EMPLOYEE MENTAL HEALTH PREDICTION USING MACHINE LEARNING

Name: Amulya Muthoju

Date: 21/11/2023

Project-3

Step 1: Prototype Selection

Abstract:

People whose occupation is to work all day sitting at a desk has a high possibility of having long-term serious health-related consequences due to the sedentary lifestyle they spend. 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. Medical predictive analytics will reform the healthcare field broadly. 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.

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 events.

The main objective is to provide a systematic review, critical review, and summary of the machine learning techniques that are being used to predict, diagnose, and identify mental health problems. It would also give attention to the challenges and limitations of applying the machine learning techniques in this area.

1. Problem Statement:

Mental health in the workplace is a critical aspect of employee well-being and organizational success. Recognizing and addressing mental health issues early can lead to improved employee satisfaction, productivity, and overall workplace atmosphere. The aim is to develop a machine learning model that predicts the mental health status of employees based on various factors.

The goal is to create a predictive model that can identify potential mental health issues among employees. By analyzing relevant features, the model should be able to provide insights into the mental health status of individuals, enabling organizations to take proactive measures to support their employees.
This has the potential to revolutionize the way organizations approach employee well-being by providing a data-driven method to identify and address mental health issues. The predictive model can serve as an early warning system, allowing companies to implement targeted interventions and support systems for their employees, ultimately fostering a healthier and more productive work environment.

2. Market/customer/Business Need Assessment

Creating the conditions for employee mental health and wellbeing is the right thing to do for many reasons, but among them are the significant business benefits. In addition to the advantages for people, the paybacks for organizations and economies are undeniable. Improved employee mental health is good for business—fully eight scientific studies prove it—from the University of California to the University of South Australia and published in journals of economic studies to occupational studies.

Good business requires balancing multiple issues—continuity over time balanced with adaptability as well as clear and compelling leadership direction balanced with participation and input from employees. And perhaps most of all, strategic choices about the right investments for payback to business, people, customers and communities.

Organizations can't invest in everything—so it's reasonable to assess where to put time, energy and resources. And the evidence is clear that programs, processes and systems which support both physical health and mental health are very good business.

But beyond the overall economic impact, organizations can make strategic investments in mental health and wellbeing which impact their outcomes.

3. Target Specifications and characterization

Understanding who you are targeting with your app is the key to developing a mental health application because every age group has different mental states. For instance, a young adult going through a heartbreak might be more shattered than an adult going through the same thing. This is because of the different maturity levels and experiences of both age groups.

3.1 Organizations

Investing in the mental health of employees is a sound business decision, leading to increased productivity and business outcomes. Many employers want to do more but need guidance on how to do so. Mental health in the workplace has often become synonymous with free yoga

classes, fruit boxes and awareness morning teas. While these are simple to implement and high profile, the evidence shows they have little material impact on mental health outcomes.
The available evidence suggests that the best returns are from businesses focusing on prevention and early intervention. Based on evidence and consultation with experts, we propose a framework for organizations that focuses on building strong foundations – through good job design, strong management capability and a supportive organizational culture. This will assist in preventing mental ill health before it becomes a problem. It requires appropriate support for all in the organization to thrive, and individualized support for those that need it.

