SOMAIYA
VIDYAVIHAR

K J Somaiya Institute of Technology
An Autonomous Institute Permanently Affiliated to the University of Mumbai

DEPARTMENT OF INFORMATION TECHNOLOGY

Synopsis of Minor Project On

Enhancing ADHD detection through AI

Prepared By:

Bhumika Baria (Roll No. 02)
Yashi Depani (Roll No. 08)
Priti Prasad (Roll No. 36)
Rhea Laloo (Roll No. 57)

Under the guidance of:

Prof.Sarita Rathod

Department of Information Technology

Academic Year: 2023-2024

Autonomy Syllabus Scheme-II - Semester VI (TY - IT)

SOMAIYA
VIDYAVIHAR

K J Somaiya Institute of Technology
An Autonomous Institute Permanently Affiliated to the University of Mumbai

DEPARTMENT OF INFORMATION TECHNOLOGY

CERTIFICATE

This is to certify that following students:

	Roll No. / Seat No.
Bhumika Baria	02
Yashi Depani	08
Priti Prasad	36
Rhea Laloo	57

have submitted PBL – Minor Project I Report on *"Enhanced ADHD Detection through AI"* as the partial fulfillment for the requirement of Third Year of Engineering (6^{th} Semester) in T.Y. - Information Technology under my guidance during the academic year 2023-2024.

Prof. Sarita Rathod
Project Guide
Assistant Professor
Department of Information Technology

Dr. Radhika Kotecha
Head of Department
Professor
Department of Information Technology

Date of Examination: __________________

Signature of Internal Examiner

Signature of External Examiner

Table of Contents

Acknowledgement

In performing our project, we had to take the help and guideline of some professors, who deserve our greatest gratitude. We express our profound gratitude to **Prof. Sarita Rathod**, our project guide and Minor-Project coordinator, for her invaluable guidance and meticulous attention to detail, which have been instrumental in the successful completion of our assignment.Our special gratitude to **Dr. Radhika Kotecha**, Head of Department of Information Technology, **Dr. Sunita Patil**, Vice- Principal and **Dr. Vivek Sunnapwar**, Principal of K. J. Somaiya Institute of Technology for mentoring out for our minor project throughout numerous consultations. We would also like to expand our deepest gratitude to all those who have directly and indirectly guided us in this project. Many people, especially our seniors have made valuable comment suggestions on this proposal which gave us inspiration to improve our project **'Enhancing ADHD Detection using AI'**. We would like to acknowledge that this project was completed entirely by our group. Henceforth, we thank all the people for their help directly and indirectly to complete our project.

Abstract

The landscape of Attention Deficit Hyperactivity Disorder (ADHD) detection is marked by a pressing need for a universally optimal algorithm. Currently, the absence of such a standardized approach leads to the proliferation of multiple algorithms, contributing to inconsistencies in accuracy and posing significant challenges to precise and consistent ADHD identification. This critical gap in standardized methodologies impedes effective detection, underscoring the urgent necessity for the development of reliable methodologies. In response to this imperative, the demand for Artificial Intelligence (AI)-based technologies to support ADHD detection emerges as paramount. This paper delves into the pressing issue surrounding the current state of ADHD detection, emphasizing the shortcomings in existing methodologies and the imperative for AI-driven solutions. Through an extensive exploration of existing literature, we elucidate the necessity for machine learning-based AI systems to address the deficiencies in ADHD detection. By leveraging AI technologies, particularly machine learning algorithms, this paper proposes a paradigm shift towards enhanced ADHD detection characterized by improved diagnostic accuracy and consistency. Central to our approach is the recognition of the complex interplay of factors contributing to ADHD and the inherent variability in symptom presentation among individuals. Traditional ADHD detection methods often rely on subjective assessments and standardized questionnaires, leading to inconsistencies in diagnosis. In contrast, AI-driven systems have the potential to integrate diverse sources of data, including behavioral, cognitive, and neuroimaging metrics, to provide a comprehensive and objective assessment of ADHD. Moreover, our paper underscores the need for robust validation and standardization protocols to ensure the reliability and generalizability of AI-based ADHD detection systems. By synthesizing insights from existing research, we identify key challenges and opportunities in the development and deployment of AI-driven solutions for ADHD detection. We propose a holistic framework that integrates advanced machine learning techniques with domain-specific expertise to optimize diagnostic accuracy while minimizing false positives and false negatives. Furthermore, we advocate for interdisciplinary collaboration among researchers, clinicians, and technologists to harness the full potential of AI in ADHD detection. By fostering synergy between domain knowledge and technological innovation, we can accelerate the translation of AI-driven solutions from research laboratories to clinical practice. In conclusion, this paper presents a comprehensive overview of the landscape of ADHD detection and underscores the transformative potential of AI-driven methodologies. By addressing the limitations of existing approaches and leveraging the capabilities of AI technologies, we aim to advance the field of ADHD detection towards greater accuracy, consistency, and clinical utility.

Chapter 1: Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder characterized by persistent patterns of inattention, hyperactivity, and impulsivity, often presenting challenges in academic, occupational, and social functioning. Timely and accurate identification of ADHD is essential for implementing effective interventions and support strategies, as early intervention can significantly improve long-term outcomes for individuals affected by the disorder. The conventional diagnostic process for ADHD typically involves subjective assessments based on standardized questionnaires and clinical interviews.

In recent years, advances in machine learning algorithms have opened up new avenues for refining ADHD detection by leveraging diverse datasets containing biological, psychological, and behavioral markers associated with the disorder. Machine learning techniques, such as decision trees, support vector machines, and neural networks, offer unique advantages in analyzing complex patterns within multidimensional data, thus enhancing the ability to discern subtle differences indicative of ADHD. The integration of machine learning into the diagnostic process holds promise for improving both the accuracy and efficiency of ADHD assessments. By systematically analyzing large volumes of data from various sources, machine learning algorithms can identify patterns and associations that may not be readily apparent to human clinicians, thereby augmenting diagnostic decision-making and reducing the likelihood of diagnostic errors.

Furthermore, machine learning-based ADHD detection has the potential to facilitate more personalized and tailored interventions for individuals with ADHD. By providing insights into underlying neurobiological mechanisms and individual differences in symptom profiles, machine learning algorithms can inform the development of targeted interventions that address specific needs and challenges faced by each patient. However, the adoption of machine learning-driven diagnostic systems in healthcare settings raises ethical considerations that must be carefully addressed.

