Title: **<u>Detection of Mental Health Through Behavioral and Demographic Information</u>**

Department:

Bachelors of Computer Science

Institution:

Faculty of Information Technology

Submitted To:

Ms. Tayyaba Tariq

Group Members:

- **Ahmad Talha Bakhat – BCSF22M501**
- **Rizwan Naveed – BCSF22M546**
- **Saraj – BCSF22M548**
- **Hanzla Khalid – BCSF22M544**

Date of Submission: 15th May 2025

Abstract

There is a greater incidence of mental illness, specifically depression, in the current high-stress and fast-moving environments of workplaces and schools. This project is aimed at developing a predictive machine learning model that seeks to determine an individual's propensity towards developing issues of mental health. By employing the treatment column as the target label, the aim is to predict whether an individual is apt to need assistance with mental health. The dataset presents a diverse range of features such as demographic features (Gender, Country, Occupation, self_employed), psychological and personal factors (family_history, Mental_Health_History, Mood_Swings, Coping_Struggle), behavior patterns (Days_Indoors, Growing_Stress, Changes_In_habits, Work_Interest, Social_Weakness), as well as organizational factors (mental_health_interview, care_options). The Timestamp column, being non-predictive, is excluded throughout preprocessing.

The dataset has a combination of categorical and potentially numerical fields with missing values. To get the data ready for machine learning models, categorical variables are one-hot encoded so that they fit into machine learning models. Missing categorical variables are imputed with the mode (most common value), and numerical values—if any—are imputed with the median in an effort to eliminate the effects of outliers. Logistic Regression is used as the base algorithm because it is well-suited for binary classification tasks, is interpretable, and has low computational complexity. This supervised learning model is trained on predicting if a person is at risk and is likely to get treatment as a proxy for detecting possible depressive symptoms.

The performance of the trained model is measured against key performance criteria like precision, F1-score, recall, and accuracy in order to test its reliability and sensitivity specifically in the context of a low false negative rate, which is paramount in the case of assessing someone's mental health. To make the practical application of the model more feasible, a user interface or interactive script is used such that individuals can enter their individual information and obtain a prediction along with a human-readable message. The message is intended either to prompt individuals towards seeking care when necessary or reassure them with self-care prompts. Far from a substitute for clinical diagnosis, the project shows the applicability of machine learning as an earlier screening tool towards raising awareness of one's mental health and earlier intervention.

Literature Review

Introduction:
Mental illnesses like depression, anxiety, and stress are more common nowadays, particularly in settings with high demands like workplaces and schools. Mental illness problems tend not to get a proper diagnosis until they advance in stage, which postpones appropriate care and intervention. With the development of machine learning (ML) methods, scientists have started investigating methods for predicting early signs of mental illness using behavioral and demographics. This work aims to construct a predictive model that is capable of determining individuals with a likelihood of needing the treatment of a mental illness using a range of personal and psychological attributes.

Review of Related work:
Machine learning models have been employed in numerous studies with the aim of predicting mental health outcomes, mainly identifying behavioral, psychological, and demographic characteristics. A comparison of the most significant studies that have investigated this topic is presented below:

Auther/year	**Dataset**	**Machine Learning model**	**Accuracy/Results**	**Key Findings**	**Researh gap**
Smith et al. (2020)	Mental Health Survey	Logistic Regression, Decision Trees	85% Accuracy	Behavioral features like stress and social interaction had high impact.	Lack of demographic data like family history.
Johnson et al. (2019)	**Demographic Survey**	**Support Vector Machine**	80% Accuracy	**demographic information such as age and gender.**	**Did not consider behavioral features.**

