Advancing cerebral palsy research: an in-depth exploration of machine learning approaches (#101603)

First submission

Guidance from your Editor

Please submit by **6 Jul 2024** for the benefit of the authors (and your token reward) .

Literature review article
Please read the 'Structure and Criteria' page to understand the reviewing criteria for this Literature review article.

Custom checks
Make sure you include the custom checks shown below, in your review.

Author notes
Have you read the author notes on the guidance page?

Image check
Check that figures and images have not been inappropriately manipulated.

If this article is published your review will be made public. You can choose whether to sign your review. If uploading a PDF please remove any identifiable information (if you want to remain anonymous).

Files

Download and review all files from the materials page.

9 Figure file(s)
6 Table file(s)

Custom checks

Systematic review or meta analysis

- Have you checked our policies?
- Is the topic of the study relevant and meaningful?
- Are the results robust and believable?

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **STUDY DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Is the review of broad and cross-disciplinary interest and within the scope of the journal?
- Has field been reviewed recently. It there a good reason for this review (different viewpoint, audience etc.)?
- Introduction adequately introduces the subject and makes audience and motivation clear.

STUDY DESIGN

- Article content is within the Aims and Scope of the journal.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.
- Is the Survey Methodology consistent with a comprehensive, unbiased coverage of the subject? If not, what is missing?
- Are sources adequately cited? Quoted or paraphrased as appropriate?
- Is the review organized logically into coherent paragraphs/subsections?

VALIDITY OF THE FINDINGS

- **Impact and novelty is not assessed.** Meaningful replication encouraged where rationale & benefit to literature is clearly stated.
- Conclusions are well stated, linked to original research question & limited to supporting results.
- Is there a well developed and supported argument that meets the goals set out in the Introduction?
- Does the Conclusion identify unresolved questions / gaps / future directions?

Standout reviewing tips

The best reviewers use these techniques

Tip	Example
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Advancing cerebral palsy research: an in-depth exploration of machine learning approaches

Anjuman Nahar [1], **Sudip Paul** [2], **Manob Jyoti Saikia** [Corresp. 3]

[1] Department of Biomedical Engineering, North Eastern Hill University, Shillong, Meghalaya, India
[2] Department of Biomedical Engineering, North Eastern Hill University, Shilong, Meghalaya, India
[3] Department of Electrical Engineering, University of North Florida, Jacksonville, FL, United States

Corresponding Author: Manob Jyoti Saikia
Email address: manob.saikia@unf.edu

Background. This narrative review aims to explore advances in the field of cerebral palsy (CP) focusing on Machine Learning (ML) models. The study analyzed ML models used in CP research to find out their computing efficiency and compare how well various algorithms perform. The results of this study offers new perspectives on ML methods for CP identification, classification, prediction and management. We found that ML models are primarily used to predict risk factors and symptoms that may lead to CP, followed by screening CP cases, classification, and diagnosis. **Methodology.** A total of 31 studies were identified on ML for CP from 2013 to 2023. The ML models used for prediction are MLP, RF, CNN, DT, and Ensemble. **Results.** RF is mainly used for classifying movements and deformities due to CP. SVM, DT, RF, and KNN show 100% accuracy in exercise evaluation. RF and DT show 94% accuracy in the classification of Gait patterns, Multilayer Perceptron (MLP) shows 84% accuracy in the classification of CP children, Bayesian causal forests (BCF) has 74% accuracy in predicting the average treatment effect on various orthopedic and neurological conditions. Neural networks are 94.17% accurate in diagnosing CP using eye images. **Conclusion.** Clinical data are primarily used in ML models in the CP field, accounting for almost 47%. With the rise in popularity of machine learning techniques, there has been a rise in interest in developing automated and data-driven approaches to explore the use of ML in CP.

Advancing cerebral palsy research: an in-depth exploration of machine learning approaches

Anjuman Nahar[1], Sudip Paul[2], Manob Jyoti Saikia[3]

[1]*Department of Biomedical Engineering, North-Eastern Hill University, Shillong, Meghalaya 793022, India.*

[2]*Department of Biomedical Engineering, North-Eastern Hill University, Shillong, Meghalaya 793022, India.*

[3]*Department of Electrical Engineering, University of North Florida, Jacksonville, FL 32224, USA.*

Corresponding Author:
Manob Jyoti Saika[3]

[3]*Department of Electrical Engineering, University of North Florida, Jacksonville, FL 32224, USA.*

Email address: manob.saikia@unf.edu

Abstract

Background. This narrative review aims to explore advances in the field of cerebral palsy (CP) focusing on Machine Learning (ML) models. The study analyzed ML models used in CP research to find out their computing efficiency and compare how well various algorithms perform. The results of this study offers new perspectives on ML methods for CP identification, classification, prediction and management. We found that ML models are primarily used to predict risk factors and symptoms that may lead to CP, followed by screening CP cases, classification, and diagnosis.

Methodology. A total of 31 studies were identified on ML for CP from 2013 to 2023. The ML models used for prediction are MLP, RF, CNN, DT, and Ensemble.

Results. RF is mainly used for classifying movements and deformities due to CP. SVM, DT, RF, and KNN show 100% accuracy in exercise evaluation. RF and DT show 94% accuracy in the classification of Gait patterns, Multilayer Perceptron (MLP) shows 84% accuracy in the classification of CP children, Bayesian causal forests (BCF) has 74% accuracy in predicting the average treatment effect on various orthopedic and neurological conditions. Neural networks are 94.17% accurate in diagnosing CP using eye images.

Conclusion. Clinical data are primarily used in ML models in the CP field, accounting for almost 47%. With the rise in popularity of machine learning techniques, there has been a rise in interest in developing automated and data-driven approaches to explore the use of ML in CP.

Introduction

Cerebral palsy (CP) is characterized by a diverse range of mobility and posture abnormalities that are permanent but not irreversible due to injury to the developing brain. Individuals may struggle with communication, behavior, vision, hearing, nutrition, pain, and sleep in addition to their issues. In industrialized nations, CP is thought to affect

1.4 to 1.8 out of every 1000 live births, compared to 2.95 to 3.4 out of every 1000 live births in low- and middle-income countries. Lifetime effects of CP include decreased independence in daily living activities, play, and involvement in educational, social, and community activities [1].

The surveillance of cerebral palsy in Europe (SCPE) has given a standardized CP classification dividing them into three major groups: spastic (unilateral or bilateral spastic), dyskinetic (dystonic or choreoathetosis), and ataxic [2]. The most frequent impairments in children with cerebral palsy are motor impairments, speech impairments, pain, intellectual impairments, sensory abnormalities, epilepsy, and behavioral issues. The most significant problem is motor impairments primarily caused by spasticity. Aberrant motor functions cause altered movement and posture. In addition to equinus deformity and hand dysfunction, it might result in hip discomfort or dislocation [3, 4]. Strength, balance, coordination, sensory processing, and selective motor control are common difficulties for kids with cerebral palsy. Additionally, unlike children who are usually growing, they cannot learn motor patterns [2, 3].

