

ABSTRACT

Mental health has emerged as a crucial area of concern in **modern healthcare**, with millions of individuals affected by conditions such as depression, anxiety, bipolar disorder, and stress-related illnesses. Despite growing awareness, early identification and intervention remain a challenge due to the stigma surrounding mental health, the subjective nature of self-reporting, and **limited access to clinical** diagnosis. To address these challenges, this project explores the application of machine learning (ML) techniques for risk assessment and early detection of mental health issues.

This study utilizes publicly available datasets containing **psychological, demographic, and behavioral attributes** such as age, gender, work-life balance, social interactions, and past mental health history. Various supervised learning algorithms—such as Logistic Regression, Random Forest, Support **Vector Machine (SVM)**, and Artificial Neural Networks—are applied to classify individuals into risk categories (e.g., low, moderate, and high risk). Data preprocessing, feature selection, and model optimization techniques are employed to enhance predictive performance and reduce bias.

The proposed system demonstrates that machine learning can be an effective tool for mental health screening, **achieving high levels of accuracy, precision, and recall**. The model can assist healthcare providers, counselors, and organizations in **identifying at-risk individuals** for timely intervention. Ethical considerations such as data privacy, informed consent, and algorithmic fairness are emphasized throughout the project to ensure responsible AI use in sensitive domains like mental health.

In conclusion, this work highlights the potential of data-driven approaches to supplement traditional clinical methods, **enabling scalable and proactive mental health care solutions**. Future enhancements may include real-time risk tracking through wearable devices, integration with electronic health records, and the incorporation of **Natural Language Processing (NLP)** for deeper psychological insight from text-based data.

TABLE OF CONTENTS

| Chapters | Page No. |
|---------------------------------|----------|
| 1 Introduction | |
| 1.1 Background of the problem | 1 |
| 1.2 Motivation | 1 |
| 1.3 Problem statement | 2 |
| 1.4 Objectives | 2 |
| 2 Literature Survey | 3-4 |
| 3 System Requirements | 5 |
| 3.1 Hardware Requirements | |
| 3.2 Software Requirements | |
| 4 Dataset Description | 6 |
| 4.1 Source of dataset | |
| 5 Methodology | 7-8 |
| 5.1 Block Diagram | |
| 6 Implementation | 9-12 |
| 6.1 Code | |
| 7 Results and Evaluation | 13-15 |
| 7.1 Screenshot | |
| 8 Conclusion | 16 |
| References | 17 |

LIST OF FIGURES

| Figures | Page No. |
|--|-----------------|
| Figure 1.1: Block diagram | 8 |
| Figure 2.1: Output Interface | 13 |
| Figure 2.2: Feature Importances | 14 |
| Figure 2.3: Age Distribution by treatment status | 14 |
| Figure 2.4: Heat Map | 15 |

LIST OF TABLES

| Table No. | Table Name | Page. No |
|------------------|-----------------------|-----------------|
| Table 3.1.1 | Hardware Requirements | 5 |
| Table 3.1.2 | Software Requirements | 5 |
| Table 4.1 | Dataset Description | 6 |