1.1 Motivation

The motivation behind exploring machine learning algorithms for the refinement of ADHD detection stems from the pressing need for timely and accurate identification of this neurodevelopmental disorder. ADHD significantly impacts cognitive functions, attention span, and impulse control, posing challenges in academic, occupational, and social domains. Furthermore, the complexity and variability of ADHD symptoms across individuals underscore the necessity for more sophisticated and nuanced diagnostic approaches. Traditional methods of ADHD assessment often rely on subjective observations and standardized questionnaires, which may not capture the full spectrum of symptoms or account for individual differences in presentation. As a result, there is a growing recognition of the limitations of conventional diagnostic practices and a corresponding demand for more objective and data-driven approaches.

1.2 Problem Analysis

- Inconsistencies in Accuracy: Multiple existing algorithms exhibit inconsistencies in accuracy when it comes to detecting ADHD.
- Challenges in Identifying Individuals with ADHD: The variation in algorithms exacerbates the challenges in accurately identifying individuals with ADHD.
- Need for Unified and Reliable Approach: The discrepancies in existing algorithms underscore the need for concerted efforts towards developing a unified and reliable approach to ADHD diagnosis.

1.3 Objectives

- To Assess various ADHD detection algorithms: To evaluate and compare different ADHD detection algorithms to determine their efficacy and accuracy in identifying individuals with ADHD
- To Compare algorithm strengths and weaknesses in diverse datasets: To analyze and contrast the strengths and weaknesses of different ADHD detection algorithms across diverse datasets to identify their performance variations and limitations in accurately detecting individuals with ADHD
- To Choose a consistently high-performing algorithm: To select an algorithm that consistently demonstrates high performance across various evaluation metrics and datasets for the accurate detection of individuals with ADHD.
- To Develop a standardized ADHD detection process: To create a standardized ADHD detection process that integrates the selected high-performing algorithm, ensuring consistent and reliable identification of individuals with ADHD across clinical settings.

1.4 Scope

The scope of this paper on enhanced ADHD detection through AI involves exploring the potential of machine learning algorithms to refine the identification of ADHD, including an assessment of various AI techniques, evaluation of their performance metrics, and discussion of ethical considerations, with the ultimate goal of improving diagnostic accuracy and tailoring interventions for individuals with ADHD.

Chapter 2: Literature Review

1.1 Related Work

A comprehensive literature survey on enhanced ADHD detection through AI illuminates a dynamic landscape characterized by innovative techniques aimed at refining the identification and management of Attention Deficit Hyperactivity Disorder. Researchers and practitioners are actively exploring a diverse range of AI-driven approaches to enhance diagnostic accuracy and tailor interventions for individuals with ADHD. These approaches encompass various machine learning algorithms, including deep learning models, decision trees, and support vector machines, each offering unique advantages in analyzing heterogeneous datasets associated with ADHD symptoms. Moreover, the literature survey highlights the emergence of novel techniques such as multimodal data fusion, which integrates information from multiple sources such as behavioral, neuroimaging, and genetic data to improve the robustness and reliability of ADHD detection.

A. Machine Learning in ADHD and Depression Mental Health Diagnosis

In 2023, research in the domain of mental health diagnosis, particularly focused on Attention Deficit Hyperactivity Disorder (ADHD) and depression, has witnessed significant advancements in utilizing machine learning techniques, notably Support Vector Machines (SVMs) and Convolutional Neural Networks (CNNs), particularly for analyzing Electroencephalography (EEG) data. Studies have reported high classification accuracies, reaching up to 99.58% with SVMs, highlighting the potential of these methods in aiding diagnosis. However, challenges persist, including limited dataset sizes, privacy concerns, and the spectrum of behavior variations inherent in mental health disorders, which pose obstacles to model generalization. To address these challenges, researchers have utilized datasets such as ADHD-200, which includes structural and resting-state functional Magnetic Resonance Imaging (s-MRIs, rs-fMRIs), to augment training data and improve model robustness. Despite these challenges, the integration of machine learning in mental health diagnosis represents a promising avenue for improving diagnostic accuracy and supporting personalized treatment strategies.

B. Detection of ADHD Disorder in Children Using Layer-Wise

In 2024, a study focused on the detection of Attention Deficit Hyperactivity Disorder (ADHD) in children utilized a Convolutional Neural Network (CNN) with Layer-Wise Relevance Propagation (LRP) to achieve an accuracy of 94.52% for ADHD classification. The study utilized the First-National-EEG-Data-Analysis-Competition-with-Clinical-Application dataset. However, challenges such as a relatively small dataset size and limited generalizability due to the specific task used were noted, highlighting the need for further research to address these limitations and enhance the robustness of ADHD detection models.

C. Subspace learning based classification of ADHD patients

In 2023, a study on the classification of Attention Deficit Hyperactivity Disorder (ADHD) patients utilized subspace learning and linear regression techniques, achieving a high accuracy of 94.6% on the ADHD-200 database. The study employed subspace learning methods in conjunction with linear regression to effectively discern patterns and features within the ADHD-200 dataset, facilitating accurate classification of ADHD patients. This research underscores the potential of subspace learning approaches in enhancing the diagnostic capabilities for ADHD and highlights the significance of utilizing large-scale datasets for robust model development and validation.

D. Comprehensive review of EEG data classification techniques for ADHD detection using machine learning and deep learning

In 2023, a comprehensive review focused on EEG data classification techniques for the detection of Attention Deficit Hyperactivity Disorder (ADHD) utilizing a variety of machine learning and deep learning algorithms, including decision trees, Naive Bayes, k-means, support vector machines (SVM), and K Nearest Neighbor (KNN) algorithm. Notably, a study employing support vector machines (SVM) reported achieving high accuracy of 84.6% in ADHD detection. However, the review primarily concentrated on hospital diagnostics, such as MRI and EEG data, potentially limiting generalizability by neglecting other data sources. While the specific mention of this limitation was not found in the paper, it underscores the importance of considering diverse data sources for robust ADHD detection models.