CP signs and symptoms typically appear in the early periods of infancy, yet it takes an average of two years for CP to be diagnosed. Because infants have a higher chance of recovering from brain damage than adults, early detection and treatment are essential for people with CP. Identifying high-risk neonates, tracking neurodevelopment, and predicting CP can all be aided by neuroimaging, motor evaluation, and neurological exams. The infant brain's structural alterations can be detected using neuroimaging techniques, including magnetic resonance imaging (MRI) and cranial ultrasonography. These techniques can also be used to track lesions' progression and evaluate treatment benefits [2]. Although CP cannot be cured, medication, surgery, and other interventions like physical, occupational, speech, and behavioral therapy lead to a significant functional outcome to make the child functionally independent [4].

However, only qualified medical professionals can carry out such an evaluation. General movement evaluations carried out by medical professionals based on visual observation are frequently influenced by observer fatigue and subjective impressions. It is necessary to develop a systematic model to deliver accurate and quick prediction outcomes and provide accurate personalized care [4].

In recent years, machine learning has become a robust tool with enormous potential in the healthcare industry. By utilizing machine learning techniques, it is possible to predict their specific treatment outcomes. With machine learning algorithms, large-scale data analysis, the extraction of significant patterns, and the development of individualized models that forecast the success of a certain intervention for a given patient are all possible [5]. Random Trees, Support Vector Machines, Multilayer perception, artificial neural networks, direct matching, virtual twins, and Bayesian causal forests are some of the ML models that have been increasingly applied to the field of CP [6-8].

Despite the advancements in therapeutic approaches and medical technology, diagnosing, classifying, and managing cerebral palsy (CP) still presents difficulties. Conventional approaches are frequently arbitrary, laborious, and dependent on the knowledge of medical specialists. The potential of machine learning to handle enormous datasets with efficiency, objectivity, and speed can result in more prompt and accurate diagnosis as well as individualized treatment strategies.

A thorough analysis that summarizes the state of machine learning applications in CP research is desperately needed. This review fills that gap by methodically going over the body of literature, pinpointing the best machine learning models, and emphasizing areas in need of more investigation. This study attempts to improve future research and therapeutic practices by offering a comprehensive evaluation of current research, thereby improving the lives of people with CP.

The target audience for this narrative review is broad and includes a range of stakeholders in the cerebral palsy (CP) area. Understanding how machine learning (ML) models improve diagnosis accuracy and optimise treatment plans can be helpful for medical professionals who identify and treat children with cerebral palsy (CP). The results of this study can be used by clinical researchers to direct their research into the causes, development, and management of cerebral palsy. Data scientists may create and improve prediction models with the help of this insightful knowledge on using machine learning (ML) algorithms in healthcare, particularly in clinical practice (CP). Although the review is mainly technical, patients and carers can benefit from summaries and implications of the findings to better grasp the role of developing technology in enhancing quality of life and possible future breakthroughs in CP care.

In this study, we aim to investigate the use of machine learning in cerebral palsy. Traditional CP diagnosis and evaluation techniques mainly rely on subjective and time-consuming clinical observations and neurological tests. Machine learning techniques have created new opportunities for improving the precision and effectiveness of CP diagnosis and prognosis. Hence we provide a thorough assessment of recent research on the use of machine learning in cerebral palsy in this paper, covering a variety of applications from early diagnosis and classification to individualized treatment plans. We explore the difficulties encountered and the possible advantages that machine learning algorithms offer in enhancing the lives of people with CP. The potential for machine learning approaches to analyze complicated medical data to help with diagnosis and treatment planning has been established. This study compares the effectiveness of various machine learning (ML) methods in addressing the difficulties associated with CP diagnosis and prognosis.

Survey Methodology

An in-depth review was conducted to understand better the use of machine learning models in children with cerebral palsy. Search Engines used during the review included electronic databases like Pubmed for accessing biomedical and life sciences, IEEE Xplore for technical literature in computer, Google Scholar for a broad range of academic publications, Scopus and web of science for multidisciplinary high impact journals. Inclusion criteria included articles containing keywords such as Cerebral palsy, machine learning approaches, outcome response, identification, classification, diagnosis, and treatment prediction. We selected Full-text papers, Clinical trials, RCT, Systematic reviews, narrative reviews, and meta-analyses published in English. Studies were included if they reported the application of ML techniques for CP patients. Peer reviewed articles from 2013 to 2023 were only included for the review. Each search was constructed using the Boolean operators AND and OR to optimize search criteria. Exclusion criteria for the review included studies not directly related to CP. Editorials, opinion pieces, and non-peer-reviewed articles were also excluded.

From the combination of database searches, 25 papers were identified. EndNote 20 was used to collate all relevant papers and remove any duplicates. After reduplication, 20 papers met the requirements for screening, and were used for analysis. To be included in the review, each paper was screened for title and abstract followed by full-text screening for identifying the inclusion and exclusion criteria.

By adhering to this structured approach, the review ensures a comprehensive and unbiased coverage of the literature, capturing the current state of research and advancements in the application of machine learning to cerebral palsy.

Results

After selecting the articles matching the needs of this study, we obtained 24 articles that discussed cerebral palsy with machine learning. Table 1 is the result of the articles discussed in this study. Of the 24 research articles included, several reports show the classification model used in CP, such as the Random Forest (RF) model used in eleven research studies, Decision tree (DT) algorithms used in five research studies, Support Vector Machine (SVM) used two studies, Logistic Regression (LR) was used in four studies, K-Nearest Neighbour (KNN) was used 3 studies in the research and so on. One Ensemble model is also used in this study to predict treatment in hand function in cerebral palsy patients. Research articles that have been collected are grouped based on the model used in the study, which is presented in Table 2. Different performance evaluation metrics were used across the studies to assess the effectiveness of the machine learning models, and commonly reported metrics included accuracy, sensitivity, specificity, area under the curve (AUC), precision, and recall. These metrics provided insights into the models' predictive accuracy and discriminative ability. Several studies identified significant predictive factors and assessed their importance in treatment outcome prediction for unilateral cerebral palsy. Commonly observed predictors included clinical assessments, demographic characteristics, neuroimaging data, genetic markers, etc. The importance of these factors varied across studies, with some features showing a more significant influence on treatment response prediction than others. A comparison of the machine learning approaches revealed variations in predictive performance and feature selection. While specific models demonstrated high accuracy and robustness, others exhibited limitations in terms of generalizability and interpretability. Despite the advancements in machine learning models for treatment prediction in unilateral cerebral palsy, several research gaps and limitations were identified. These included the lack of standardized outcome measures, limited external validation of models, small sample sizes, and challenges associated with model interpretability.

Additionally, there was a need for more longitudinal studies to explore the dynamic nature of treatment response and the potential impact of time-dependent predictors. Based on the reviewed studies, several emerging trends and future directions were identified. These included integrating multimodal data sources, developing explainable AI models, exploring causal inference approaches, and needing large-scale collaborative studies to validate and refine the existing machine learning models. Machine learning and computational intelligence have been applied to various medical challenges to aid medical practitioners in making decisions. This is particularly significant and intriguing in the developing field of personalized medicine, which is frequently defined as providing "the right patient with the right drug at the right dose at the right time" and tailoring medical care to unique patient characteristics, needs, and preferences. Machine learning models used in cerebral palsy for identifying risk factors, classification of patterns in CP children, prediction of treatment outcome, and diagnosis of CP are briefed in Table 2. Table 3 briefly provides a list of the type of data used for corresponding ML models in the field of CP. Figure 1 provides an estimate of the data used in ML for CP research, and Fig 2 provides the estimates of ML models and their corresponding data.

