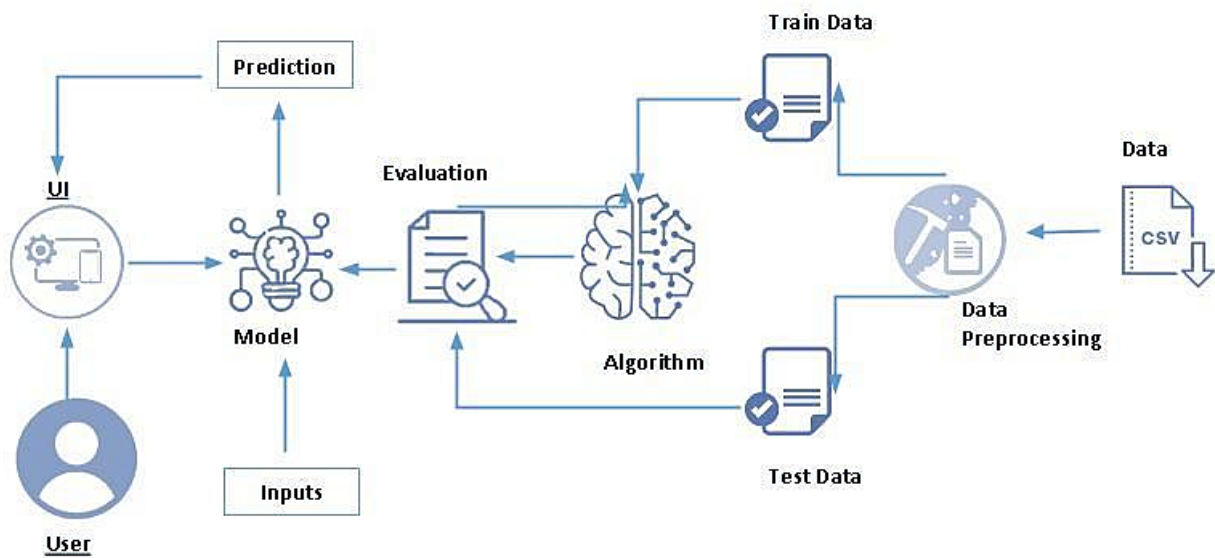


**SESHADRI RAO GUDLAVALLERU
ENGINEERING COLLEGE**



Mental Health Prediction

Using IBM Watson



Submitted By:
V.N.V.Harshitha
(19481A0281)

Content

1. INTRODUCTION

a. Abstract

The use of this project. What can be achieved using this.

b. Overview

A brief description about your project

2. LITERATURE SURVEY

a. Existing problem

Existing approaches or method to solve this problem

b. Proposed solution

What is the method or solution suggested by you?

3. THEORITICAL ANALYSIS

a. Block diagram

Diagrammatic overview of the project.

b. Hardware / Software requirements

Hardware and software requirements of the project

4. EXPERIMENTAL INVESTIGATIONS

Analysis or the investigation made while working on the solution

5. FLOWCHART

Diagram showing the control flow of the solution

6. RESULT

Final findings (Output) of the project along with screenshots.

7. ADVANTAGES & DISADVANTAGES

List of advantages and disadvantages of the proposed solution

8. CONCLUSION

Conclusion summarizing the entire work and findings.

9. FUTURE SCOPE

Enhancements that can be made in the future.

10. BIBILOGRAPHY

References of previous works or websites visited analysis about the project.

1. INTRODUCTION

1.1 Abstract:

The increase of mental health problems and the need for effective medical health care have led to an investigation of machine learning that can be applied in mental health problems. This report presents a recent systematic review of machine learning and deep learning approaches in predicting mental health problems. Furthermore, we will discuss the challenges, limitations, and future directions for the application of machine learning in the mental health field. We collect research articles and studies that are related in predicting mental health problems by searching reliable databases. Then, we categorize the collected research articles based on the mental health problems such as schizophrenia, bipolar disorder, anxiety and depression, posttraumatic stress disorder, and mental health problems among children. Discussing the findings, we reflect on the challenges and limitations faced by the researchers on machine learning in mental health problems. Additionally, we provide concrete recommendations on the potential future research and development of applying machine learning and deep learning in the mental health field. The main purpose of the Mental Health Prediction system is to predict whether a person needs to seek Mental health treatment or not based on inputs provided by them.