Name	**Year**	**Algorithm**	**Results**	**Limitations**	**Dataset**
Machine Learning in ADHD and Depression Mental Health Diagnosis	2023	SVMs, CNNs (for EEG data) CNNs, SVMs (for EEG data)	High classification accuracies (up to 99.58% with SVMs)	Limited dataset size, privacy concerns, spectrum of behavior variations, challenges in model generalization	ADHD-200 dataset: rs-fMRIs
Detection of ADHD Disorder in Children Using Layer-Wise	2024	Convolutional Neural Network (CNN) with Layer-Wise Relevance Propagation (LRP)	Achieved an accuracy of 94.52% for ADHD classification	Relatively small dataset size. Limited generalizability due to specific tasks used.	First-National-EEG-Data-Analysis-Competition-with-Clinical-Application dataset
Subspace learning based classification of ADHD patients	2023	Subspace learning, Linear regression	Achieved 94.6% accuracy on the ADHD-200 database	Subspace learning, Linear regression	ADHD-200 database
Comprehensive review of EEG data classification techniques for ADHD detection using machine learning and deep learning	2023	decision trees, Naive Bayes, k-means, support vector machines (SVM), and K Nearest Neighbor (KNN) algorithm.	A study using support vector machines (SVM) achieved high accuracy (84.6%)	- Review focuses on hospital diagnostics (MRI, EEG), potentially limiting generalizability by neglecting other data sources.	Not explicitly mentioned in paper.

1.2 Existing System

The existing systems for enhanced ADHD detection through AI primarily leverage machine learning and deep learning techniques to improve diagnostic accuracy and facilitate personalized interventions for individuals with Attention Deficit Hyperactivity Disorder. These systems typically utilize diverse datasets encompassing behavioral, neuroimaging, and genetic data to train and validate predictive models for ADHD diagnosis. Examples of existing AI-driven ADHD detection systems include those that employ support vector machines (SVM),

convolutional neural networks (CNN), and ensemble learning methods. These systems aim to analyze complex patterns and associations within the data to identify distinctive biomarkers and behavioral features indicative of ADHD. Additionally, some systems integrate real-time monitoring and feedback mechanisms to track individuals' attentional states and behavioral patterns, enabling timely interventions and support strategies. Despite the advancements in AI-based ADHD detection, challenges such as limited dataset sizes, privacy concerns, and model generalization issues persist, highlighting the need for ongoing research and development in this field.

Chapter 3: Proposed System

The proposed system aims to revolutionize ADHD detection through the integration of advanced machine learning techniques and rigorous evaluation methodologies. This provides an overview of the methodology employed to develop and evaluate the system, delineating key phases involved in the process. Beginning with the preparatory step, the system sources the ADHD detection dataset from Data World and undergoes preprocessing to ensure data quality. Subsequently, the performance measurement phase entails the implementation and training of machine learning algorithms such as Random Forest, SVM, and Gaussian Naive Bayes. The evaluation of algorithmic efficacy against ground truth labels forms the cornerstone of this phase, with metrics including accuracy, precision and recall employed for comprehensive assessment. Following this, the validation and evaluation phase scrutinizes the reliability and correctness of algorithmic predictions. Through meticulous comparison of performance metrics, the most effective solution for ADHD detection is identified, paving the way for future research and clinical applications. This introduction serves as a roadmap for the subsequent detailed exploration of each phase, offering insights into the development and evaluation of the proposed ADHD detection system.

3.1 Proposed Approach and Details

Our approach to enhancing ADHD detection through AI involves a systematic methodology encompassing several key phases, each meticulously designed to ensure robustness and accuracy in the system's performance.

Phase I: Preparatory Step

In this initial phase, the ADHD detection dataset is acquired from Data World, ensuring diversity and relevance to the task at hand. Preprocessing steps are then applied to the dataset to enhance its quality and suitability for training machine learning models. These preprocessing steps include handling missing values, which involves imputation techniques such as mean or median imputation, normalizing features to ensure uniform scaling across different features, and addressing class imbalances through techniques like oversampling or undersampling to mitigate any biases that may arise during model training.

Phase II: Model Selection and Training

The next phase focuses on selecting appropriate machine learning algorithms for ADHD detection and training them using the preprocessed dataset. We choose to implement and train three widely used classification algorithms: Random Forest, Support Vector Machine (SVM), and Gaussian Naive Bayes. Random Forest is chosen for its ensemble learning capabilities, SVM for its effectiveness in handling high-dimensional data and non-linear relationships, and Gaussian Naive Bayes for its simplicity and efficiency in probabilistic classification tasks. The dataset is split into training, validation, and testing sets to facilitate model training and evaluation. Hyperparameter tuning is performed using techniques such as grid search with cross-validation to optimize the performance of each algorithm.

Phase III: Performance Evaluation

Once the models are trained, their performance is evaluated using a comprehensive set of evaluation metrics including accuracy, precision, recall, F1-score, Receiver Operating Characteristic (ROC) curve, and Area Under the Curve (AUC). These metrics provide valuable insights into the effectiveness of each algorithm in accurately detecting ADHD. Additionally, visualizations such as confusion matrices and ROC curves are generated to further analyze and interpret the performance of the models.

Phase IV: Model Validation and Selection

In this final phase, the reliability and correctness of algorithmic predictions are validated to ensure the robustness of the system. The performance metrics obtained from the evaluation phase are meticulously analyzed to identify the most effective algorithm for ADHD detection. The algorithm demonstrating the highest accuracy and precision is selected as the most suitable for deployment in real-world scenarios. Insights derived from this evaluation guide the identification of the best algorithm for enhancing ADHD detection through AI, potentially influencing future research and clinical applications.