While evaluating the performance of used machine learning techniques in cerebral palsy in Table 2, we have compared the accuracy measures of used algorithms to select the best one for future use. Machine learning in cerebral palsy has been used to identify the risk factors, predict the disease, diagnosis, and treatment response. Figure 3 explains that ML models in CP are mainly used for prediction.

Identification and diagnosis using ML Models

The primary diagnostic method for CP identification still relies on traditional clinical assessment components like delayed motor milestones, asymmetry of movement, or abnormal muscle tone, scales like The General Movements Assessment, The Hammersmith Infant Neurological Examination, and neurological data like MRI. However, these approaches are arbitrary, time-consuming, and expensive. However, as machine learning techniques have become more prevalent, there has been an increase in interest in creating automated and data-driven methods to recognize and diagnose CP, improving accuracy and effectiveness [24]. A list of some of the ML models used for identifying symptoms and diagnosing CP is given in Table 4.

Table 3 explains the identification pattern in cerebral palsy cases by ML models. Most of the ML in this section identifies physical activities in CP children, followed by identifying CP cases and disease conditions in children that might lead to CP. Figure 4 shows that the Random forest model is the most accurate ML model for identifying movements and disease-causing factors.

Classification using ML Models

It is essential to correctly classify CP subtypes to comprehend underlying mechanisms, forecast outcomes, and create individualized therapies. Automatically identifying CP subtypes based on various data sources using machine learning algorithms has shown remarkable results, enabling more accurate diagnosis and individualized treatment plans [26]. Table 5 presents some of the studies done in the classification field in cerebral palsy children.

Table 5 explains the list of ML models used to classify various components in CP. ML models in CP classify humam gait patterns and conditions that lead to the disease. In Figure 5 we can see that Decision trees and Random forests are accurate classifying models in cases of CP.

Predictions using ML Models

By helping healthcare professionals foresee probable complications, anticipate the course of the condition, and develop individualized treatment plans, prediction plays a crucial role in managing cerebral palsy (CP). In CP, machine learning has become a potent tool for predictive modeling, utilizing vast amounts of data and advanced algorithms to produce precise prognostic insights. This section focuses on how machine learning methods have been used to forecast various CP-related outcomes, from treatment outcomes and disease progression to functional products and adaptive solutions. Traditionally the prediction of the most accurate risk factors of CP is the observational general movement assessment (GMA) and cerebral imaging. However, they either rely on qualitative perception, necessitating extensive training and clinical expertise (GMA), or call for costly tools. Machine learning models in such conditions remain cost-effective and accurate prediction tools [24]. Due to the neuromuscular involvement in hemiplegic cerebral palsy, which surface EMG determines, it is strongly encouraged to assess the recruitment of muscles using myoelectric-signal analysis. In this case, additional, expensive, and complex attributes are needed to identify the gait components. In these circumstances, the machine learning model significantly anticipates the locomotion process [28]. Table 6 lists the ML models used for predicting disease and its features, followed by treatment responses in CP. Figure 6 explains that prediction models indicate the condition, motor problems due to the disease, moments, and treatment responses. The ML models used for disease prediction are RF and LR; for movement prediction, GBR and CNN; for motor problems, predictions are MLP and RF; for treatment prediction, are Ensemble, BR, and DT models

Machine learning is a field that teaches computers to handle data efficiently. Sometimes, evaluating information can be challenging, even after thoroughly examining the data. This is where machine learning comes into play. The demand for machine learning has risen due to the availability of numerous datasets. To address various data problems, machine learning relies on different algorithms. The specific algorithm depends on the nature of the issue, including factors such as the type of problem and the number of variables involved. Figure 7 demonstrates the kinds of models that are used in machine learning.

The machine learning endeavor of inferring a function from labeled training data is known as supervised learning. A collection of training examples makes up the training data. Supervised machine learning algorithms require external assistance for their operation. The input dataset is bifurcated into training and testing datasets. Within the training dataset, an output variable serves as the target for prediction or classification purposes.

Unsupervised learning is how a computer learns to infer a function representing a hidden structure from "unlabeled" data. Unsupervised learning differs from supervised learning and reinforcement learning in that the examples given to the learner are not labeled. Therefore the accuracy of the structure produced by the pertinent algorithm cannot be evaluated. Reinforcement learning is when a computer program interacts with a lively environment to accomplish a specific task. As the program moves through its issues, feedback in the form of rewards and penalties is given [32]. The workflow of supervised machine learning, unsupervised, and reinforced ML models is described in Figure 8**.**

Some well-known supervised machine learning algorithms are SVM, DT, and Naïve Bias. Support Vector Machines (SVM) are flexible; they can be used for Regression (prediction) as well as classifying data (classification). Drawing a line to categorize data helps divide items into distinct categories. As it learns from examples during training, SVM's task is to become highly adept at drawing these lines. Decision Tree is a valuable tool that may be used for value prediction or grouping data into categories. Like a flowchart, the data is divided into smaller groups depending on specific characteristics. The chart's endpoints represent conclusions, while the numerous spots along the route represent the division of the data.

While a Regression Tree works with forecasting data, a Classification Tree is like making decisions with a Yes or No response. Another model of supervised learning is Naïve Bias. Simple conditional probability is used in this approach. It uses training data to modify a probability table that serves as its model. This "probability table" was created based on the data features' qualities. This table can be used to determine the probability associated with various classes when making a new prediction. Unsupervised machine learning is mainly used for clustering. K-means stands out as a simple unsupervised learning method created to address the well-known clustering problem. Using an easy-to-understand methodology, the procedure divides a given dataset into a predetermined number of clusters. The main idea is to determine 'k' centers, each corresponding to a group. The location of these centers is crucial because it considerably affects the results. It is advisable to place them as far apart as possible to get the best effects. A neural network is like a chain of steps that tries to uncover hidden connections within a dataset by imitating how our brain works. In unsupervised learning, the web lacks hints about the expected outcomes. Its primary task is to sort data based on similarities. The network examines how different inputs relate to each other and groups them accordingly [33, 34]. Fig 9 represents some of the models.

Discussion

A promising development in the medical industry is the application of machine learning to the recognition, diagnosis, and prediction of cerebral palsy. This report analyzed numerous studies and research initiatives that showed how machine learning algorithms might improve the accuracy and effectiveness of CP management. The following discussion illustrates the difficulties and potential possibilities in this developing field while summarising the main findings, discussing the implications, and highlighting the challenges.

Random forests have shown excellent results for precision in CP identification and diagnosis. Machine learning algorithms have also demonstrated the ability to predict motor development delays and identify newborns more likely to acquire cerebral palsy. Thanks to these predictive models, the ability to intervene early facilitates early intervention therapy and lessens the condition's potential effects. These models offer an objective and data-driven approach to subtype classification, reducing the subjectivity associated with traditional clinical assessments. However, successfully implementing these models in clinical settings requires rigorous validation and the establishment of clear guidelines for their interpretation and integration [12-20]. By enabling personalized therapeutic recommendations, machine learning has the potential to revolutionize cerebral palsy treatment planning. Predictive models can suggest the best treatments for each patient by examining their unique patient data, such as clinical history, neuroimaging findings, and treatment responses. This individualized strategy optimizes the use of healthcare resources while simultaneously improving treatment outcomes. However, when implementing personalized therapy recommendations, ethical issues, including data protection and patient permission, are essential [8][10][11][13][22][23][29][30].