1.2 Overview:

Mental Health First Aid teaches participants how to notice and support an individual who may be experiencing a mental health or substance use concern or crisis and connect them with the appropriate employee resources. Employers can offer robust benefits packages to support employees who go through mental health issues. That includes Employee Assistance Programs, Wellness programs that focus on mental and physical health, Health and Disability Insurance, or flexible working schedules or time off policies. Organizations that incorporate mental health awareness help to create a healthy and productive work environment that reduces the stigma associated with mental illness, increases the organizations' mental health literacy, and teaches the skills to safely and responsibly respond to a co-worker's mental health concern. The main purpose of the Mental Health Prediction system is to predict whether a person needs to seek Mental health treatment or not based on inputs provided by them. We will be using classification algorithms such as Logistic Regression, KNN, Decision tree, Random Forest, AdaBoost, GradientBoost, and XGBoost. We will train and test the data with these algorithms. From this, the best model is selected and saved in pkl format. We will also be deploying our model locally using Flask.

2.LITERATURE SURVEY

2.1 Existing problem

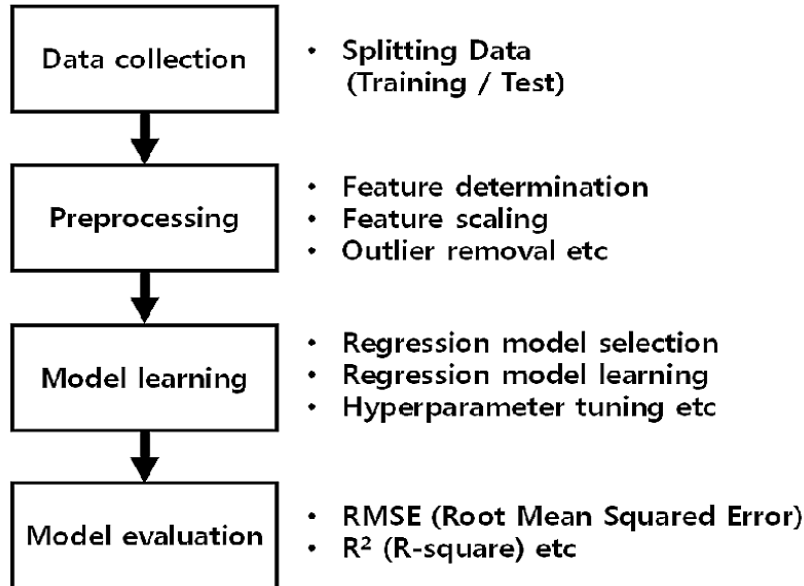
Mental illness is a health problem that undoubtedly impacts emotions, reasoning, and social interaction of a person. These issues have shown that mental illness gives serious consequences across societies and demands new strategies for prevention and intervention. To accomplish these strategies, early detection of mental health is an essential procedure. Mental illness is usually diagnosed based on the individual self-report that requires questionnaires designed for the detection of the specific patterns of feeling or social interactions. With proper care and treatment, many individuals will hopefully be able to recover from mental illness or emotional disorder .

2.2 Proposed solution

Machine learning is a technique that aims to construct systems that can improve through experience by using advanced statistical and probabilistic techniques. It is believed to be a significantly useful tool to help in predicting mental health. It is allowing many researchers to acquire important information from the data, provide personalized experiences, and develop automated intelligent systems. The widely used algorithms in the field of machine learning such as support vector machine, random forest, and artificial neural networks have been utilized to forecast and categorize the future event

