SESHADRI RAO GUDLAVALLERU
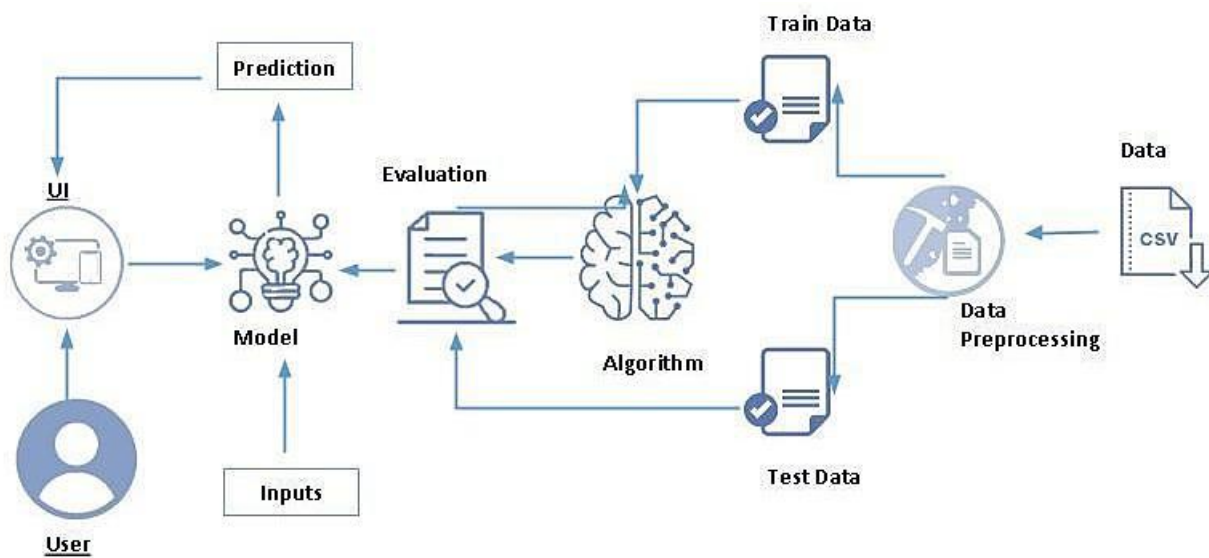
ENGINEERING COLLEGE

Mental Health Predicfion
Using IBM Watson

Submited By:

s.saichran
(19481A0272)

Content

1. **INTRODUCTION**

   a. Abstract

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

   b. Overview

   A brief descripfion about your project

2. **LITERATURE SURVEY**

   a. Exisfing problem

   Exisfing approaches or method to solve this problem

   b. Proposed solufion

   What is the method or solufion suggested by you?

3. **THEORITICAL ANALYSIS**

   a. Block diagram

   Diagrammafic overview of the project.

   b. Hardware / Sohware requirements

   Hardware and sohware requirements of the project

4. **EXPERIMENTAL INVESTIGATIONS**

   Analysis or the invesfigafion made while working on the solufion

5. **FLOWCHART**

   Diagram showing the control flow of the solufion

6. **RESULT**

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

7. **ADVANTAGES & DISADVANTAGES**

   List of advantages and disadvantages of the proposed solufion

8. **CONCLUSION**

   Conclusion summarizing the enfire 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 effecfive medical health care have led to an invesfigafion of machine learning that can be applied in mental health problems. This report presents a recent systemafic review of machine learning and deep learning approaches in predicfing mental health problems. Furthermore, we will discuss the challenges, limitafions, and future direcfions for the applicafion of machine learning in the mental health field. We collect research arficles and studies that are related in predicfing mental health problems by searching reliable databases. Then, we categorize the collected research arficles based on the mental health problems such as schizophrenia, bipolar disorder, anxiety and depression, postraumafic stress disorder, and mental health problems among children. Discussing the findings, we reflect on the challenges and limitafions faced by the researchers on machine learning in mental health problems. Addifionally, we provide concrete recommendafions on the potenfial future research and development of applying machine learning and deep learning in the mental health field. The main purpose of the Mental Health Predicfion 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 parficipants how to nofice 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 fime off policies. Organizafions that incorporate mental health awareness help to create a healthy and producfive work environment that reduces the sfigma associated with mental illness, increases the organizafions' 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 Predicfion 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 classificafion algorithms such as Logisfic 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.
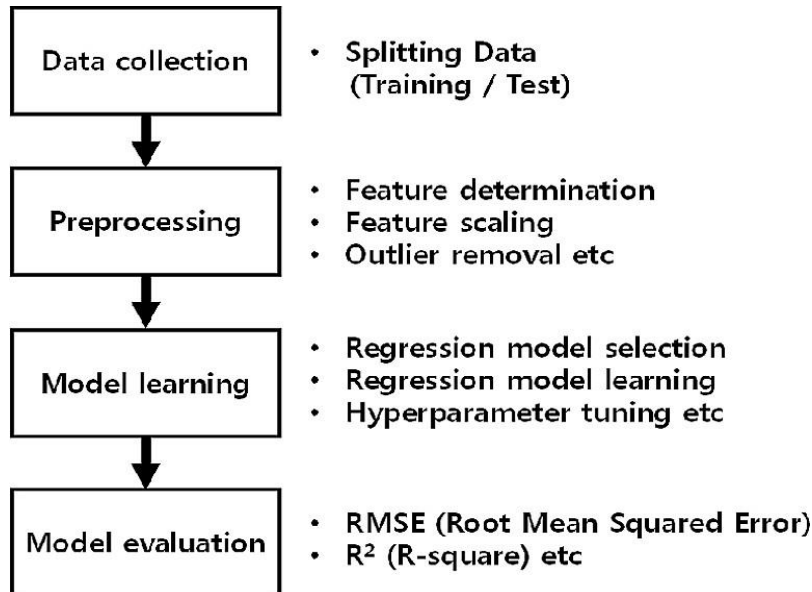
2.LITERATURE SURVEY