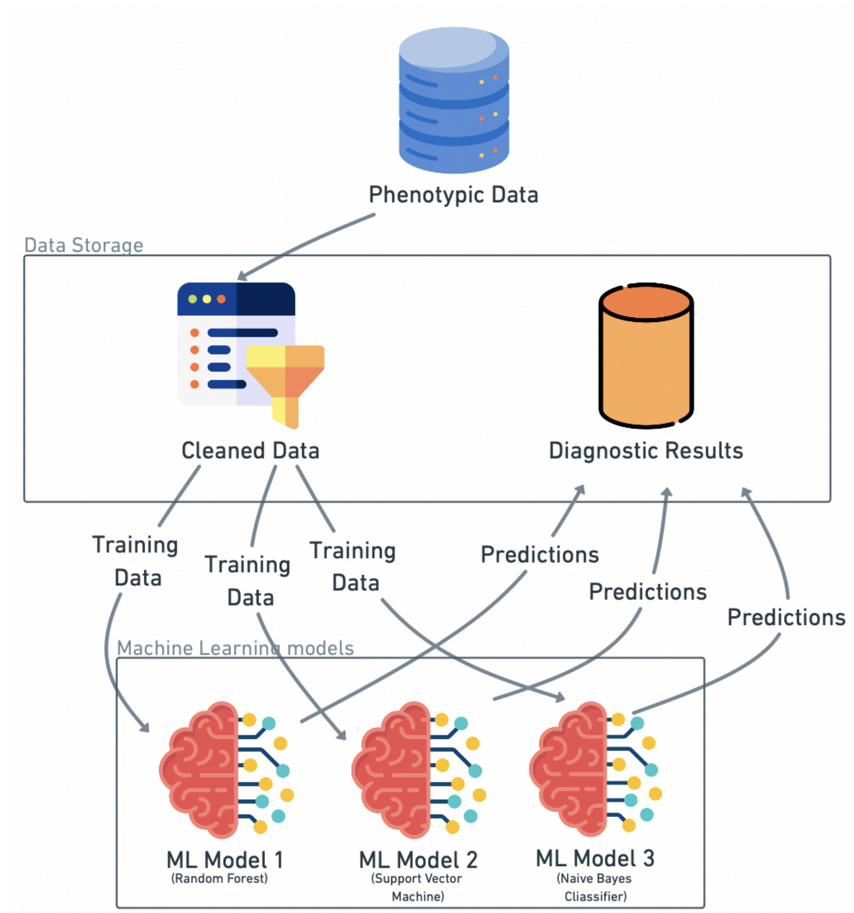

Fig 3.1.1: system architecture

Fig 3.1.1 delineates the system architecture, providing a visual representation of the structural framework and interconnections among its components, elucidating the design and organization of the system under examination.

3.2 Innovation in Idea

- **Dynamic Model Selection**: Develop an intelligent system that dynamically selects the most suitable algorithm based on the characteristics of the input data. This system could analyze data patterns in real-time and automatically choose between SVM, Naive Bayes, or Random Forest to optimize prediction accuracy and efficiency for different types of ADHD-related outcomes.
- **Adaptive Feature Engineering**: Implement a novel feature engineering approach that adapts to the specific needs of ADHD prediction tasks. This could involve incorporating domain-specific knowledge about ADHD symptomatology, neuropsychological measures, and treatment responses into the feature extraction process to enhance the discriminative power of the models.
- **Transfer Learning for ADHD Subtypes**: Explore transfer learning techniques to leverage knowledge from related tasks or domains to improve ADHD subtype classification. By pretraining models on datasets from related conditions or neurodevelopmental disorders, you can transfer learned representations to the ADHD classification task, potentially improving generalization performance, especially for rare subtypes or comorbidities.
- **Interactive Model Interpretation**: Develop interactive visualization tools that allow clinicians to explore and interpret the decision-making processes of the AI models. By visualizing feature importance, decision boundaries, and prediction probabilities in an intuitive and interactive manner, clinicians can gain insights into how the models arrive at their predictions and validate their clinical relevance.
- **Domain Adaptation for Multisite Dat**a: Address the challenge of model generalization across different clinical sites or populations by incorporating domain adaptation techniques into the training process. By aligning feature distributions across diverse datasets, you can improve the robustness and transferability of the models, enabling more effective deployment in real-world clinical settings with heterogeneous patient populations.
- **Active Learning for Data-Efficient Training:** Implement active learning strategies to iteratively select the most informative samples for labeling, thereby reducing the labeling burden and improving model performance with limited annotated data. By actively querying unlabeled data points that are most uncertain or informative, you can accelerate the training process and optimize model performance for ADHD prediction tasks.

3.3 Timeline

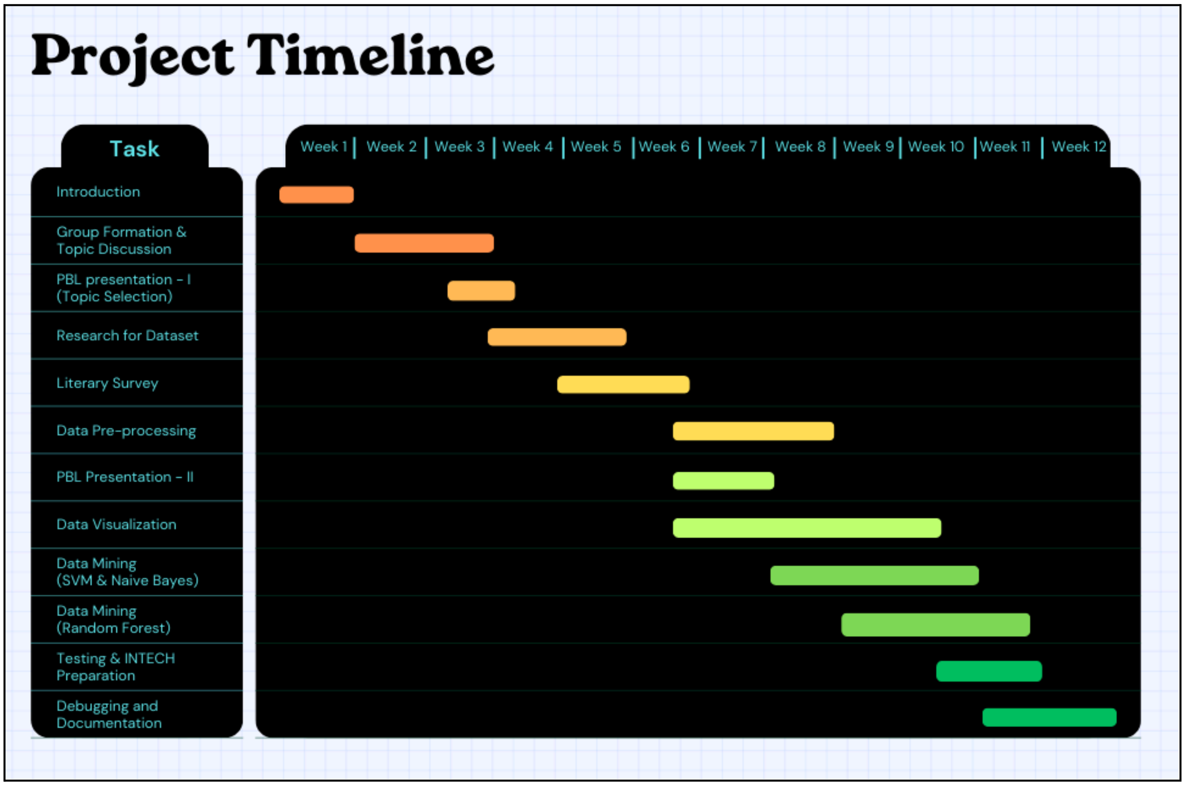

Fig.3.3.1: Timeline chart

Fig. 3.3.1 presents a timeline chart, illustrating the chronological sequence of events or milestones within a specified timeframe. This visual representation offers a clear and concise overview of key activities, tasks, or developments, facilitating easy understanding and interpretation of temporal relationships and dependencies.