3.THEORITICAL ANALYSIS

3.1 Block diagram



3.2 Hardware / Software Requirements

Recommended System Requirements

- Processors: Intel® Core™ i5 processor 4300M at 2.60 GHz or 2.59 GHz (1 socket, 2 cores, 2 threads per core), 8 GB of DRAM Intel® Xeon® processor E5-2698 v3 at 2.30 GHz (2 sockets, 16 cores each, 1 thread per core), 64 GB of DRAM Intel® Xeon Phi™ processor 7210 at 1.30 GHz (1 socket, 64 cores, 4 threads per core), 32 GB of DRAM, 16 GB of MCDRAM (flat mode enabled)
- Disk space: 2 to 3 GB
- Operating systems: Windows® 10, macOS*, and Linux*

Minimum System Requirements

- Processors: Intel Atom® processor or Intel® Core™ i3 processor
- Disk space: 1 GB
- Operating systems: Windows* 7 or later, macOS, and Linux
- Python* versions: 3.9

Software requirements:

Anaconda navigator:

Anaconda is an open-source distribution for python and R. It is used for data science, machine learning, deep learning, etc. With the availability of more than 300 libraries for data science, it becomes fairly optimal for any programmer to work on anaconda for data science

Pycharm:

PyCharm is a dedicated Python Integrated Development Environment (IDE) providing a wide range of essential tools for Python developers, tightly integrated to create a convenient environment for productive Python, web, and data science development.

4.EXPERIMENTAL INVESTIGATIONS

Logistic Regression:

This type of statistical model (also known as logit model) is often used for classification and predictive analytics. Logistic regression estimates the probability of an event occurring, such as voted or didn't vote, based on a given dataset of independent variables. Since the outcome is a probability, the dependent variable is bounded between 0 and 1. In logistic regression, a logit transformation is applied on the odds—that is, the probability of success divided by the probability of failure.

DecisionTreeClassifier:

Decision Tree is a white box type of ML algorithm. It shares internal decision-making logic, which is not available in the black box type of algorithms such as Neural Network. Its training time is faster compared to the neural network algorithm. The time complexity of decision trees is a function of the number of records and number of attributes in the given data. The decision tree is a distribution-free or non-parametric method, which does not depend upon probability distribution assumptions. Decision trees can handle high dimensional data with good accuracy.

KNeighborsClassifier:

K-Nearest Neighbors, or KNN for short, is one of the simplest machine learning algorithms and is used in a wide array of institutions. KNN is a non-parametric, lazy learning algorithm. When we say a technique is non-parametric, it means that it does not make any assumptions about the underlying data. In other words, it makes its selection based off of the proximity to other data points regardless of what feature the numerical values represent. Being a lazy learning algorithm implies that there is little. Therefore, we can immediately classify new data points as they present themselves. to no training phase.