As mentioned in the literature, the challenges and limitations associated with using machine learning (ML) in the context of CP mainly revolve around the need for high-quality datasets for model training. ML algorithms heavily rely on reliable datasets. However, gathering well-curated datasets for CP can be challenging. The availability of representative data, such as records, imaging data, and longitudinal patient information, is often limited. This limitation can potentially impact the performance and generalizability of ML models in CP.

Furthermore, if the training data is biased and not representative of the CP population, it may lead to predictions from the model for specific demographic groups. Achieving fairness and reducing bias in ML models concerning CP is a task. Another issue with ML models in CP is their ability to generalize beyond the training data without overfitting. Overfitting occurs when ML models become overly specialized to the training data and perform poorly when faced with unknown data. Striking a balance between model complexity and generalization is particularly challenging when working with CP datasets. To mitigate overfitting, regularization techniques, and comprehensive model evaluation are necessary. While machine learning has made enormous strides in improving the diagnosis and prognosis of cerebral palsy, specific issues still need to be resolved.

The generalizability of models can be impacted as we are limited by data quality, sample size, and accessibility to various datasets. Deploying complicated ML and deep learning models in clinical practice may be constrained by overfitting and interpretability problems. To acquire the trust of clinicians and simplify their inclusion into healthcare systems, future research should concentrate on creating more interpretable and explicable models.

Conclusions

In conclusion, detecting, diagnosing, and predicting cerebral palsy using machine learning constitutes a significant achievement in the medical industry. These predictive models have shown promise in improving accuracy, enabling early intervention, and improving treatment planning. However, cooperation between physicians, researchers, and data scientists is essential to entirely using machine learning in CP management. To successfully integrate predictive models into routine clinical practice and ultimately improve outcomes and quality of life for people with cerebral palsy, addressing ethical issues, reducing bias, and carrying out thorough validation studies are essential.

Acknowledgements

The authors acknowledge the support from the Department of Biomedical Engineering, NEHU and Department of Electrical Engineering, University of North Florida.

References

1. Jackman, M., et al., *Interventions to improve physical function for children and young people with cerebral palsy: international clinical practice guideline.* Developmental Medicine & Child Neurology, 2022. **64**(5): p. 536-549.
2. Sadowska, M., B. Sarecka-Hujar, and I. Kopyta, *Cerebral palsy: current opinions on definition, epidemiology, risk factors, classification and treatment options.* Neuropsychiatric disease and treatment, 2020: p. 1505-1518.
3. Graham, D., S.P. Paget, and N. Wimalasundera, *Current thinking in the health care management of children with cerebral palsy.* Medical Journal of Australia, 2019. **210**(3): p. 129-135.
4. Zhang, J., *Multivariate Analysis and Machine Learning in Cerebral Palsy Research.* Front Neurol, 2017. **8**: p. 715.
5. Song, C., et al., *Application and research progress of machine learning in the diagnosis and treatment of neurodevelopmental disorders in children.* Frontiers in Psychiatry, 2022. **13**: p. 960672.
6. Zhang, J., *Multivariate analysis and machine learning in cerebral palsy research.* Frontiers in neurology, 2017. **8**: p. 715.
7. Schwartz, M.H., A.J. Ries, and A.G. Georgiadis, *Estimating the Efficacy of Common Treatments in Children and Young Adults Diagnosed with Cerebral Palsy Using Three Machine Learning Algorithms.* medRxiv, 2021: p. 2021.10. 06.21264624.
8. Schwartz, M.H., A.J. Ries, and A.G. Georgiadis, *Short-term causal effects of common treatments in ambulatory children and young adults with cerebral palsy: three machine learning estimates.* Scientific reports, 2022. **12**(1): p. 7818.
9. Slijepcevic, D., et al., *Explainable Machine Learning in Human Gait Analysis: A Study on Children With Cerebral Palsy.* IEEE Access, 2023.
10. Bertoncelli, C.M., et al., *Identifying Postural Instability in Children with Cerebral Palsy Using a Predictive Model: A Longitudinal Multicenter Study.* Diagnostics, 2023. **13**(12): p. 2126.
11. Ozates, M.E., et al., *Machine learning-based prediction of joint moments based on kinematics in patients with cerebral palsy.* Journal of Biomechanics, 2023. **155**: p. 111668.
12. von Elling-Tammen, L., et al., *Predicting Gross Motor Function in Children and Adolescents with Cerebral Palsy Applying Artificial Intelligence Using Data on Assistive Devices.* Journal of Clinical Medicine, 2023. **12**(6): p. 2228.
13. Krechowicz A, Deniziak S, Kaczmarski D. Machine learning approach to gait deviation prediction based on isokinetic data acquired from biometric sensors. Gait & Posture. 2023 Mar 1;101:55-9.14.
14. Al-Sowi, A.M., et al., *Cerebral Palsy classification based on multi-feature Analysis using machine learning.* Informatics in Medicine Unlocked, 2023. **37**: p. 101197.
15. Hua, A., et al., *Evaluation of machine learning models for classifying upper extremity exercises using inertial measurement unit-based kinematic data.* IEEE Journal of Biomedical and Health Informatics, 2020. **24**(9): p. 2452-2460.
16. Palkowski, A., et al. *Basic evaluation of limb exercises based on electromyography and classification methods*. in *2018 International Interdisciplinary PhD Workshop (IIPhDW)*. 2018. IEEE.
17. Kim, J.-Y., et al., *Analysis of machine learning-based assessment for elbow spasticity using inertial sensors.* Sensors, 2020. **20**(6): p. 1622.
18. Illavarason, P., J. Arokia Renjit, and P. Mohan Kumar, *Medical Diagnosis of Cerebral Palsy Rehabilitation Using Eye Images in Machine Learning Techniques.* J Med Syst, 2019. **43**(8): p. 278.
19. Goodlich, B. and M. BExSc, *Machine learning algorithms for the automatic detection and classification of physical activity in children with cerebral palsy who use mobility aids for ambulation.* 2019.
20. Ahmadi, M., et al., *Machine learning algorithms for activity recognition in ambulant children and adolescents with cerebral palsy.* Journal of neuroengineering and rehabilitation, 2018. **15**(1): p. 1-9.
21. Hussain, A.N., et al., *Machine Learning Approach Based on Automated Classification for Leg Rehabilitation.*

22. Parmar, P. and B.T. Morris. *Measuring the quality of exercises*. in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2016. IEEE.
23. van den Boorn, H., *Using a machine learning approach to predict the outcome of intervention programs in children with unilateral cerebral palsy.* 2016.
24. Ihlen, E.A., et al., *Machine learning of infant spontaneous movements for the early prediction of cerebral palsy: A multi-site cohort study.* Journal of clinical medicine, 2019. **9**(1): p. 5.
25. Turova, V., et al., *Machine learning models for identifying preterm infants at risk of cerebral hemorrhage.* PloS one, 2020. **15**(1): p. e0227419.
26. Zhang, Y. and Y. Ma, *Application of supervised machine learning algorithms in the classification of sagittal gait patterns of cerebral palsy children with spastic diplegia.* Computers in biology and medicine, 2019. **106**: p. 33-39.
27. Abbas, R., et al. *Classification of foetal distress and hypoxia using machine learning approaches*. in *Intelligent Computing Methodologies: 14th International Conference, ICIC 2018, Wuhan, China, August 15-18, 2018, Proceedings, Part III 14*. 2018. Springer.
28. Morbidoni, C., et al., *Machine-learning-based prediction of gait events from EMG in cerebral palsy children.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2021. **29**: p. 819-830.
29. Afifi, J., *Prediction of Cerebral Palsy in Very Preterm Infants*. 2021.
30. Krechowicz, A., S. Deniziak, and D. Kaczmarski, *Machine learning approach to gait deviation prediction based on isokinetic data acquired from biometric sensors.* Gait & Posture, 2023. **101**: p. 55-59.
31. Oliveira, L.B.T.G.d., *Using Machine Learning to Predict Mobility Improvement of Patients after Therapy: A Case Study on Rare Diseases*. 2023.
32. Sharma, D., & Kumar, N. (2017). A review on machine learning algorithms, tasks, and applications. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, *6*(10), 2278-1323.
33. Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]*, *9*(1), 381-386.
34. Ray, S. (2019, February). A quick review of machine learning algorithms. In *2019 International Conference on machine learning, big data, cloud and parallel computing (COMITCon)* (pp. 35-39). IEEE.

