NITHIN NARAYANAN P B
NITHINPB010@GMAIL.COM
AMRITA VISHWA VIDYAPEETHAM
TAMIL NADU
ARTIFICIAL INTELLIGENCE INTERNSHIP
12/06/23- 06/08/2023

MACHINE LEARNING FOR MENTAL FITNESS TRACKING

- Mental health issues are becoming increasingly prevalent in today's society. Depression and anxiety are among the most common mental health disorders, affecting millions of people worldwide. Unfortunately, many individuals who suffer from these conditions do not receive the help they need due to a lack of awareness or stigma surrounding mental illness. This is where mental fitness tracking comes in. By monitoring one's mental health on a regular basis, individuals can gain insight into their emotional well-being and take proactive steps to improve it. Mental fitness tracking can also help identify potential mental health issues early on, allowing for prompt intervention and treatment. The results of this project could have a significant impact on the way mental health is tracked and managed. The model could be used to provide personalized feedback and recommendations to help people improve their mental fitness. It could also be used to identify people who are at risk of developing mental health problems, so that they can receive early intervention.

AGENDA

- In this project, we will be discussing the need for mental fitness tracking and introducing our solution to address this growing problem.
- We will identify the target audience for our mental fitness tracking solution, explain its value proposition, and discuss the technical aspects of our solution. We will also present the results of our solution and its impact on the target audience, and finally, we will discuss the potential future directions of mental fitness tracking and its potential to impact mental health on a larger scale.

PROJECT OVERVIEW

- **Purpose**: The project aims to develop an innovative machine learning-based system for tracking and enhancing mental fitness on an individual level. By integrating diverse data sources and leveraging predictive modeling, the system provides real-time insights and personalized recommendations to support users' mental well-being.

- **Scope:**

1. Data Acquisition and Preprocessing: Gathering and merging two datasets containing prevalence and disease share information for different mental health disorders. Data is cleaned, transformed, and prepared for modeling.
2. Feature Engineering: Identifying relevant features and mapping them to suitable labels. Utilizing a Label Encoder for converting non-numeric labels into numeric representations.
3. Model Selection and Hyperparameter Tuning: Employing the Gradient Boosting Regressor algorithm to create predictive models. Tuning hyperparameters such as learning rate, max depth, and number of estimators to optimize model performance.
4. Model Training and Evaluation: Splitting the data into train, validation, and test sets. Training the model on the training set and evaluating its performance using Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R2) metrics on the validation set.
5. Insights and Reporting: Generating insights into the relationship between mental health disorders and DALY. Presenting the project findings through performance metrics and model analysis.

- **Objectives**:

1. Develop a robust mental health tracking system for predicting DALY associated with mental health disorders.
2. Explore the impact of various mental health disorders on overall health outcomes.
3. Implement an effective data preprocessing pipeline to handle merging, cleaning, and transformation of datasets.
4. Optimize model performance through hyperparameter tuning and model selection.
5. Provide valuable insights into the significance of mental health within the realm of public health.
6. Create a project report that documents the entire process, from data preprocessing to model evaluation, for future reference and dissemination.

WHO ARE THE END USERS OF THIS PROJECT?

- The target audience and end users of the Mental Health Tracking System encompass a diverse group of stakeholders, each with unique characteristics, needs, and potential benefits from the developed solution.
- **1. Public Health Officials and Policy Makers:** These professionals are responsible for formulating public health policies and strategies. They require accurate insights into the prevalence and impact of mental health disorders to allocate resources effectively, design intervention programs, and make informed policy decisions. The system's predictive models can assist them in understanding the distribution of DALY across different mental health disorders, enabling evidence-based policy development.
- **2. Healthcare Providers:** Medical practitioners, psychiatrists, psychologists, and counselors need comprehensive information on mental health disorders to provide appropriate care and treatment. The system's predictions can guide them in identifying high-risk patient populations and tailoring treatment plans to address specific disorders, ultimately leading to more personalized and effective interventions.
- **3. Researchers and Academia:** Researchers and academics studying mental health epidemiology and its relation to overall health outcomes can benefit from the system's insights. The system can serve as a valuable tool for conducting further research, validating hypotheses, and contributing to the scientific understanding of mental health's impact on public health.
- **4. Non-Governmental Organizations (NGOs) and Advocacy Groups:** NGOs and advocacy groups focused on mental health awareness and support can utilize the system's findings to raise awareness about the importance of mental well-being. They can leverage the data to advocate for improved mental health resources, reduce stigma, and promote early intervention and prevention strategies.
- **5. General Public and Individuals:** The general public and individuals seeking information about mental health can access the system's user-friendly visualizations and reports. By understanding the significance of different mental health disorders in terms of DALY, individuals can make informed lifestyle choices, prioritize mental well-being, and seek timely help when needed.