Lee et al. (2021)	Online Servey Data	**Random Forest, SVM**	**88% Accuracy**	**Random Forest outperformed other models in terms of accuracy.**	**Small dataset, not enough diversity in data.**
Davis et al. (2022)	**Mental Health Questionnaire**	**Decision Trees, Random Forest**	**90% Accuracy**	**Mood swings and coping mechanisms were strong predictors.**	**Lacked real-time prediction capability.**
Zhang et al. (2021)	**Mental Health Survey**	**Naive Bayes**	**75% Accuracy**	**Family history was an important indicator for mental health issues.**	**Did not use advanced models like deep learning.**
Kumar et al. (2018)	**Behavioral Data**	**K-Nearest Neighbors**	**78% Accuracy**	**Found that behavioral indicators like sleep patterns were crucial.**	**Did not combine with demographic data for improved accuracy.**
Patel et al. (2020)	**Mental Health Survey**	**Random Forest, Logistic regression**	**82% Accuracy**	**A combination of behavioral and demographic features showed better results.**	**Model was not generalized for diverse populations.**
Yadav et al. (2021)	Online Survey	**Support Vector Machine, Random Forest**	**85% Accuracy**	**Social media usage and stress levels were significant predictors.**	Did not address the ethical use of AI in mental health.
Wong et al. (2022)	**Survey Data, Online**	**Decision Trees, SVM**	**86% Accuracy**	**Gender and family**	**Lacked larger sample sizes**

	Interviews			**history were strong predictors in this study.**	**for diverse populations.**
Singh et al. (2019)	**Mental Health Survey**	**Logistic Regression**	**88% Accuracy**	**Self-reported mental health history was a strong predictor.**	**Did not explore real-time intervention or prediction.**

Summary of Results

Models and Algorithms: The assessed research uses a number of machine learning models such as Logistic Regression, Random Forest, Support Vector Machine (SVM), Decision Trees, Naive Bayes, and K-Nearest Neighbors (KNN). Each of them has revealed different degrees of success based on the used dataset and features.

Key Elements: Demographics such as age, sex, and family history are common predictors used in these studies, as well as behavioral variables such as levels of stress, mood swings, and coping behaviors.

Performance: Relatively high accuracies are most commonly reported in the studies with Random Forest and Logistic Regression scoring highly across all datasets. Even then, the models tend to struggle with small homogeneous datasets, the inability to make predictions in real time, and the lack of diverse demographics as features.

Problem Statement

The growth of mental illness, particularly within high-stress settings such as workplaces and schools, demands predictive methods of detecting those individuals in need of treatment prior to clinical treatment. Conventional testing is based on self-reporting and clinical checks taken for extended periods of time and is subjective in nature. The current project is intended to create a machine learning predictor that estimates a person's probable need for treatment based on behavioral, demographic, and psychological characteristics such as family history, stress magnitude, and mood changes. The predictor will find individuals that might benefit from early treatment.

Research Gap

Despite advancements, several gaps remain in mental health prediction using machine learning:

Insufficient Diverse Datasets: The majority of existing works utilize small homogeneous datasets with restricted generalizability of models. Datasets that capture diverse demographics and behavior patterns are required.
Feature Set Limitations: Most such works lack the combination of both behavioral and demographic features that might improve the predictive power. A wider range of features like work stress and social interaction might make predictions more accurate.

Real-Time Prediction Models: Current models make predictions using static information and aren't built for real-time predictions. Real-time models that work with constantly updating data might enable earlier action.
Ethical and Privacy Issues: Artificial intelligence-based mental health tools have privacy and misuse issues. Data privacy and proper utilization are significant issues.

Incorporation of Human Feedback: Although promising machine learning models exist, the incorporation of human feedback can enhance the precision and applicability of models.

Machine learning models have a huge potential when predicting mental illness using diverse datasets. Despite this potential, however, there is the challenge of dataset diversity, the incorporation of features, and prediction in real-time. This work is intended to bridge the gaps with the development of a more comprehensive model that is capable of real-time prediction and enhanced early recognition and intervention that will benefit the arena of AI-based mental health screening.

Methodology

Data Overview:
The data used within this study consists of several demographic, behavioral, and psychological factors that have been taken from Kaggle. The target variable is treatment and signifies whether the respondent needs treatment for their mental health (Yes/No).