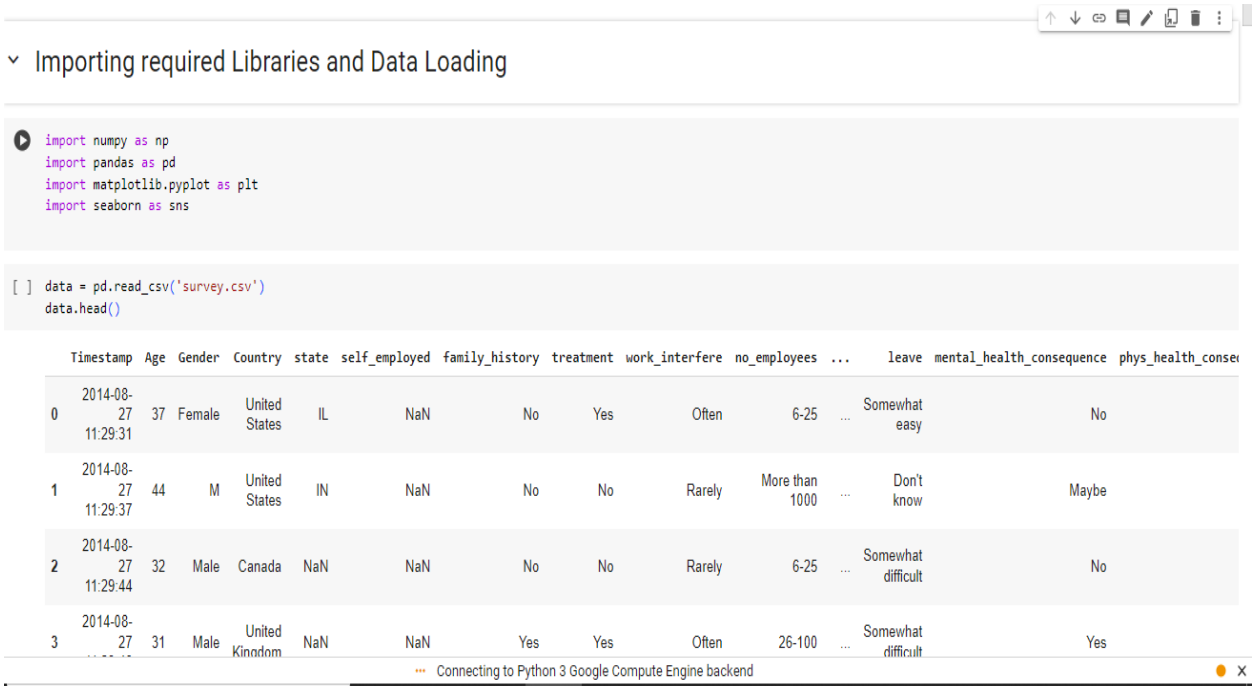
Where can businesses start to improve employee wellbeing;

- Prioritize mental health as much as physical health and safety in the workplace;
- Undertake an assessment of where their organization is currently sitting with mental health, and implement ongoing monitoring; and
- Invest in the management capability of staff to identify and support workers with mental ill-health.

4. External Search

https://www.researchgate.net/
https://www.jetir.org/papers/JETIR2306614
https://www.corporatewellnessmagazine.com/article/revolutionizing-employee-wellness-how-ai-and-machine-learning-are-making-a-difference
https://www.vantagefit.io/blog/employee-wellness-apps/

Data collection



Importing required Libraries and Data Loading

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
[ ] data = pd.read_csv('survey.csv')
    data.head()
```

	Timestamp	Age	Gender	Country	state	self_employed	family_history	treatment	work_interfere	no_employees	...	leave	mental_health_consequence	phys_health_consec
0	2014-08-27 11:29:31	37	Female	United States	IL	NaN	No	Yes	Often	6-25	...	Somewhat easy	No	
1	2014-08-27 11:29:37	44	M	United States	IN	NaN	No	No	Rarely	More than 1000	...	Don't know	Maybe	
2	2014-08-27 11:29:44	32	Male	Canada	NaN	NaN	No	No	Rarely	6-25	...	Somewhat difficult	No	
3	2014-08-27	31	Male	United Kingdom	NaN	NaN	Yes	Yes	Often	26-100	...	Somewhat difficult	Yes	

Connecting to Python 3 Google Compute Engine backend

Let's see some basic information about data