Figure 1

Machine learning models used for classification in cerebral palsy cases

Fig. 5: Machine learning models used for classification in cerebral palsy cases

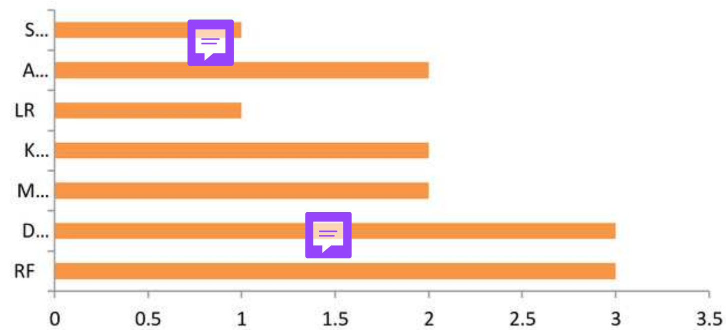

Figure 2

Some machine learning models

Fig 9: Some machine learning models

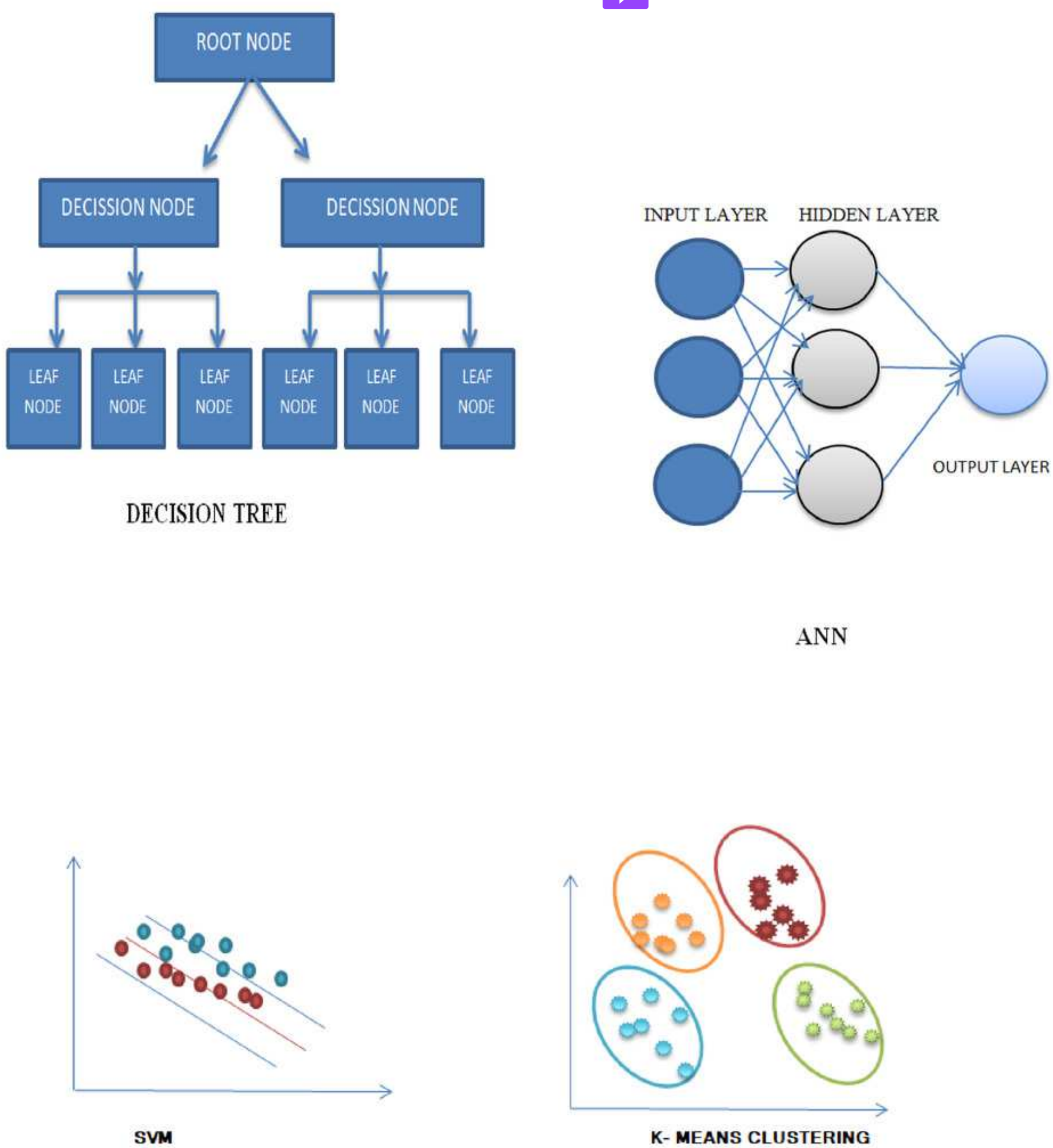

Figure 3

The types of models that are used in machine learning

Fig 7: The types of models that are used in machine learning

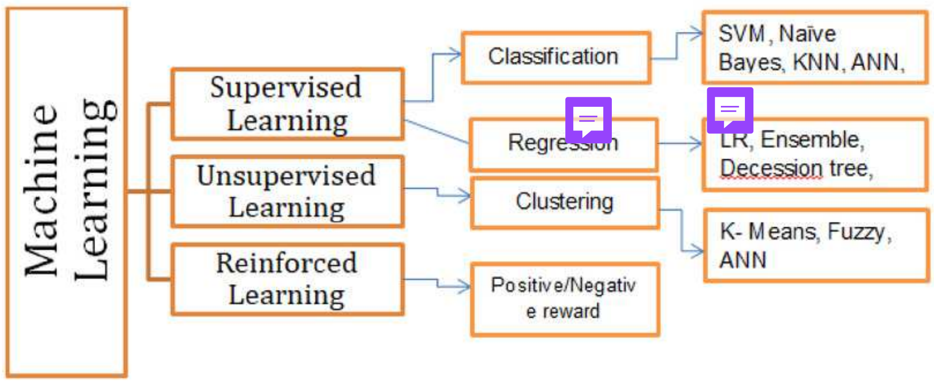

Figure 4

Use of ML models for prediction in cerebral palsy cases

Fig 6: Use of ML models for prediction in cerebral palsy cases

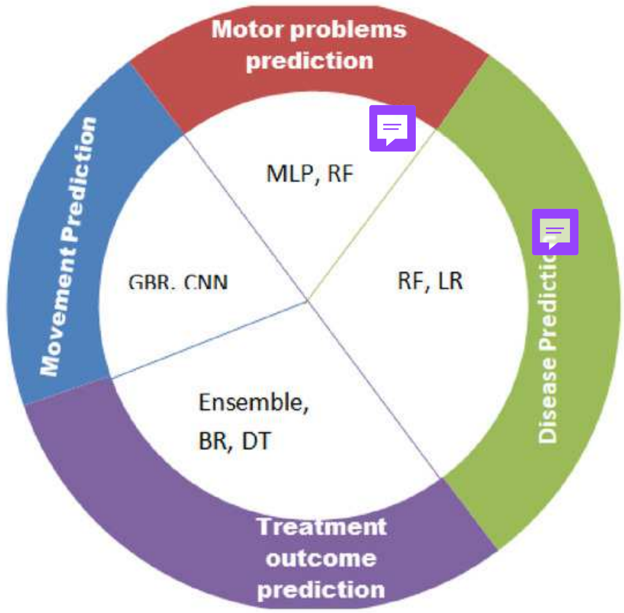

Figure 5

workflow of machine learning models