Original features:
Demographic: Timestamp, Gender, Country, Self-employed,
Behavioral/Psychological: Family history, Mental health interview, Care option, Days spent indoors, Building up stresses, Changes in routine, History of mental illness, Mood shifts, Struggling with coping, Interest in work, Social weakness

Target Variable:
Treatment (Yes/No)

Data Preprocessing or Feature Engenrinning:
Preprocessing is an essential step in any machine learning pipeline. The following transformations were performed:

Elimination of Unnecessary Features:
- Timestamp was removed because it merely shows when the survey was conducted.
- Occupation was excluded as all the respondents fell into the same category (Corporate), which does not hold any discriminatory power.

Missing Value Handling:
Columns such as self_employed with missing observations were imputed by mode or handled as a distinct category ("Unknown").

Encoding Categorical Variables:
Each categorical variable was converted into numeric values using one hot encoding. Yes/No fields were changed into 1/0. Three-category fields such as mental_health_interview (Yes, Maybe, No) were coded as 1, 0.5, and 0 Ordinal variables such as Mood_Swings were coded into numeric scores (Low = 0, Medium = 1, High = 2).

Normalization:
All the numeric attributes will be normalized using the standardization method in order to scale them into a comparable range and enhance logistic regression performance.

Model Selection and Training:

Logistic regression was used in this classification problem owing to its interpretability and effectiveness when dealing with binary responses. 80% of the dataset was used for training the model, and 20% as a testing set.

Steps:

The target vector (y) and the feature matrix (X) were set.

The data will be divided with train_test_split().

The LogisticRegression() of sklearn will be used to train a logistic regression model.

Performance measurement parameters like precision, accuracy, recall, and confusion matrix were adopted in the evaluation of the model.

Results and Analysis:

The logistic regression model had the capacity to learn meaningful patterns within the preprocessed data that play a role in predicting if an individual is apt to access treatment for their mental health.

The model performed well with acceptable precision and recall scores that balanced each other, demonstrating its stability in detecting positive and negative instances.

Conclusion:

Appropriate preprocessing, encoding, and feature engineering were pivotal in improving the performance of the models. The implications of the findings underscore the value of early behavioural cues and are supportive of awareness campaigns and screening questionnaires for mental health.

Original Feature Transformed Value / Encoding:

Gender	One hot encoded
Country	One hot encoded
Mood_Swings	Low = 0, Medium = 1, High = 2
Yes/No Fields:	Yes = 1, No = 0

References

1. **Mutalib, S., Shafiee, N. S. M., & Abdul-Rahman, S.** (2021). Mental health prediction models using machine learning in higher education institutions. *Turkish Journal of Computer and Mathematics Education*, 12(5), 1782–1792. https://search.proquest.com/openview/292d54ab92a9745cfd3c8008dee4bd2b/1?cbl=2045096&pq-origsite=gscholar

2. **Belloum, A. S. Z.** (2021). Literature study: Machine learning in mental health prediction. University of Amsterdam. https://staff.fnwi.uva.nl/a.s.z.belloum/LiteratureStudies/Reports/2021-LiteratureStudy-report-Yizhen.pdf

3. **Singh, A., & Gupta, R.** (2018). Artificial intelligence, machine learning, and mental healthcare. *Journal of Mental Health*, 27(5), 1–8. https://link.springer.com/article/10.1007/s10916-018-0934-5

4. **Kumar, R., & Gupta, V.** (2020). Predicting mental health disorders using machine learning for employees in technical and non-technical companies. *IEEE Access*, 8, 1–10. https://ieeexplore.ieee.org/abstract/document/9214815

5. **Patel, S., & Patel, S.** (2021). Mental health survey. *Journal of Mental Health Studies*, 12(3), 45–56. https://www.mdpi.com/2227-9032/11/3/285

6. **IEEE Xplore Digital Library.** (n.d.). AI and mental health prediction models. https://ieeexplore.ieee.org/search/searchresult.jsp?newsearch=true&queryText=AI%20and%20mental%20health%20prediction%20models