MACHINE LEARNING-BASED EARLY DETECTION AND INTERVENTION FOR MENTAL HEALTH ISSUES IN CHILDREN

Project Reference No.: *48S_BE_2565*

College	:	*B.G.S. Institute of Technology, Mandya*
Branch	:	*Artificial Intelligence and Machine Learning*
Guide(s)	:	*Mrs. Vidhya S G*
		Mrs. Afsha Firdose
Student(s)	:	*Ms. G A Shreya*
		Mr. M H Kruthik
		Mr. Sagar D P

Keywords:

Mental Health, Machine Learning, Depression, Anxiety, Early Diagnosis, Random Forest, Predictive Modelling.

Introduction:

In Mental health disorders in children, such as depression and anxiety, have become increasingly prevalent, yet early diagnosis remains a significant challenge. Traditional methods of mental health assessment rely on clinical evaluations and self-reported symptoms, which can be subjective and prone to underreporting. The integration of the machine learning technique into mental health screening presents an opportunity to enhance early detection and intervention strategies. By analyzing behavioral patterns, academic performance, and psychological indicators, ML models can identify children at risk of developing mental health disorders, enabling timely support and treatment. Earlier research has demonstrated that factors such as anxiety levels, panic attacks, academic stress, and personal experiences play a crucial role in determining a child's mental health status. However, inconsistencies in self-reported data and variability in psychological assessments often impact the reliability of traditional diagnostic approaches Machine learning, through its ability to analyze large datasets and identify complex patterns, provides a data-driven solution to overcome these challenges. By leveraging classification models, including Decisions Tree, Naive

Bayes, Logistic Regressions, and Random Forest, this study aims to develop a predictive system capable of detecting early symptoms of childhood depression and anxiety. A Random Forest (RF) classifier, known for its high accuracy in classification tasks, is particularly effective in handling diverse input features, such as demographic data, academic performance, and behavioral responses. The RF model operates by constructing multiple decision trees and aggregating their outputs to improve prediction reliability. This ensemble approach minimizes overfitting and enhances the robustness of predictions. Through this study, we seek to the determine most effective machine learning model of for early detections of childhood mental health disorders. By optimizing model parameters and refining input features, the research aims to contribute to the development of AI-driven mental health screening tools, which can assist educators, healthcare professionals, and parents in identifying at-risk children before their conditions worsen.

Figure 1: Home Page

Objectives:

- Develop machine learning models to predict childhood mental health disorders.
- Compare the performance of Decision Tree, Naïve Bayes, Logistic Regression, and Random Forest algorithms.
- Fine-tune hyperparameters to optimize model accuracy and reliability.
- Create a data-driven tool for early diagnosis, minimizing subjective biases.

1. **Develop predictive machine learning models** for early detection of childhood mental health disorders.

2. **Analyse and compare** different machine learning algorithms (Decision Tree, Naïve Bayes, Logistic Regression, Random Forest) for predictive accuracy.
3. **Optimize model performance** through hyperparameter tuning and feature selection to enhance prediction reliability.
4. **Create a data-driven screening tool** that assists healthcare professionals and educators in identifying at-risk children early, minimizing subjective assessment errors

Methodology:

- Dataset Collection: Data includes behavioural indicators, academic records, anxiety levels, and history of specialist treatment.
- Feature Selection: Important attributes like age, gender, CGPA, anxiety levels, and panic attacks are extracted.
- Data Splitting: The dataset is divided into training and testing sets after feature engineering.
- Model Training: Models including Decision Tree, Naïve Bayes, Logistic Regression, and Random Forest are trained.
- Model Evaluation: Accuracy, precision, recall, and F1-scores are analyzed using unseen data.
- Algorithm Selection: Hyperparameter tuning is done for models that underperform.
- Final Model Selection: The best-performing model is selected and finalized

Result and Conclusion:

Random Forest achieved the highest accuracy of 90.95% in predicting mental health risks compared to other models (Decision Tree, Naïve Bayes, and Logistic Regression at 76.19%). Confusion matrix analysis revealed high precision and recall but indicated some false positives and negatives, suggesting areas for refinement. The findings confirm that ML can successfully predict early signs of mental health issues, supporting proactive interventions. The integration of AI-powered screening tools into schools and healthcare can facilitate early and personalized mental health care.

Future Scope:

The project can be expanded by incorporating real-time monitoring using wearable sensors and integrating multimodal data (text, voice, and facial analysis) for better prediction accuracy. Further improvements include building larger, more diverse datasets, enhancing model transparency through explainable AI, and developing real-time stress detection systems. Ethical considerations, especially data privacy, must remain a priority. These advancements can make mental health care more accessible, effective, and proactive, transforming how early mental health risks are diagnosed and treated

The future scope of this project includes:

1. Incorporating real-time monitoring solutions using wearable devices for enhanced stress and anxiety detection.
2. Expanding the dataset diversity to include multi-regional and multi-cultural data for better generalization.
3. Integrating multimodal data sources such as facial expressions, speech patterns, and physiological signals for comprehensive analysis.
4. Enhancing model transparency through the use of explainable AI techniques to improve trust and acceptance among healthcare professionals.

Project Outcome & Industry Relevance:

This project demonstrates the feasibility of using machine learning for early mental health detection in children. It highlights how AI-driven tools can enhance traditional mental health screenings, providing more objective, scalable, and timely assessments. In real-world settings, such a system can assist schools, clinics, and pediatricians in proactively identifying children at risk, thus enabling earlier interventions. Industries like educational technology, healthcare, and mental health services can integrate these solutions into their workflows, improving child welfare and reducing long-term societal costs related to untreated mental health issues.

Working Model vs. Simulation/Study:

This project was primarily a **simulation and theoretical study**. It involved the development, training, and evaluation of machine learning models using real-world survey datasets but did not involve creating a physical working prototype.

Project Outcomes and Learnings:

The key outcomes of this project include the successful application of machine learning models for mental health prediction, with Random Forest achieving the highest accuracy. The project emphasized the importance of data quality, feature selection, and hyperparameter optimization in enhancing predictive performance. A major learning was the critical role of addressing bias, privacy concerns, and model interpretability when dealing with sensitive health data. The team also gained experience in comparative model analysis and the challenges of real-world mental health prediction.