Fig 8: workflow of machine learning models

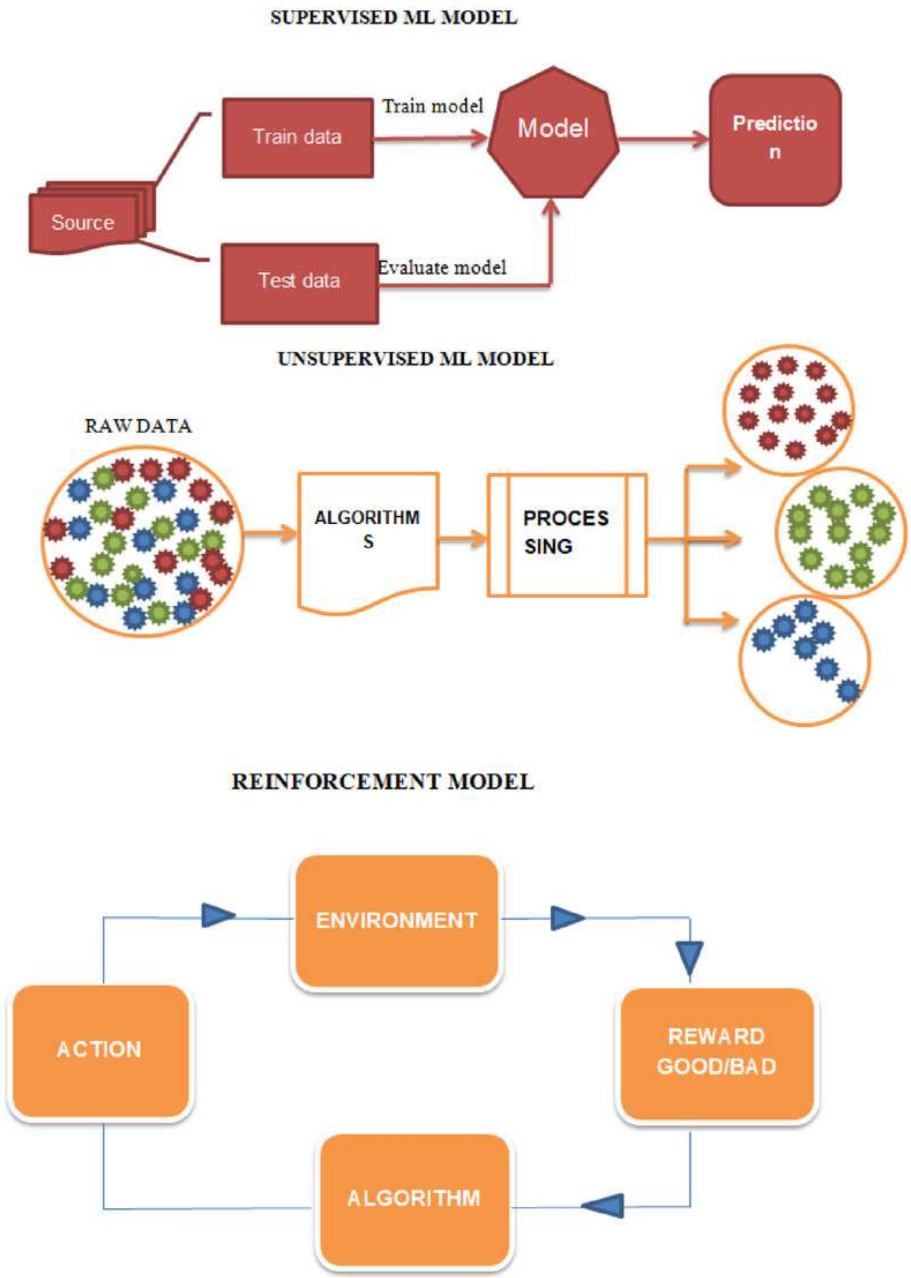

Figure 6

Estimates of data used in ML models in cerebral palsy research

Fig 1. Estimates of data used in ML models in cerebral palsy research

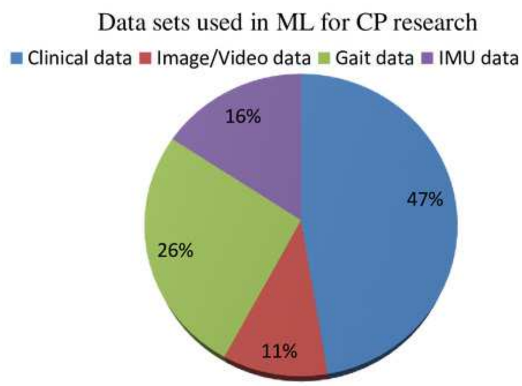

Figure 7

Estimates of ML models and their corresponding types of data used in cerebral palsy research

Fig 2. Estimates of ML models and their corresponding types of data used in cerebral palsy research

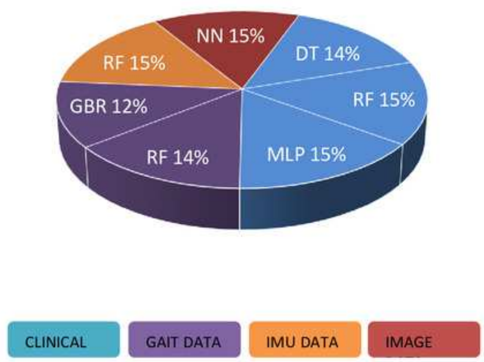

Figure 8

Machine learning models used for identification and diagnosis of cerebral palsy

Fig. 4: Machine learning models used for identification and diagnosis of cerebral palsy

