," Proc. of the IEEE 8th Int. Conf. on Advanced Comp. and Comm. Systems (ICACCS 2022), pp. 782–786, Coimbatore, India, 2022.
[45] T. Tuncer, S. Dogan and A. Subasi, "LEDPatNet19: Automated Emotion Recognition Model Based on Non-linear LED Pattern Feature Extraction Function Using EEG Signals," Cognitive Neurodynamics, vol. 16, no. 4, pp. 779–790, Aug. 2022.
[46] N. A. Mashudi, N. Ahmad and N. M. Noor, "Classification of Adult Autistic Spectrum Disorder Using Machine Learning Approach," IAES Int. J. of Artificial Intelligence, vol. 10, no. 3, pp. 743–751, 2021.
[47] M. F. Rabbi et al., "A Convolutional Neural Network Model for Early-stage Detection of Autism Spectrum Disorder," Proc. of the 2021 IEEE Int. Conf. on Information and Communication Technology for Sustainable Development (ICICT4SD 2021), pp. 110–114, Dhaka, Bangladesh, Feb. 2021.
[48] G. P. Pralhad et al., "Dyslexia Prediction Using Machine Learning," Proc. of the 2021 1st IEEE Int. Conf. on Artificial Intelligence and Machine Vision (AIMV 2021), DOI: 10.1109/AIMV53313.2021.9671004, Gandhinagar, India, 2021.
[49] Y. H. Chen et al., "Early Detection of Autism Spectrum Disorder in Young Children with Machine Learning Using Medical Claims Data," BMJ Health Care Inform, vol. 29, no. 1, p. 100544, Sep. 2022.
[50] S. Karimi-Shahraki and M. Khezri, "Identification of Attention Deficit Hyperactivity Disorder Patients Using Wavelet-based Features of EEG Signals," JIPET J. of Intelligent Procedures in Electrical Technology, vol. 12, no. 47, pp. 29–40, [Online], Available: http://jipet.iaun.ac.ir/, 2021.
[51] N. Kaur and G. Gupta, "Refurbished and Improvised Model Using Convolution Network for Autism Disorder Detection in Facial Images," Indonesian Journal of Electrical Engineering and Computer Science, vol. 29, no. 2, pp. 883–889, DOI: 10.11591/ijeecs.v29.i2.pp883-889, Feb. 2023.
[52] S. Raj and S. Masood, "Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques," Procedia Computer Science, pp. 994–1004, DOI: 10.1016/j.procs.2020.03.399, 2020.
[53] Z. Sherkatghanad et al., "Automated Detection of Autism Spectrum Disorder Using a Convolutional Neural Network," Frontiers in Neuroscience, vol. 13, DOI: 10.3389/fnins.2019.01325, Jan. 2020.
[54] M. Rakić, M. Cabezas, K. Kushibar, A. Oliver and X. Lladó, "Improving the Detection of Autism

Spectrum Disorder by Combining Structural and Functional MRI Information," NeuroImage: Clinical, vol. 25, DOI: 10.1016/j.nicl.2020.102181, Jan. 2020.

[55] M. Moghaddari, M. Z. Lighvan and S. Danishvar, "Diagnose ADHD disorder in Children Using Convolutional Neural Network Based on Continuous Mental Task EEG," Computer Methods and Programs in Biomedicine, vol. 197, DOI: 10.1016/j.cmpb.2020.105738, Dec. 2020.

[56] Z. Zhao et al., "Applying Machine Learning to Identify Autism with Restricted Kinematic Features," IEEE Access, vol. 7, pp. 157614–157622, DOI: 10.1109/ACCESS.2019.2950030, 2019.

[57] R. Yaghoobi, S. Azadi and P. Keshavarzi, "Loss Detection of Recurrence Rate in the EEG Signals of Children with ADHD," NIScPR Online Periodicals Repository, JSIR, vol.78, no. 04, April 2019.

[58] G. A. Senthil et al., "A Novel Analysis and Detection of Autism Spectrum Disorder in Artificial Intelligence Using Hybrid Machine Learning," Proc. of the IEEE Int. Conf. on Innovative Data Communication Technologies and Application (ICIDCA), pp. 291–296, Uttarakhand, India, 2023.

[59] A. E. Alchalabi et al., "FOCUS: Detecting ADHD Patients by an EEG-based Serious Game," IEEE Transactions on Instrumentation and Measurement, vol. 67, no. 7, pp. 1512–1520, Jul. 2018.

[60] M. R. Mohammadi et al., "EEG Classification of ADHD and Normal Children Using Non-linear Features and Neural Network," Biomedical Engineering Letters, vol. 6, no. 2, pp. 66–73, May 2016.

[61] A. J. Hao, B. Lianghua He and H. Yin, "Discrimination of ADHD Children Based on Deep Bayesian Network," Proc. of 2015 IET Int. Conf. on Biomedical Image and Signal Processing (ICBISP 2015), DOI: 10.1049/cp.2015.0764, Beijing, China, 2015.