3.4 Roles and Responsibilities

Data Preprocessing	Rhea Laloo
Model Development and Evaluation	Priti Prasad
Data Acquisition and Visualization	Bhumika Baria
Documentation	Yashi Depani

3.5 Software Lifecycle Model

SDLC Cycle represents the process of developing software. SDLC framework includes the following steps:

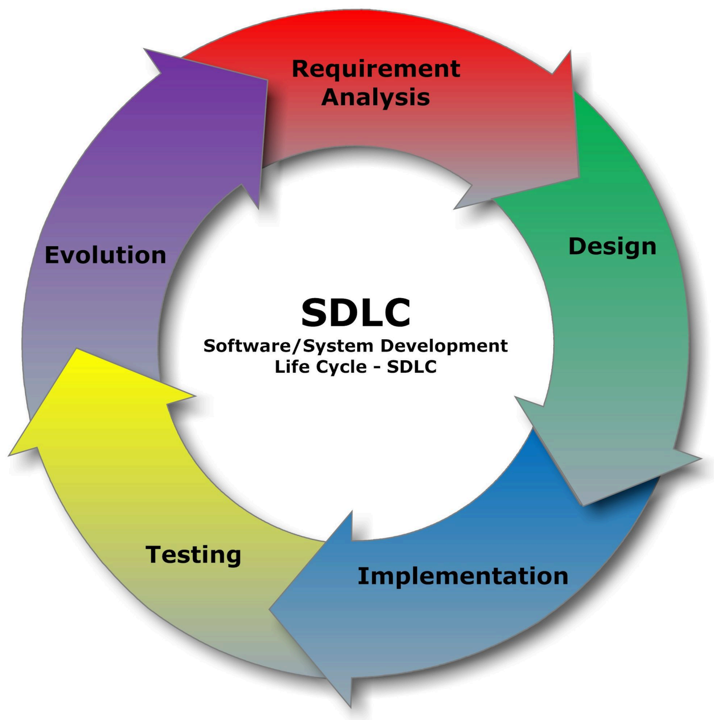

Fig. 3.5.1 Software development lifecycle model

The above fig. depicts the software lifecycle model which includes the main stages i.e Planning and requirement analysis, Designing the software, Developing the project , Testing, Deployment and Maintenance

Stage 1: Planning and requirement analysis

To understand the nature of the programs to be built, we firstly understood the information domain for the software, its functions, interfaces and performance. With the requirements given to me I decided to make the interfaces like forms to allow users to enter the information and other interfaces to view the details. The information domain is decided as the attributes of various tables.

Stage 2: Designing the software

In this process we translated the requirements into representation of the software that can be accessed for quality before we start coding. The interfaces are designed in VISUAL BASIC and are tested for quality and requirements. This phase is the product of the last two, like inputs from the customer and requirement gathering. We started creating individual templates of each webpage and structure (backend working) for our project.

Stage 3: Developing the project

The next phase was actual development where the programming is built. The implementation of design begins concerning writing code. We followed the coding guidelines described collectively earlier and programming tools like compilers, interpreters, debuggers, etc. are used to develop and implement the code.

Stage 4: Testing

The testing process focuses on the internals of the software, ensuring that all statements have been tested, and on functional externals, that is conducting tests to uncover errors and ensure that defined input will produce actual results that agree with the required results.

Stage 5: Evolution/Deployment

Once the software is certified, and no bugs or errors are stated, then it can be deployed. Whether through server installations, distribution channels, or online platforms, deployment heralds the software's availability for use by its intended audience, ensuring that it delivers on its promised capabilities while maintaining robustness and reliability in real-world scenarios.

Chapter 4: Implementation Details and Results

4.1 Technology Stack

Programming Languages: Python
Machine Learning Libraries: Scikit-learn
Data Visualization Tools: Matplotlib, Seaborn , Tableau
Data Preprocessing Tools: Pandas, Scikit-learn
Development Environment: Jupyter Notebook, Google Colab, IDEs (VSCode),

4.2 Implementation Parameters

The implementation parameters for enhancing ADHD detection through AI encompass various technical and operational considerations critical for the successful deployment and operation of the system. These parameters include:

1. **Algorithm Selection:** The choice of machine learning algorithms, including Random Forest, Support Vector Machine (SVM), and Gaussian Naive Bayes, plays a crucial role in the system's implementation. Each algorithm's suitability for ADHD detection and its computational requirements must be carefully evaluated to ensure optimal performance.
2. **Data Preprocessing:** The preprocessing techniques employed to clean and prepare the dataset significantly impact the accuracy and reliability of the machine learning models. Parameters such as handling missing values, normalizing features, and addressing class imbalances must be defined and optimized to enhance data quality and model effectiveness.
3. **Model Training and Optimization:** Parameters related to model training, such as the selection of hyperparameters and the choice of optimization algorithms, are essential for maximizing the performance of the machine learning models. Techniques such as grid search with cross-validation and hyperparameter tuning are employed to fine-tune the models and improve their accuracy and generalization capabilities.
4. **Infrastructure and Environment:** The choice of programming language, development environment, and hardware infrastructure (e.g., cloud-based platforms or on-premises servers) are crucial implementation parameters. The system is implemented using Python programming language within the Google Colab environment, leveraging the scikit-learn library for machine learning tasks.
5. **Monitoring and Maintenance:** Parameters related to system monitoring and maintenance are essential for ensuring the ongoing reliability and effectiveness of the ADHD detection system. Proactive monitoring techniques, fault detection mechanisms, and predictive maintenance strategies are implemented to minimize downtime and maximize system uptime.
6. **Evaluation Metrics:** Parameters defining the evaluation metrics used to assess the performance of the system are critical for measuring its effectiveness in ADHD detection. Metrics such as accuracy, precision, and recall are computed to gauge the system's diagnostic capabilities and inform decision-making processes.