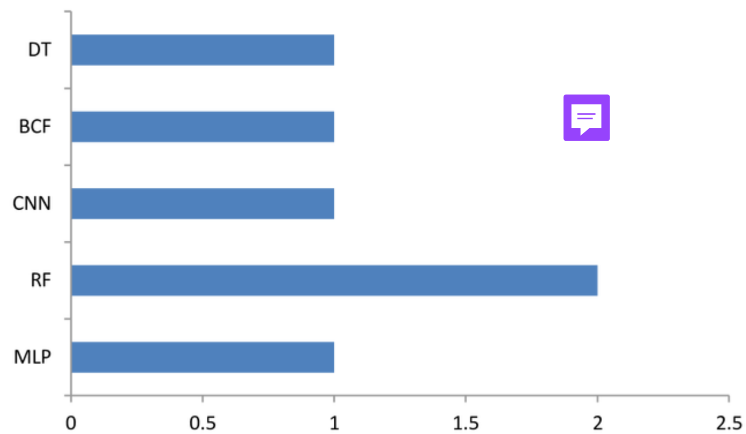

Figure 9

Use of machine learning models in cerebral palsy

Fig 3: Use of machine learning models in cerebral palsy

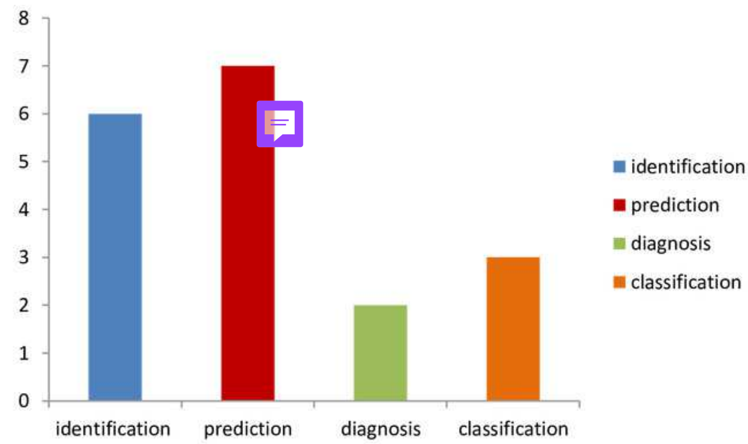

Table 1(on next page)

List of articles selected for analysis in this study

1 Table 1: List of articles selected for analysis in this study

Study	Year	Algorithm applied	Type of Data	Objectives	Outcome	Accuracy
Slijepcevic D et al. [9]	2023	CNN, self-normalizing neural networks, RF, DT	Gait Data	Classification of human gait patterns related to cerebral palsy (CP).	When contrasted, random forests and decision trees produce better outcomes and concentrate more on clinically pertinent regions.	93.4%
Carlo Marioi Bertoncelli [10]	2023	TT-PredictMed	Clinical data	Using a Predictive Model to Predict Postural Instability in Children with Cerebral Palsy	The predictive model's average accuracy was 82%, consistent with current research on applying machine learning models in the clinical setting.	82%
Mustafa Erkam Ozates et. al [11]	2023	CNN models	Gait Data	Joint moment prediction from kinematics	Joint movement kinematics may be predicted b the CNN model in cerebral palsy children.	nRMSE=18.02%-13.58%
L von Elling-Tammen et. al. [12]	2023	Feed-forward neural net (FNN) RF SVM Extreme Gradient Boosting (XG Boost)	Clinical data	To determine the precision of calculating the degree of gross motor impairment in kids and teenagers with CP	The most accurate algorithm was the random forest one.	nRMSE =10.1%
Krechowicz A et. al. [13]	2023	Adaptive Boosting Regression, KNN, DT, Regression, Random Forest Regression, and Gradient Boost Regression.	Gait Data	Identification of gait deviations in cerebral palsy children	The gradient-boosting regression model produced the best outcome.	79%
Abrar M. Al-Sowi [14]	2023	MLP, Naïve Bayes (NB) Random Tree (RT) and SVM	Clinical Data	Classification of cerebral palsy children using machine learning	Multilayer Perceptron (MLP) accurately classifies cerebral palsy.	84%
Michael H. Schwartz [8]	2022	Direct matching, virtual twins, and Bayesian causal forests	Clinical Data	To predict the average treatment effect on the treated 13 common orthopedic and neurological treatments using well-establish causal inference approaches.	Compared to other causal inference approaches, BCF performed remarkably well and offered more precise and accurate treatment predictions.	74%
Andrew Hua et al. [15]	2022	RF, LinearSVC, KNN), and MLP	IMU data	To identify machine learning models for categorizing kinematic data obtained from an IMU-based device while performing nine different upper extremity activities.	The RF models had the most excellent accuracy for categorizing kinematic data obtained from an IMU-based device while performing nine different upper extremity activities.	98.6%
Aleksander Palkowski [16]	2022	SVM, DT, RF,KNN	Clinical Data	To identify an automated limb exercise evaluation mechanism based	All the models achieved 100% accuracy in classifying whether an exercise was executed	100%

				on machine learning techniques.	well.	
JY Kim et al. [17]	2020	DT, RF, SVM, linear discriminant Analysis, and MLP	IMU data	To identify wearable technology and machine learning models to create a clinically helpful index while also giving rehabilitation patients a chance to track their level of spasticity even in settings outside healthcare facilities.	RF performed well among all the models.	95.4%
P. Illavarason [18]	2019	SVM, RF, and NN	Eye images	To diagnose cerebral palsy using eye images.	Neural Network(NN) is found to be the most accurate.	94.17%
Goodlich B et. al. [19]	2019	DT, SVM and RF	IMU data	To identify and test machine learning models for automatically detecting and categorizing Physical Activity types in CP children who utilize ambulation assistance.	RF model was the most accurate in automatically detecting and categorizing Physical Activity types in CP children who utilize ambulation assistance urate.	74%
Matthew Ahmadi [20]	2018	RF, SVM, and (BDT)	Clinical Data	To identify and evaluate ML models for automatically identifying Physical Activity in ambulant CP children.	SVM provided significantly better classification accuracy.	82% - 89%
Jing Zhang [6]	2017	SVM, NN, and Ada Boosted the decision tree and dynamic time warping.	Clinical Data	To identify the quality of exercises prescribed to CP	AdaBoosted decision tree performed the best with high classification accuracies.	90% - 94%
Ayat Naji Hussain et. al. [21]	2017	RF, DT, and KNN,	Gait Data	To classify foot diseases and find the accuracy of disease detection and diagnosis.	100% for Random Forest (RF), Decision Tree (DT), and k-nearest neighbors (KNN) and 98% for Logistic Regression.	98% - 100%
Paritosh Parmar [22]	2016	SVM, single and double-NN, boosted decision trees, and dynamic time warping (DTW),	Clinical data	To predict the quality of an exercise and judge if it was "good" or "bad."	The Ada Boosted tree fared the best proving the viability of exercise quality evaluation.	94.68
HG van den Boorn [23]	2016	SVR, ADA Boost regressor, RF, linear regression and Bayesian regression	Clinical data and EEG data	To predict the results of treatment plans for children with hand function impairment among cerebral palsy.	ML prediction is more accurate than Hand capacity test predictions.	All have low RMSE than Clinical tests.