[62] M. De Oliveira Meira et al., "Evaluating Brain Regions that Characterize Attention Deficit/Hyperactivity Disorder Based on SPECT Images and Machine Learning Models," Proc. of the IEEE Int. Joint Conf. on Neural Networks, DOI: 10.1109/IJCNN55064.2022.9892968, 2022.

[63] S. K. Khare, N. B. Gaikwad and V. Bajaj, "VHERS: A Novel Variational Mode Decomposition and Hilbert Transform-based EEG Rhythm Separation for Automatic ADHD Detection," IEEE Transactions on Instrumentation and Measurement, vol. 71, DOI: 10.1109/TIM.2022.3204076, 2022.

[64] M. Aceves-Fernandez and M. A. Aceves-Fernandez, "Methodology Proposal of ADHD Classification of Children Based on Cross Recurrence Plots Methodology Proposal of ADHD Classification of Children based on Cross Recurrence Plots," Non Linear Dynamics, vol. 104, pp. 1491-1505, 2021.

[65] S. Sangeetha, R. Uma and R. Valarmathi, "Dyslexia Biomarker Detection Using Machine Learning," Proc. of the 2022 Int. Conf. on Communication, Computing and Internet of Things (IC3IoT 2022), DOI: 10.1109/IC3IOT53935.2022.9767973, 2022.

[66] A. Parashar et al., "Machine Learning Based Framework for Classification of Children with ADHD and Healthy Controls," Intelligent Automation and Soft Computing, vol. 28, no. 3, pp. 669– 682, 2021.

[67] B. Wingfield et al., "A Predictive Model for Paediatric Autism Screening," Health Informatics J., vol. 26, no. 4, pp. 2538–2553, DOI: 10.1177/1460458219887823, Dec. 2020.

[68] S. Liu et al., "Deep Spatio-temporal Representation and Ensemble Classification for Attention Deficit/Hyperactivity Disorder," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 29, pp. 1–10, DOI: 10.1109/TNSRE.2020.3019063, 2021.

[69] R. Sujatha et al., "A Machine Learning Way to Classify Autism Spectrum Disorder," International Journal of Emerging Technologies in Learning, vol. 16, no. 6, pp. 182–200, 2021.

[70] T. Akter et al., "Machine Learning-Based Models for Early Stage Detection of Autism Spectrum Disorders," IEEE Access, vol. 7, pp. 166509–166527, DOI: 10.1109/ACCESS.2019.2952609, 2019.

[71] Y. Jayawardana et al., "Analysis of Temporal Relationships between ASD and Brain Activity through EEG and Machine Learning," Proc. of the 2019 IEEE 20th Int. Conf. on Information Reuse and Integration for Data Science (IRI), DOI: 10.1109/IRI.2019.00035, Los Angeles, USA, 2019.

[72] S. Pulido-Castro et al., "Ensemble of Machine Learning Models for an Improved Facial Emotion Recognition," Proc. of the 2021 IEEE URUCON, pp. 512–516, Montevideo, Uruguay, 2021.

[73] P. Thi Viet Huong et al., "Ensemble Learning in Detecting ADHD Children by Utilizing the Non-linear Features of EEG Signal*," Proc. of the 2nd Int. Conf. on Human-centered Artificial Intelligence (Computing4Human 2021), [Online], Available: http://ceur-ws.org, 2021.

[74] R. Holker and S. Susan, "Computer-aided Diagnosis Framework for ADHD Detection Using Quantitative EEG," Part of Book: Lecture Notes in Computer Science (Including Sub-series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 229– 240, DOI: 10.1007/978-3-031-15037-1_19, Springer, 2022.

[75] S. Batsakis et al., "Data-driven Decision Support for Adult Autism Diagnosis Using Machine Learning," Digital, vol. 2, no. 2, pp. 224–243, DOI: 10.3390/digital2020014, Jun. 2022.

[76] N. Hamedi et al., "An Effective Connectomics Approach for Diagnosing ADHD Using Eyes-open Resting-state MEG," Proc. of the 11th IEEE Int. Conf. on Computer Engineering and Knowledge (ICCKE 2021), pp. 110–114, DOI: 10.1109/ICCKE54056.2021.9721443, Mashhad, Iran, 2021.

[77] A. JothiPrabha, R. Bhargavi and B. V. Deepa Rani, "Prediction of Dyslexia Severity Levels from Fixation and Saccadic Eye Movement Using Machine Learning," Biomed Signal Process Control, vol. 79, DOI: 10.1016/j.bspc.2022.104094, Jan. 2022.

[78] S.-F. Liang et al., "Differentiation between Resting-state fMRI Data from ADHD and Normal Subjects: Based on Functional Connectivity and Machine Learning," Proc. of the 2012 Int. Conf. on Fuzzy Theory and Its Applications (iFUZZY2012), DOI: 10.1109/iFUZZY.2012.6409719, Taiwan, 2012.

[79] Y. Ding, Y. Li, S. Li, Z. Fan and L. Wang, 2011 8th Int. Conf. on Fuzzy Systems and Knowledge Discovery (FSKD 2011) Proceedings, pp. 1349-2073, Shanghai, China, 26-28 July 2011.