By defining and optimizing these implementation parameters, the system can be effectively deployed and operated to enhance ADHD detection through AI, ultimately improving patient outcomes and healthcare efficiency.

4.3 Preliminary Results

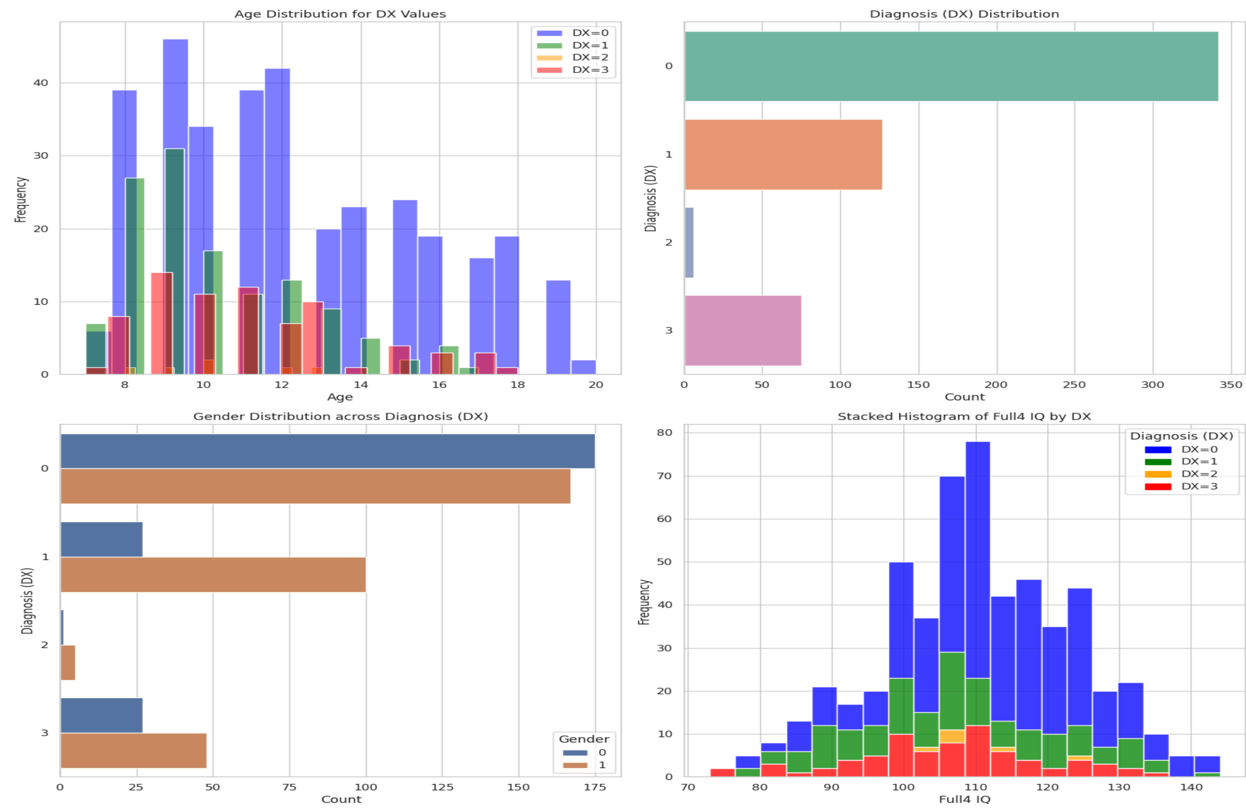

Fig 4.3.1 - Data distribution with respect to diagnosis

Fig 4.3.1 showcases the distribution of data relative to diagnosis, offering a visual representation of how different diagnostic categories are distributed within the dataset, providing insights into the prevalence or distribution of various conditions or outcomes.

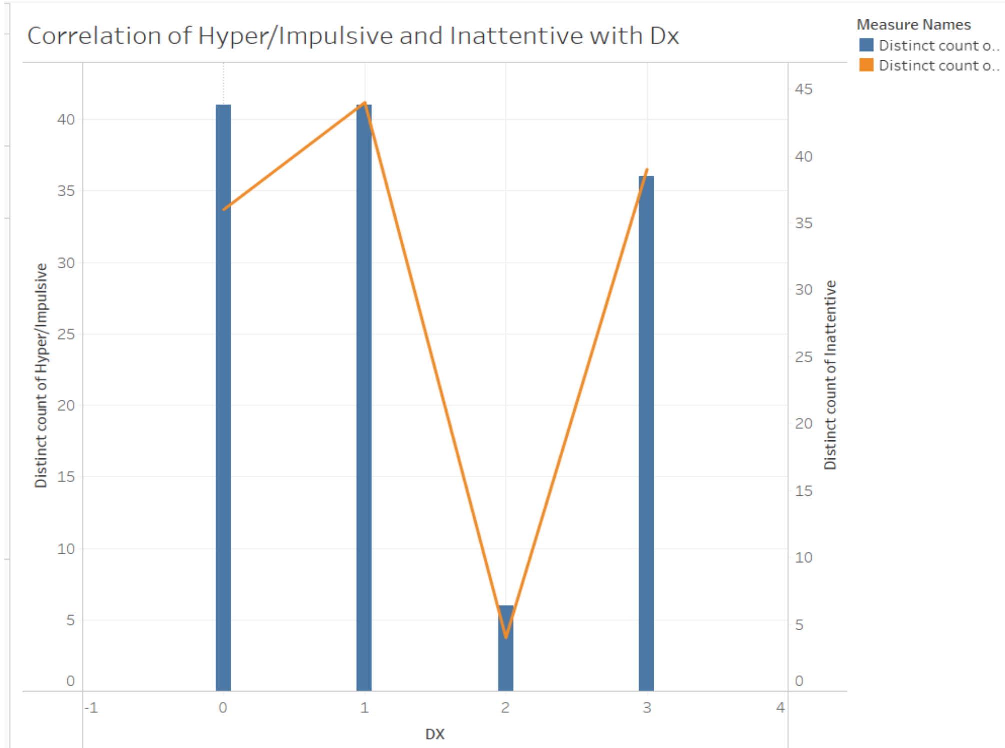

Fig 4.3.2 - Correlation between Hyper/Impulsive and Inattentive with Diagnosis

Fig 4.3.2 illustrates the correlation between hyper/impulsive and inattentive symptoms with diagnosis, providing valuable insights into the relationship between these ADHD symptom categories and the diagnosis outcome.