2.1 Exisfing problem

Mental illness is a health problem that undoubtedly impacts emofions, reasoning, and social interacfion of a person. These issues have shown that mental illness gives serious consequences across sociefies and demands new strategies for prevenfion and intervenfion. To accomplish these strategies, early detecfion of mental health is an essenfial procedure. Mental illness is usually diagnosed based on the individual self-report that requires quesfionnaires designed for the detecfion of the specific paterns of feeling or social interacfions. With proper care and treatment, many individuals will hopefully be able to recover from mental illness or emofional disorder .

2.2 Proposed solufion

Machine learning is a technique that aims to construct systems that can improve through experience by using advanced stafisfical and probabilisfic techniques. It is believed to be a significantly useful tool to help in predicfing mental health. It is allowing many researchers to acquire important informafion 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 arfificial neural networks have been ufilized to forecast and categorize the future event

3. THEORITICAL ANALYSIS

3.1 Block diagram

```
┌──────────────────┐      • Splitting Data
│ Data collection  │        (Training / Test)
└──────────────────┘
         │
         ▼
┌──────────────────┐      • Feature determination
│  Preprocessing   │      • Feature scaling
└──────────────────┘      • Outlier removal etc
         │
         ▼
┌──────────────────┐      • Regression model selection
│  Model learning  │      • Regression model learning
└──────────────────┘      • Hyperparameter tuning etc
         │
         ▼
┌──────────────────┐      • RMSE (Root Mean Squared Error)
│ Model evaluation │      • R² (R-square) etc
└──────────────────┘
```

3.2 Hardware / Sohware 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 DRAMIntel® Xeon® processor E5-2698 v3 at 2.30 GHz (2 sockets, 16 cores each, 1 thread per core), 64 GB of DRAMIntel® 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
- Operafing systems: Windows® 10, macOS*, and Linux*

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

Sohware requirements:
Anaconda navigator:
Anaconda is an open-source distribufion 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 opfimal 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 essenfial tools for Python developers, fightly integrated to create a convenient environment for producfive Python, web, and data science development.

4.   EXPERIMENTAL INVESTIGATIONS

Logisfic Regression:

This type of stafisfical model (also known as logit model) is ohen used for classificafion and predicfive analyfics. Logisfic regression esfimates 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 logisfic regression, a logit transformafion 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 fime is faster compared to the neural network algorithm. The fime complexity of decision trees is a funcfion of the number of records and number of atributes in the given data. The decision tree is a distribufion-free or non-parametric method, which does not depend upon probability distribufion assumpfions. 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.
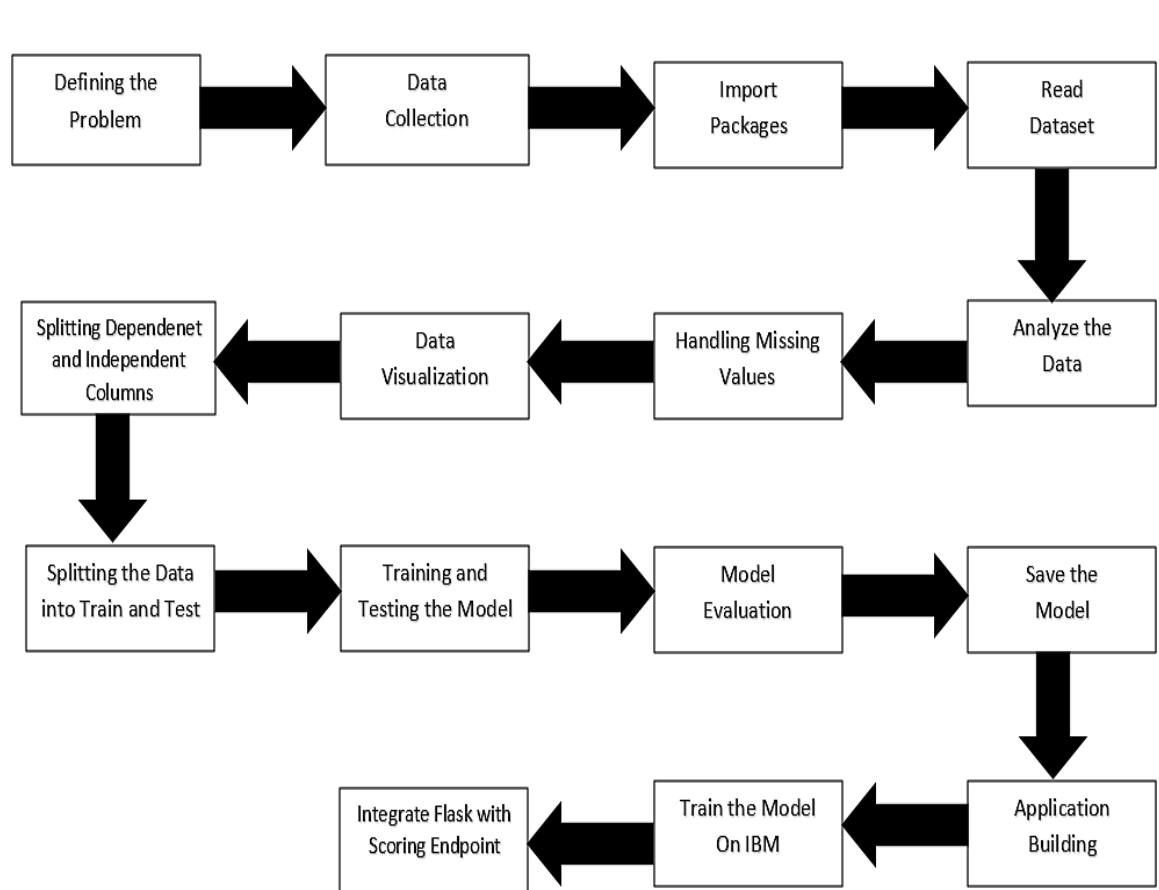
XGBClassifier:

The XGBoost stands for eXtreme Gradient Boosfing, which is a boosfing algorithm based on gradient boosted decision trees algorithm. XGBoost applies a beter regularizafion technique to reduce overfifing, and it is one of the differences from the gradient boosfing.The 'xgboost' is an open-source library that provides machine learning algorithms under the gradient boosfing methods. The xgboost.XGBClassifier is a scikit-learn API compafible class for classificafion.
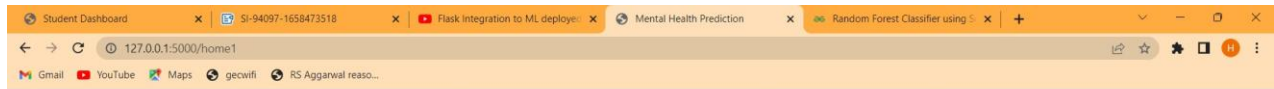
RandomForestClassifier:

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.
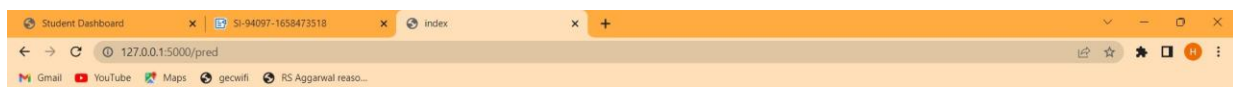
5. FLOWCHART

```
Defining the        →    Data           →    Import         →    Read
Problem                  Collection          Packages            Dataset
                                                                   ↓
Splitting Dependenet ←   Data           ←    Handling Missing ←  Analyze the
and Independent          Visualization       Values              Data
Columns
    ↓
Splitting the Data  →    Training and   →    Model          →    Save the
into Train and Test      Testing the Model   Evaluation          Model
                                                                   ↓
Integrate Flask with ←   Train the Model ←   Application
Scoring Endpoint         On IBM              Building
```

6.   RESULT

C   127.0.0.1:5000/pred

M   Gmail   D   YouTube   Maps O   0 RS Aggarwal reaso...

leave

Somewhat easy ▾

Do you think that discussing a physic a1 health issiJe with your employer would 4aave negative consequence