[80] A. Alim and M. H. Imtiaz, "Automatic Identification of Children with ADHD from EEG Brain Waves," Signals, vol. 4, no. 1, pp. 193–205, DOI: 10.3390/signals4010010, Mar. 2023.

[81] EEE, IEEE Engineering in Medicine and Biology Society and IEEE Signal Processing Society, IEEE ISBI 2020 Int. Symposium on Biomedical Imaging (2020 Symposium Proc.), Iowa City, Iowa, 2020.

[82] IEEE, International Symposium on Biomedical Imaging, 2018 IEEE 15th Int. Symposium on Biomedical Imaging (ISBI 2018), DOI: 10.1109/ISBI39256.2018, 2018.

[83] H. Alkahtani et al., "Developing System-based Artificial Intelligence Models for Detecting the Deficit Hyperactivity Disorder," Mathematics, vol. 11, no. 22, DOI: 10.3390/math11224698, Nov. 2023.

[84] G. Sharma and A. M. Joshi, "SzHNN: A Novel and Scalable Deep Convolution Hybrid Neural Network Framework for Schizophrenia Detection Using Multichannel EEG," IEEE Transactions on Instrumentation, vol. 71, DOI: 10.1109/TIM.2022.3212040, 2022.

[85] A. Sharma and P. Tanwar, "Identification of Autism Spectrum Disorder (ASD) from Facial Expressions Using Deep Learning," Proc. of the 2022 IEEE Int. Conf. on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON 2022), pp. 478–484, Faridabad, India, 2022.

[86] S. Dalai, Jadavpur University, Institute of Electrical and Electronics Engineers, Kolkata Section and IEEE Signal Processing Society, Kolkata Chapter, Proc. of 2018 IEEE Applied Signal Processing Conference (ASPCON), Venue: Jadavpur University Main Campus, Kolkata, India, Dec. 7-9, 2018.

[87] A. A. Torres-García et al., Biosignal Processing and Classification Using Computational Learning and Intelligence: Principles, Algorithms and Applications, 1st Edn., ISBN: 9780128201251, Elsevier, 2022.

ملخص البحث:

تقدّم هذه الدراسة مراجعة شاملة للأدبيات السابقة تركّز على استخدام تقنيات التّعلّم الآلي والتّعلّم العميق لتوقُّع اضطرابات النّموّ العصبي والكشف عنها؛ مثل التّخلّف العقلي، واضطراب طيْف التّوحُّد، واضطراب نقْص الانتباه وفرط النّشاط، والديسليكسيا، وغيرها. حيث تفتقر المراجعات المتوافرة إلى النّقاشات التّفصيلية لخوارزميات التّعلّم الآلي ومجموعات البيانات ومؤشِّرات الأداء المستخدمة في توقُّع اضطرابات النّموّ العصبي والكشف عنها، وتسعى هذه الدّراسة إلى جَسْر هذه الفجوة عبر تناول جانبين منفصلين هما التّوقّع والكشْف. كما تهدف الدّراسة إلى البحث في آخر ما توصّلت إليه الأبحاث العلمية بالمنهجيات والتّحدّيات واتّجاهات البحث المستقبلية بشأن استخدام تقنيات التّعلّم الآلي والتّعلّم العميق في توقُّع اضطرابات النّموّ العصبي والكشف عنها. وتهدف إلى تصنيف الدّراسات السابقة تبعاً للجانبين الرئيسيّين المتمثّلين في توقّع اضطرابات النّموّ العصبي والكشف عنها، إلى جانب التّطرُّق إلى المنهجيات ومجموعات البيانات والمتغيرات ومؤشّرات الآداء الّتي استخدمتها الدراسات السابقة حول الموضوع.

شملت المراجعة الدّراسات المنشورة في المجلات والمؤتمرات المتخصصة والمفهرسة في Scopus في الفترة من عام 2013 إلى عام 2023. واستخدمت المراجعة مصطلحات للبحث، مثل: توقّع اضطرابات النّموّ العصبي، والكشف عن اضطرابات النّموّ العصبي، باستخدام التّعلّم الآلي. وركّز التّحليل على تحديد منهجيات التّعلّم الآلي والتّعلّم العميق، والنّماذج المجمَّعة، وأنواع مجموعات البيانات، بالإضافة إلى المتغيرات ومؤشرات الأداء المستخدمة في الدراسات السابقة. ولقد ألقت نتائج المراجعة الضّوء على أكثر تقنيات التّعلّم الآلي والتّعلّم العميق انتشاراً، والتّحدّيات المرتبطة بالبحث في هذا المجال، واتّجاهات البحث المستقبلية الرّامية إلى تحسين الخدمات المقدَّمة إلى مجتمع اضطرابات النّموّ العصبي؛ من أجل تطوير الرّعاية الصّحية للمصابين بهذه الاضطرابات عبر تقنيات توقّعٍ وكشْفٍ أفضل.