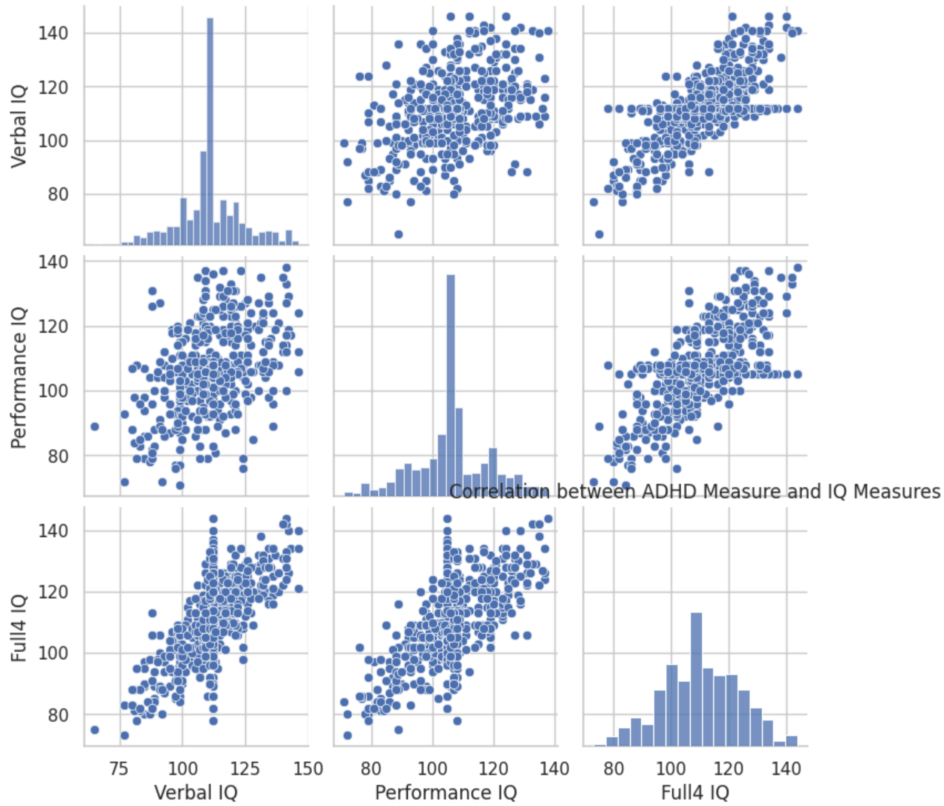

Fig 4.3.3 - correlation between ADHD measure and IQ measures

Fig 4.3.3 displays the correlation between ADHD measures and IQ measures, offering an insightful analysis of their relationship and potential associations within the context of the study or analysis.

Algorithm	Testing accuracy	Training accuracy
SVM	74.25	80
Random Forest	85.4545	87.2727
Naive bayes	57.72	60

Fig 4.3.4 - Testing and Training accuracy

Fig 4.3.4 illustrates the testing and training accuracy, providing a visual depiction of their respective trends and comparative performance over a specified period or series of iterations.

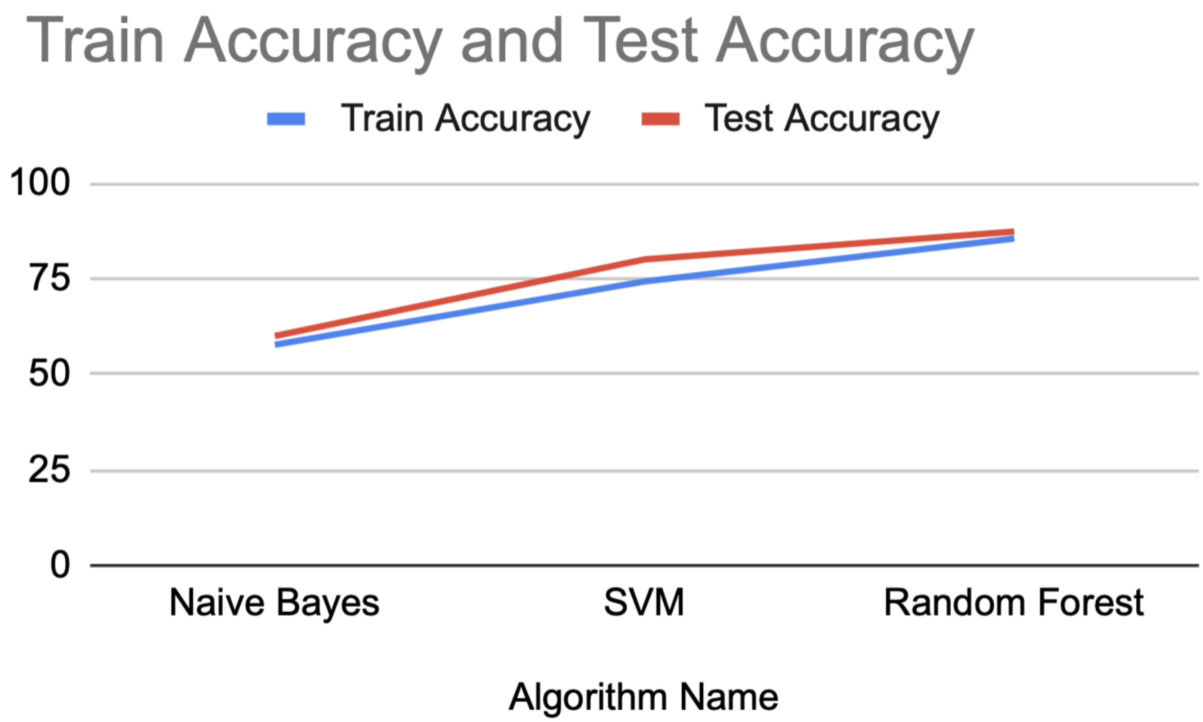

Fig 4.3.5 - Training and Testing Accuracy Line Graph

Fig 4.3.5 showcases a detailed line graph that provides a visual representation of the training and testing accuracy metrics over a specified period or set of iterations.

```
Accuracy: 0.8727272727272727
Classification Report:
              precision    recall  f1-score   support

           0       0.96      0.96      0.96        73
           1       0.81      0.81      0.81        26
           3       0.45      0.45      0.45        11

    accuracy                           0.87       110
   macro avg       0.74      0.74      0.74       110
weighted avg       0.87      0.87      0.87       110
```