YOUR SOLUTION AND ITS VALUE PROPOSITION

- Our solution is a robust Mental Health Tracking System that utilizes machine learning techniques to predict Disability-Adjusted Life Years (DALY) associated with various mental health disorders. This comprehensive system addresses the critical need for accurate insights into the impact of mental health on overall well-being, benefiting a diverse range of end users.
- **Value Proposition:**

1. **Accurate Predictive Models:** Our system employs advanced machine learning algorithms, specifically the Gradient Boosting Regressor, to create accurate predictive models. By leveraging historical prevalence and disease share data, the system provides precise estimates of DALY for different mental health disorders. This accuracy is crucial for informed decision-making by public health officials, policy makers, and healthcare providers.
2. **Evidence-Based Policy Development:** Public health officials and policy makers can utilize the system's predictions to develop evidence-based policies and strategies. The insights gained from the system allow them to allocate resources efficiently, prioritize interventions, and design targeted programs to address specific mental health disorders. This, in turn, can lead to improved overall public health outcomes.
3. **Tailored Treatment Plans:** Healthcare providers can benefit from the system's predictions to tailor treatment plans for individuals with mental health disorders. The system's insights help identify high-risk patient populations and guide the customization of interventions, resulting in more effective and personalized patient care.
4. **Informed Research and Advocacy:** Researchers, academia, and advocacy groups can leverage the system's findings to conduct in-depth research on mental health epidemiology. The accurate DALY predictions contribute to a better understanding of the relationship between mental health disorders and overall health outcomes. Advocacy groups can use the data to raise awareness, reduce stigma, and advocate for improved mental health resources.
5. **Empowering Individuals:** The system's user-friendly visualizations and reports empower the general public and individuals with valuable information about mental health. By understanding the DALY associated with different disorders, individuals can make informed decisions to prioritize mental well-being, seek early intervention, and adopt healthier lifestyles.

HOW DID YOU CUSTOMIZE THE PROJECT AND MAKE IT YOUR OWN

- In this customized project, I have taken the existing mental health tracking concept and enhanced it with a unique focus on precision and accessibility. I introduced the use of the Gradient Boosting Regressor, a powerful algorithm, for predicting DALY associated with mental health disorders, resulting in more accurate estimations. Moreover, I incorporated a user-friendly interface that generates insightful visualizations and reports, making complex data easily understandable for individuals and non-experts. This innovative approach bridges the gap between technical accuracy and user accessibility, empowering both experts and the general public to comprehend and act upon the significance of mental health in a more informed and engaging manner.

MODELLING

- In this project, I employed the Gradient Boosting Regressor as the core modeling technique. Gradient Boosting is an ensemble learning method that builds multiple decision trees sequentially, each correcting the errors of the previous one. This approach effectively captures complex relationships within the data, making it well-suited for predicting the Disability-Adjusted Life Years (DALY) associated with different mental health disorders.
- **Utilization of Technology Principles:**

1. **Data Preprocessing Pipeline:** I established a robust data preprocessing pipeline using the Pandas library, which involved merging datasets, handling missing values, and transforming non-numeric labels using LabelEncoder. This ensured that the data was clean, structured, and suitable for modeling.
2. **Hyperparameter Tuning:** Leveraging the principles of hyperparameter tuning, I fine-tuned the learning rate, max depth, and number of estimators for the Gradient Boosting Regressor. This optimization process improved model accuracy and generalization.
3. **Train-Validation-Test Split:** By partitioning the dataset into distinct train, validation, and test sets, I followed the fundamental principle of model evaluation. This enabled me to train the model on one subset, validate its performance on another, and assess its final performance on unseen data.
4. **Performance Metrics:** The use of mean squared error (MSE), mean absolute error (MAE), and R-squared (R2) scores to evaluate the model's performance adhered to best practices in regression analysis. These metrics quantified the model's accuracy, error, and explanatory power.
5. **Visualization and Reporting:** Employing Python's libraries for data visualization, I created intuitive charts and reports to present the project's findings. This facilitated clear communication of complex insights to diverse audiences, aligning with modern technology-driven communication practices.

- By integrating these techniques and principles, my solution enhances accuracy, scalability, and interpretability, while showcasing a comprehensive understanding of modern data science methodologies and their application to mental health tracking.

RESULTS

```
Gradient Boosting Mean Squared Error: 0.0292
Gradient Boosting Mean Absolute Error: 0.1154
Gradient Boosting R-squared: 0.9942
```

- The developed Mental Health Tracking System demonstrates impressive performance in predicting Disability-Adjusted Life Years (DALY) associated with various mental health disorders. The outcomes of the project highlight the effectiveness and success of the solution.
- **Quantitative Performance Metrics:**
- Gradient Boosting Mean Squared Error: 0.0292
- Gradient Boosting Mean Absolute Error: 0.1154
- Gradient Boosting R-squared: 0.9942
- These quantitative metrics reflect the high accuracy and predictive power of the Gradient Boosting Regressor model. The low Mean Squared Error (MSE) and Mean Absolute Error (MAE) values indicate minimal prediction errors, while the high R-squared (R2) score of 0.9942 signifies a remarkable ability to explain the variance in the target variable.

RESULTS AND LINK

- **Interpretation and Implications:**
- The outcomes of the project demonstrate that the developed system successfully captures the complex relationships between prevalence rates of mental health disorders and their associated DALY. The high R2 score suggests that approximately 99.42% of the variability in DALY can be explained by the model, indicating strong predictive performance.
- These results have significant implications for various stakeholders, including public health officials, healthcare providers, researchers, and advocacy groups. The accurate predictions provided by the system can inform evidence-based policy decisions, personalized treatment plans, research studies, and awareness campaigns.
- **Future Applications:**
- The success of the project opens avenues for future enhancements and applications. The system could be expanded to incorporate real-time data, enabling continuous tracking and updating of DALY predictions. Additionally, the model's architecture could be further optimized using advanced techniques like feature selection and ensemble methods to potentially improve its performance even more.
- **Link to repository:**
- MENTAL HEALTH TRACKER

THANK YOU