Table 2(on next page)

Table 2: Machine Learning methods used in this study

Table 2: Machine Learning methods used in this study

Models	Citation
DT	[6], [9], [16], [21], [22].
RF	[9], [12], [15], [16], [17], [19], [21], [24], [25], [27], [28].
SVM	[16], [20].
GBR	[13], [27].
KNN	[16], [21], [27].
MLP	[14], [28]
ANN	[26], [27]
CNN	[11]
BCF	[8]
BR	[23]
LR	[29]
Ensemble Model	[31]

Table 3(on next page)

Types of data used by ML models in cerebral palsy research [6-23]

Table 3: Types of data used by ML models in cerebral palsy research [6-23]

Type of Data	ML models
Clinical Data	DT, RF,MLP, BCF, SVM, KNN, BR,LR
Gait Data	RF, GBR, CNN
IMU Data	RF
Image Video Data	NN

Table 4(on next page)

List of ML models used for identification and diagnosis of CP

Table 4: List of ML models used for identification and diagnosis of CP

Study	Year	Algorithm applied	Objectives	Outcome
Matthew Ahmadi [20]	2018	RF, SVM, and binary decision tree (BDT)	To establish and evaluate ML models for automatically identifying Physical Activity in ambulant CP children.	SVM provided significantly better classification accuracy
Goodlich B et. al. [19]	2019	DT,SVM RF	To identify and test machine learning models for automatically detecting and categorizing Physical Activity types in CP children who utilize ambulation assistance.	RF model was the most accurate in automatically detecting and categorizing Physical activity types in CP children who utilize ambulation assistance.
L von Elling-Tammen et. al. [12]	2023	Feed-forward neural net (FNN) Random forest (RF) Support vector machine (SVM) Extreme Gradient Boosting (XG Boost)	to identify the accuracy of measuring the level of gross motor impairment in children and adolescents with CP	The random forest algorithm proved to be the most accurate.
Krechowicz A et. al. [13]	2023	Adaptive Boosting Regression, K-nearest Neighbor, Decision Tree Regression, Random Forest	Identification of gait deviations in cerebral palsy children	The best result was obtained using the gradient-boosting regression model

		Regression, and Gradient Boost Regression.		
Andrew Hua et al. [15]	2022	Random Forest (RF), LinearSVC, k-Nearest Neighbors (kNN), and Multilayer Perceptron (MLP)	To identify machine learning models for categorizing kinematic data obtained from an IMU-based device while performing nine different upper extremity activities.	The RF models had the most excellent accuracy for categorizing kinematic data obtained from an IMU-based device while performing nine different upper extremity activities.
Aleksander Palkowski [16]	2022	support vector machines, decision trees, random forests, and k-nearest neighbors	To identify an automated limb exercise evaluation mechanism based on machine learning techniques.	All the models achieved 100% accuracy in classifying whether an exercise was executed well.
JY Kim et. al [17]	2020	Decision tree, Random forests (RFs), Support vector machine, Linear discriminant Analysis, and multilayer perceptrons	To identify wearable technology and machine learning models to create a clinically helpful index while also giving rehabilitation patients a chance to track their level of spasticity even in settings outside of healthcare facilities.	RF performed well among all the models.
Jing Zhang [6]	2017	SVM, neural networks, Ada Boosted decision tree,	To identify the quality of exercises prescribed to CP	AdaBoosted decision tree performed the best with high classification accuracies

		and dynamic time warping		
Fan H et al. [24]	2018	logistic Regression and Random Forest	Comparison of ML models to identify cases of cerebral palsy from unidentified cases	RF models are a reliable and affordable method to locate cerebral palsy instances that may not have been previously identified.
Varvara Turova et. al [25]	2020	RF	To identify cerebral hemorrhage in preterm infants using the RF model.	It has good predictability to identify cerebral hemorrhage in preterm infants.

Table 5(on next page)

List of ML models used for the classification of cerebral palsy

1 Table 5: List of ML models used for the classification of cerebral palsy
2

Study	Year	Algorithm applied	Objectives	Outcome
Slijepcevic D et al. [9]	2023	Convolutional neural networks, self-normalizing neural networks, random forests, and decision trees	Classification of human gait patterns related to cerebral palsy (CP).	Random forests and decision trees achieve better results and focus more on clinically relevant regions compared.
Abrar M. Al-Sowi [14]	2023	K-Star Multilayer Perceptron (MLP) Naïve Bayes (NB) Random Tree (RT) and Support Vector Machine (SVM)	Classification of cerebral palsy children using machine learning.	Multilayer Perceptron (MLP) accurately classifies cerebral palsy.
Ayat Naji Hussain et. al. [21]	2017	Random Forest (RF), Decision Tree (DT), and k-nearest neighbors (KNN),	To classify foot diseases and find the accuracy of disease detection and diagnosis	Random Forest (RF), Decision Tree (DT), and k-nearest neighbors (KNN) and 98% for Logistic Regression
Yanxin Zhang [26]	2019	artificial neural network (ANN), discriminant Analysis, naive Bayes, decision tree, *k*-nearest neighbors (KNN), support vector machine (SVM), and random forest	To assess the effectiveness of machine learning techniques for categorizing the gait patterns of CP children.	The most accurate prediction method is ANN.
R Abbas [27]	2018	ANN, MLP, SVM, KNN, DT classifiers, RF, GBM	To classify fetuses who are	All models show good accuracy. ANN,

			suffering from oxygen deprivation using ML models.	DT classifiers, RF, and GBM, have demonstrated good accuracies using various variables.

Table 6(on next page)

List of machine learning models used for Prediction in cerebral palsy cases

Table 6: List of machine learning models used for Prediction in cerebral palsy cases

Study	Year	Algorithm applied	Objectives	Outcome
Christian Morbidoni et. al. [28]	2021	Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbors (KNN), and Multilayer Perceptron (MLP)	To predict locomotion events n hemiplegic CP	MLP and RF have good prediction accuracies
Carlo Marioi Bertoncelli [10]	2023	TT-PredictMed	Using a Predictive Model to Predict Postural Instability in Children with Cerebral Palsy	The accuracy of the predictive model was 82% on average, which is in line with recent studies on using machine learning models in the clinical field.
Mustafa Erkam Ozates et. al [11]	2023	CNN models	Joint moment prediction from kinematics	Joint movement kinematics may be predicted b the CNN model in cerebral palsy children.
Michael H. Schwartz [8]	2022	Direct matching, virtual twins, and Bayesian causal forests	To predict the average treatment effect on the treated 13 common orthopedic and neurological treatments u sing well-establish causal inference approaches.	BCF performed exceptionally well and provided more accurate and precise treatment predictions than other causal inference methods

Paritosh Parmar [22]	2016	support vector machines (SVM), single and double-layered neural networks (NN), boosted decision trees, and dynamic time warping (DTW),	To predict the quality of an exercise and judge if it was "good" or "bad."	The Ada Boosted tree fared the best proving the viability of exercise quality evaluation.
HG van den Boorn [23]	2016	Support Vector Regressor, ADA Boost Regressor, Random Forest regressor, linear Regression and Bayesian Regression	To predict the results of treatment plans for children with hand function impairment among cere bral palsy.	Bayesian Regression has good accuracy in predicting treatment plans for children with CP.
Afifi J et al. [29]	2021	random forest (RF), logistic Regression	To predict CP in very preterm infants.	Both are comparable and predictable in CP prediction.
Adam Krechowicz [30]	2023	Adaptive Boosting Regression, K-nearest Neighbor, Decision Tree Regression, Random Forest Regression, and Gradient Boost Regression.	To predict gait deviation in cerebral palsy children	Gradient Boost Regression shows better result
Oliveira LB [31]	2023	Ensemble models, Support Vector Machines, and Artificial Neural Networks	To identify a machine learning model to predict patients' motor functions after therapy in rare disorders like cerebral	Ensemble models

			palsy.	