Fig 4.3.6 - Random forest model performance metrics

Fig 4.3.6 presents a comprehensive overview of the performance metrics of the random forest model, offering insights into its efficacy and effectiveness across various evaluation criteria.

```
Testing Accuracy: 0.6

Classification Report for Testing Data:
              precision    recall  f1-score   support

           0       0.86      0.75      0.80        73
           1       0.58      0.42      0.49        26
           2       0.00      0.00      0.00         0
           3       0.00      0.00      0.00        11

    accuracy                           0.60       110
   macro avg       0.36      0.29      0.32       110
weighted avg       0.71      0.60      0.65       110
```

Fig 4.3.7 - Naive Bayes model performance metrics

Fig 4.3.7 presents a detailed analysis of the performance metrics pertaining to the Naive Bayes model, shedding light on its effectiveness and efficiency across diverse evaluation parameters.

```
Training Accuracy: 0.7425968109339408
 Testing Accuracy: 0.8

Classification Report for Testing Data:
              precision    recall  f1-score   support

           0       0.80      0.96      0.87        76
           1       0.79      0.62      0.70        24
           3       0.00      0.00      0.00        10

    accuracy                           0.80       110
   macro avg       0.53      0.53      0.52       110
weighted avg       0.73      0.80      0.76       110
```

Fig 4.3.8 - SVM model performance metrics

Fig 4.3.8 provides a comprehensive examination of the performance metrics associated with the SVM (Support Vector Machine) model, offering insights into its effectiveness and accuracy across a range of evaluation criteria.

Chapter 5: Conclusion

5.1 Conclusion

In conclusion, our project aimed to enhance ADHD detection through AI by evaluating multiple algorithms using a carefully curated dataset from Data World. Through meticulous preprocessing, we ensured the integrity and reliability of the data, laying a robust foundation for subsequent analysis. Our investigation focused on three widely used algorithms: Naive Bayes, Support Vector Machine (SVM), and Random Forest. Following extensive experimentation and performance evaluation, we observed that the Random Forest algorithm consistently outperformed the other two algorithms in terms of classification accuracy. Random Forest demonstrated superior capability in discerning patterns and features within the dataset, resulting in more accurate ADHD detection.

Our conclusion that Random Forest is the most suitable algorithm for ADHD detection through AI is supported by empirical evidence derived from our thorough analysis. By leveraging Random Forest, we can enhance the accuracy and reliability of ADHD diagnosis, thereby facilitating early intervention and tailored treatment strategies for individuals with ADHD. However, it is essential to acknowledge that the choice of algorithm may vary depending on specific factors such as dataset characteristics, computational resources, and application requirements. Therefore, while Random Forest emerges as the preferred choice in our study, further research and validation are warranted to explore its applicability in different contexts and settings.

Overall, our findings contribute valuable insights to the field of ADHD detection through AI, highlighting the potential of Random Forest as a powerful tool for improving diagnostic accuracy and advancing personalized healthcare for individuals with ADHD.

5.2 Future Scope

- **Multimedia Integration for Comprehensive Behavioral Analysis:** This involves leveraging images and video data to conduct a comprehensive analysis of behavioral patterns. It allows researchers and clinicians to capture visual cues such as facial expressions, body language, and environmental interactions, providing deeper insights into individuals' behavior. Multimedia integration enables a richer understanding of behavior by capturing non-verbal cues that may not be captured through traditional methods. It allows for more objective and comprehensive assessments, leading to more accurate diagnoses and tailored interventions.
- **Cross-Domain Impact:** This involves extending algorithms and sensor technologies developed for one neurodevelopmental disorder to others, facilitating cross-domain impact and accelerating progress in the field. This approach maximizes the utility of existing technologies and accelerates innovation by leveraging insights and methodologies across different disorders. It promotes interdisciplinary collaboration and knowledge sharing, leading to more efficient and impactful research efforts.
- **Mobile Feedback App:** This involves developing mobile applications for continuous monitoring of symptoms and providing real-time feedback to individuals with neurodevelopmental disorders and their caregivers. Mobile feedback apps offer a convenient and accessible way for individuals to track their symptoms and receive support in real-time. They empower individuals and caregivers to actively manage symptoms and make informed decisions about treatment and intervention strategies. Additionally, the data collected through these apps can contribute to research efforts, providing valuable insights into the natural progression of neurodevelopmental disorders and the effectiveness of interventions.

References

[1]Nash, Christian, Rajesh Nair, and Syed Mohsen Naqvi. 2023a. “Machine Learning in ADHD and Depression Mental Health Diagnosis: A survey.” IEEE Access 11 (January): 86297–317. https://doi.org/10.1109/access.2023.3304236.

[2]Nouri, Ali, and Zahra Tabanfar. 2023. “Detection of ADHD Disorder in Children Using Layer-Wise Relevance Propagation and Convolutional Neural Network: An EEG Analysis.” Frontiers in Biomedical Technologies (Print), December. https://doi.org/10.18502/fbt.v11i1.14507.

[3]Chen Y, Tang Y, Wang C, Liu X, Zhao L, Wang Z. ADHD classification by dual subspace learning using resting-state functional connectivity. Artif Intell Med. 2020 Mar;103:101786. doi: 10.1016/j.artmed.2019.101786. Epub 2020 Jan 13. PMID: 32143793.

[4]J. Shin, M. Maniruzzaman, Y. Uchida, M. A. M. Hasan, A. Megumi and A. Yasumura, "Handwriting-Based ADHD Detection for Children Having ASD Using Machine Learning Approaches,"in IEEE Access, vol. 11, pp. 84974-84984, 2023, doi: 10.1109/ACCESS.2023.3302903.

[5]M. Maniruzzaman, M. A. M. Hasan, N. Asai and J. Shin, "Optimal Channels and Features Selection Based ADHD Detection From EEG Signal Using Statistical and Machine Learning Techniques," in IEEE Access, vol. 11, pp. 33570-33583, 2023, doi: 10.1109/ACCESS.2023.3264266.

[6]Ahire, Nitin, Raval N. Awale, and Abhay Wagh. "Comprehensive review of EEG data classification techniques for ADHD detection using machine learning and deep learning." Romanian Journal of Pediatrics/Revista Română de Pediatrie 72, no. 2 (2023).