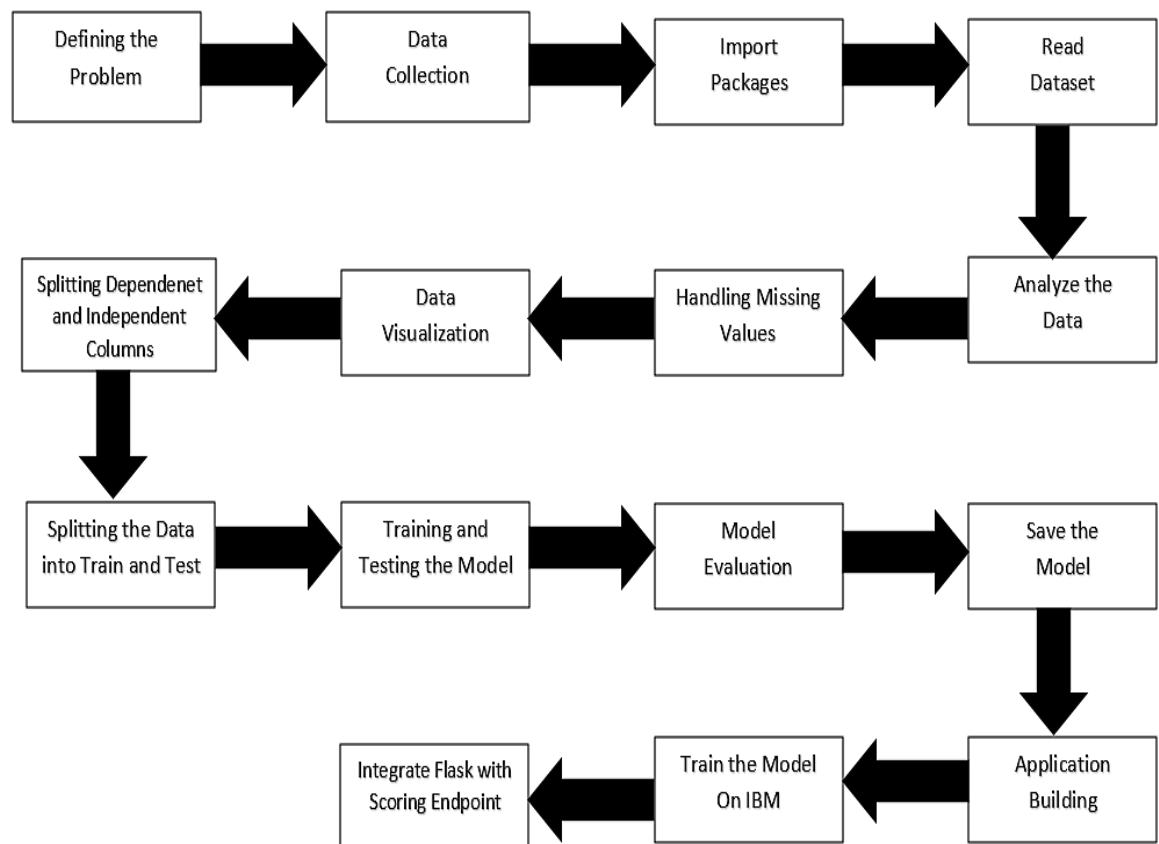
XGBClassifier:

The XGBoost stands for eXtreme Gradient Boosting, which is a boosting algorithm based on gradient boosted decision trees algorithm. XGBoost applies a better regularization technique to reduce overfitting, and it is one of the differences from the gradient boosting. The 'xgboost' is an open-source library that provides machine learning algorithms under the gradient boosting methods. The xgboost.XGBClassifier is a scikit-learn API compatible class for classification.

RandomForestClassifier:

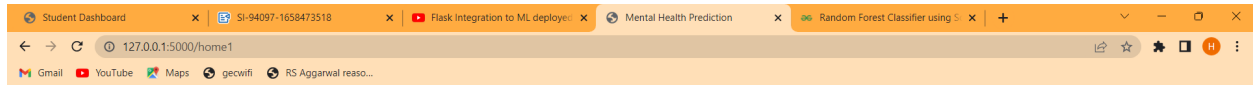
The Random forest classifier creates a set of decision trees from a randomly selected subset of the training set. It is basically a set of decision trees (DT) from a randomly selected subset of the training set and then It collects the votes from different decision trees to decide the final prediction.

5. FLOWCHART



6.RESULT

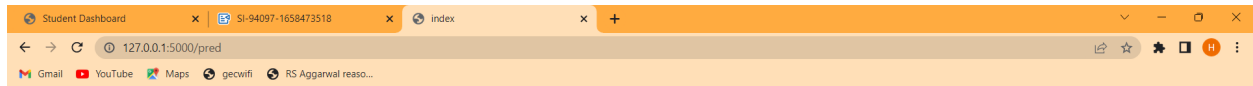
This study identified five machine learning techniques i.e k nearest neighbor classifier, logistic regression, decision tree, and stacking, and random forest. And we assessed their accuracy in identifying mental health issues.



Mental Health Prediction

Mental health affects your emotional, psychological and social well-being. It affects how we think, feel, and act. It also helps determine how we handle stress, relate to others, and make choices. In the workplace, communication and inclusion are key skills for successful high performing teams or employees. The impact of mental health on an organization can mean an increase of absent days from work and a decrease in productivity and engagement. In the United States, approximately 70% of adults with depression are in the workforce. Employees with depression will miss an estimated 35 million workdays a year due to mental illness.