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

Yes ▾

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?
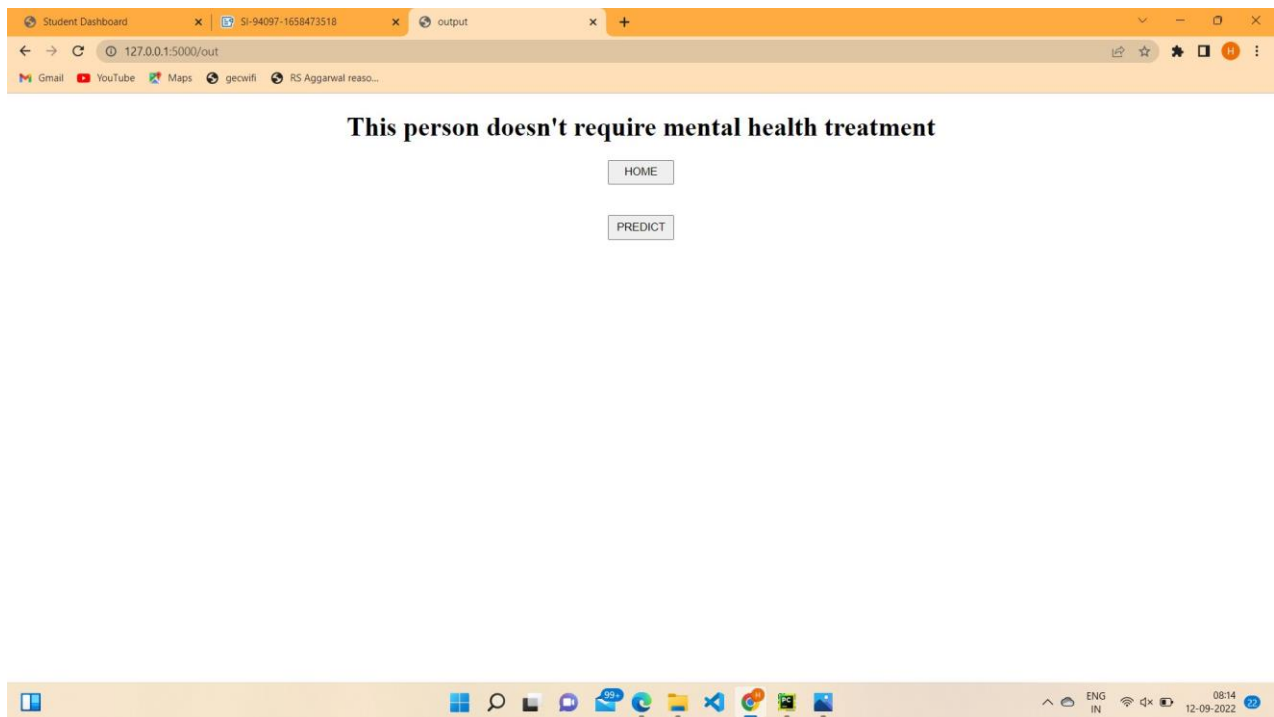
Don't know ▾

obs consequence

Yes ▾

Predict

C O   127.0.0.1:5000/out

M   Gmail   D   YouTube   Maps O   0 RS Aggarwal reaso...

This person require mental health treatment

Aher changing the Predicfion 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.

  -->Automafion of Everything

  -->Wide Range of Applicafions

  -->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 Selecfion

  -->Data Acquisifion

  -->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 sfill many solufions that can be refined. In addifion, there are sfill many problems to be discovered and tested using a wide variety of sefings 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 classificafion.

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

It is noficeable that most of the research and studies are sfill 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 acfivifies such as data cleaning and parameter tuning in order to achieve opfimal results.

Hence, it is very important for researchers to invesfigate 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 limitafions faced by the researchers need to be managed with proper care to achieve safisfactory results that could improve the clinical pracfice and decision-making.

9.   FUTURE SCOPE

We can extend the Mental Health Predicfion Project by using different constraints. Various machine learning approaches can been 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 been detected but with diverse accuracy levels depending on factors such as the used algorithm, feature set, training dataset, and so on.

10.   BIBILOGRAPHY

Installafion of Anaconda Navigator:

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

Installafion Of Pycharm Professionals:

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

Installafion Of Python Packages:

      https://www.youtube.com/embed/akj3_wTploU

Data Collecfion:

      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 Visualizafion:

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

Splifing Dependent And Independent Columns:

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

Splifing The Data Into Train And Test:

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

Training And Tesfing The Model:

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

Model Evaluation:

h tps://towardsdatascience.com/the-5-classification-evaluation-metrics-you-must-know-aa97784ff226

Flask Frame Work Reference:

https://www.youtube.com/embed/lj4I_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