Proceed



Prediction Page

1. What is your age :

2. What is your gender?

self_employed

family_history

work_interfere

No. of employees :

remote_work

tech_company

benefits

care_options



Student Dashboard

SI-94097-1658473518

Flask Integration to ML deployer

index

Random Forest Classifier using S

127.0.0.1:5000/pred

GmailYouTubeMapsgecwifIRS Aggarwal reaso...

leave
Somewhat easy

mental health consequence
Yes

Do you think that discussing a physical health issue with your employer would have negative consequence?
Yes

Would you be willing to discuss a mental health issue with your coworkers?
Yes

Would you be willing to discuss a mental health issue with your direct supervisor(s)?
Yes

Would you bring up a mental health issue with a potential employer in an interview?
Maybe

Would you bring up a physical health issue with a potential employer in an interview?
Maybe

Do you feel that your employer takes mental health as seriously as physical health?
Don't know

obs consequence
Yes

Predict

29°C Mostly cloudy

Windows Taskbar

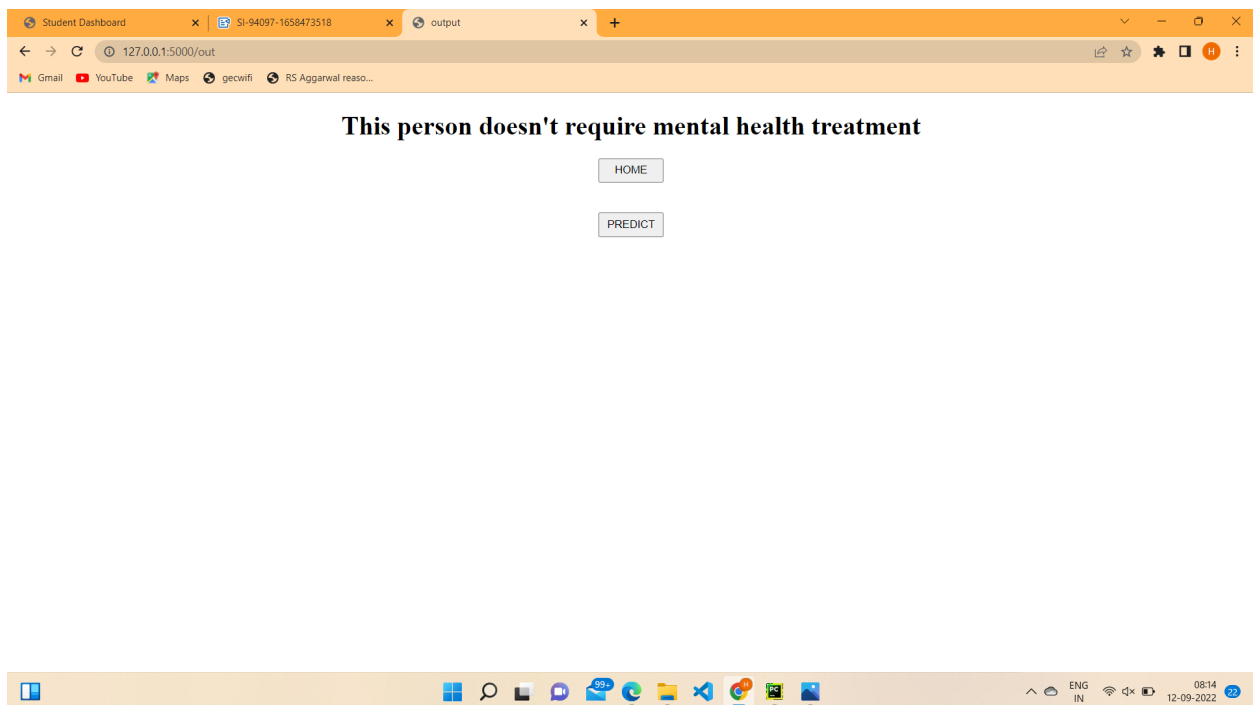
20:26 11-09-2022

This person require mental health treatment

HOME

PREDICT

After changing the Prediction values we can obtain the output based on input.



7.ADVANTAGES AND DISADVANTAGES

Advantages:

There is an endless number of advantages of Machine learning. We can take a look at the one which are really helpful.

The advantages of Machine learning tell us how using Machine learning would benefit us. So, let's have a look at the advantages of Machine Learning.

- >Automation of Everything
- >Wide Range of Applications
- >Scope of Improvement
- >Efficient Handling of Data
- >Best for Mental Health

Disadvantages:

Similar to the advantages of Machine Learning, we should also know the disadvantages of Machine Learning. If you don't know the cons, you won't know the risks of Machine Learning. So, let's have a look at these disadvantages.

- >Possibility of High Error
- >Algorithm Selection
- >Data Acquisition
- >Time and Space
- >Internet Issues

8.CONCLUSION

Many different techniques and algorithms had been introduced and proposed to test and solve the mental health problems. There are still many solutions that can be refined. In addition, there are still many problems to be discovered and tested using a wide variety of settings in machine learning for the mental health domain. As classifying the mental health data is generally a very challenging problem, the features used in the machine learning algorithms will significantly affect the performance of the classification.

The existing studies and research show that machine learning can be a useful tool in helping understand psychiatric disorders. Besides that, it may also help distinguish and classify the mental health problems among patients for further treatment. Newer approaches that use data that arise from the integration of various sensor modalities present in technologically advanced devices have proven to be a convenient resource to recognize the mood state and responses from patients among others.

It is noticeable that most of the research and studies are still struggling to validate the results because of insufficiency of acceptable validated evidence, especially from the external sources. Besides that, most of the machine learning might not have the same performance across all the problems. The performance of the machine learning models will vary depending on the data samples obtained and the features of the data. Moreover, machine learning models can also be affected by preprocessing activities such as data cleaning and parameter tuning in order to achieve optimal results.

Hence, it is very important for researchers to investigate and analyze the data with various machine learning algorithms to choose the highest accuracy among the machine learning algorithms [62]. Not only that, challenges and limitations faced by the researchers need to be managed with proper care to achieve satisfactory results that could improve the clinical practice and decision-making.

9.FUTURE SCOPE

We can extend the Mental Health Prediction Project by using different constraints. Various machine learning approaches can be implemented to predict or detect a disease at its early stages so that the treatment for it would be less complex and it would increase the probability of the patient being cured. As a result of these approaches, different types of disease can be detected but with diverse accuracy levels depending on factors such as the used algorithm, feature set, training dataset, and so on.

10.BIBLIOGRAPHY

Installation of Anaconda Navigator:

<https://www.youtube.com/embed/5mDYijMfSzs>

Installation Of Pycharm Professionals:

<https://www.youtube.com/embed/z73PyNDgVyQ>

Installation Of Python Packages:

https://www.youtube.com/embed/akj3_wTpIoU

Data Collection:

<https://www.kaggle.com/datasets/rishal005/admission-predict>

Data Pre-processing:

<https://thesmartbridge.com/documents/spsaimldocs/Datapreprocessing.pdf>

Handling Null Values:

<https://towardsdatascience.com/7-ways-to-handle-missing-values-in-machine-learning-1a6326adf79e>

Data Visualization:

<https://www.youtube.com/embed/TLdXM0A7SR8>

Splitting Dependent And Independent Columns:

https://www.youtube.com/embed/A_V6daPQZIU

Splitting The Data Into Train And Test:

<https://www.youtube.com/embed/xgDs0scjuuQ>

Training And Testing The Model:

<https://www.youtube.com/embed/yIYKR4sgzI8>

Model Evaluation:

<https://towardsdatascience.com/the-5-classification-evaluation-metrics-you-must-know-aa97784ff226>

Flask Frame Work Reference:

https://www.youtube.com/embed/lj4l_CvBnt0

Flask Refarance To Run:

<https://www.youtube.com/embed/UbCWoMf80PY>

Train The Model On IBM:

Account Creation:

https://www.youtube.com/embed/4y_zD-0Q3F8

Train Model On IBM Watson:

<https://www.youtube.com/embed/TysuP3KgSzc>

Integrate Flask With Scoring Endpoint:

<https://www.youtube.com/embed/ST1ZYLmYw2U>