```
print(data.shape)
print(data.describe())
print(data.info())
```

```
(1259, 27)
                Age
count  1.259000e+03
mean   7.942815e+07
std    2.818299e+09
min   -1.726000e+03
25%    2.700000e+01
50%    3.100000e+01
75%    3.600000e+01
max    1.000000e+11
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1259 entries, 0 to 1258
Data columns (total 27 columns):
 #   Column                     Non-Null Count  Dtype
---  ------                     --------------  -----
 0   Timestamp                  1259 non-null   object
 1   Age                        1259 non-null   int64
 2   Gender                     1259 non-null   object
 3   Country                    1259 non-null   object
 4   state                      744 non-null    object
 5   self_employed              1241 non-null   object
 6   family_history             1259 non-null   object
 7   treatment                  1259 non-null   object
 8   work_interfere             995 non-null    object
 9   no_employees               1259 non-null   object
 10  remote_work                1259 non-null   object
 11  tech_company               1259 non-null   object
 12  benefits                   1259 non-null   object
 13  care_options               1259 non-null   object
 14  wellness_program           1259 non-null   object
 15  seek_help                  1259 non-null   object
 16  anonymity                  1259 non-null   object
 17  leave                      1259 non-null   object
 18  mental_health_consequence  1259 non-null   object
 19  phys_health_consequence    1259 non-null   object
 20  coworkers                  1259 non-null   object
 21  supervisor                 1259 non-null   object
 22  mental_health_interview    1259 non-null   object
 23  phys_health_interview      1259 non-null   object
 24  mental_vs_physical         1259 non-null   object
 25  obs_consequence            1259 non-null   object
 26  comments                   164 non-null    object
dtypes: int64(1), object(26)
memory usage: 265.7+ KB
None
```

5. Bench marking alternate products

Recent years have seen an exponential development of mobile technologies aimed at improving various mental health problems. Such technologies are considered part of a new field of medicine called mobile health (*mHealth*). This term refers to health (including mental health) supported by mobile technologies. The mHealth field is booming, with a plethora of health-related apps, websites, and text messaging–support interventions being developed by the industry and being adopted by the public. However, only a small proportion of these technologies have undergone any form of empirical assessment.

This lack of app validation is a concern, even more so when studies suggest that mental health– and addiction-related apps currently available to the public, with few exceptions, offer insufficient content quality. Fortunately, recent years have seen an increase in the gathering of empirical data related to smartphone app–related interventions.

6. Applicable Regulations:

- Data Privacy Laws
- Informed Consent
- Transparency and Explainability
- Medical Device Regulations
- Ethical Guidelines
- Review and Approval Processes

7. Applicable Constraints

Developing and implementing a mental health prediction system using machine learning involves several considerations, including the need for space, budget, and expertise. Here's an overview of each aspect:

A. Space:

The space requirements for a machine learning project, including one focused on mental health prediction, can vary based on the scale and complexity of the project. If you're developing and training machine learning models, you'll need sufficient computing resources. This could range from a dedicated workstation for small-scale projects to a more extensive computing cluster for larger and more complex models. Cloud computing services, such as AWS, Google Cloud, or Azure, can be utilized to access scalable resources without the need for physical infrastructure

B. Budget:

The budget for a mental health prediction using machine learning can cover various aspects:

- **Hardware and software:** The cost of acquiring or renting hardware, GPUs (Graphics Processing Units) for model training, and relevant software licenses.
- **Data Collection and Storage:** Expenses related to obtaining and storing the necessary datasets. This may include costs associated with data acquisition, data preprocessing and storage solutions.
- **Personnel:** Budget for hiring skilled professionals, such as data scientists, machine learning engineers, and domain experts in mental health.
- **Training and Skill Development:** Investments in training and skill development programs for the team members involved in the project.
- **Infrastructure:** If deploying models in a production environment, costs associated with server infrastructure, cloud services, and maintenance.

C. Expertise:

Building a mental health prediction system using machine learning requires a multidisciplinary team with expertise in several areas:

- **Data Science and machine learning:** Individuals with expertise in data analysis, feature engineering, model selection, and evaluation.
- **Domain Knowledge:** Professionals with a deep understanding of mental health, psychology, or psychiatry are crucial for developing models that are clinically relevant and sensitive to the nuances of mental health data.
- **Software Development:** Skilled software engineers who can implement and integrate machine learning models into applications or systems.
- **Ethics and Privacy:** Experts who understand the ethical implications of working with sensitive health data and can ensure compliance with relevant regulations.

8. Business model

Employees need, and increasingly demand, resources to help them cope with mental health problems. If companies make mental health services more accessible and intervene in the workplace in ways that improve well-being, they will simultaneously make investments that will provide real improvements in employee outcomes and consequently in company performance.

More large employers are providing onsite medical care. Providing care onsite cuts out employee travel time and can save costs. Company-paid doctors are often less expensive than fee-for-service arrangements or care provided in settings with large facility charges.

Now there is a growing movement to make behavioral health services available at the workplace as well. According to the Business Group on Health, one-third of employers with more than 5,000 employees said they would offer behavioral health counseling on-site in 2020, a big increase from the one-fifth that did so in 2018. Presumably, this trend—together with increased access to virtual care—will continue when employees return to traditional office settings post-COVID-19.

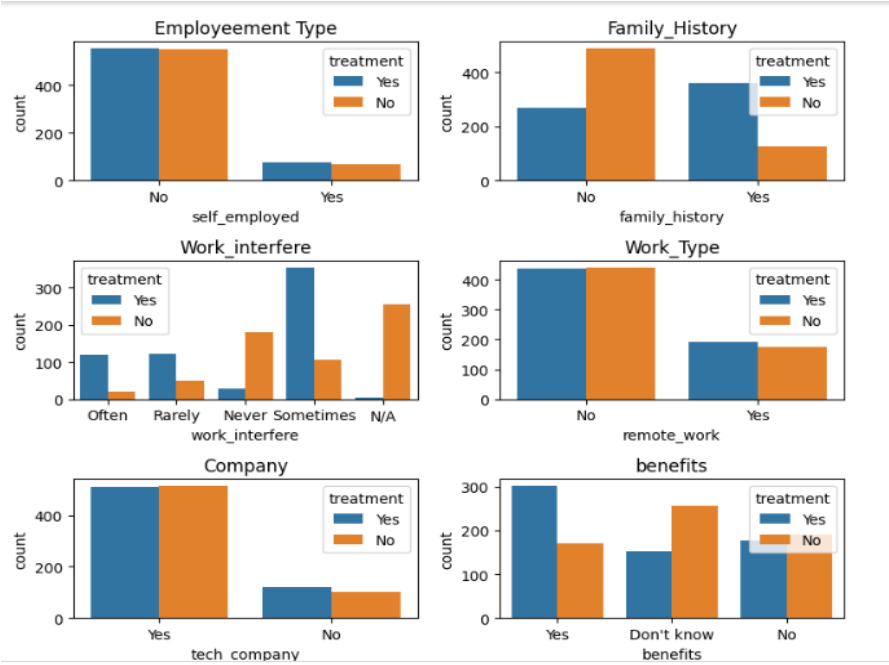
9. Concept generation

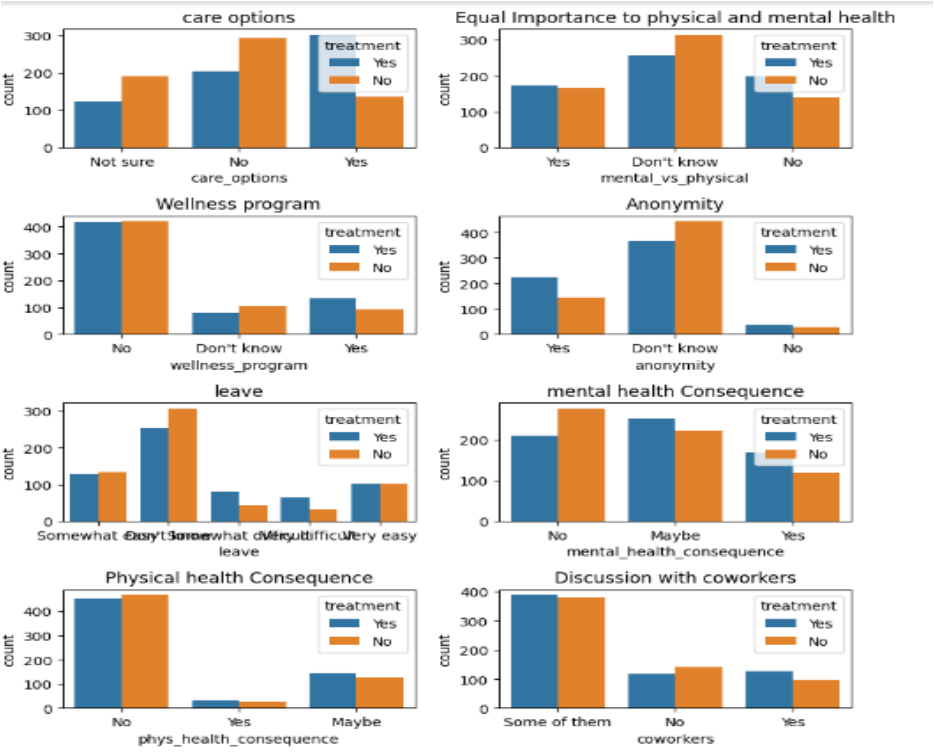
Technology is evolving round the clock in recent times. This has resulted in job opportunities for people all around the world. It comes with a hectic schedule that can be detrimental to people's mental health. So during the Covid-19 pandemic, mental health has been one of the most prominent issues, with stress, loneliness, and depression all on the rise over the last year. Diagnosing mental health is difficult because people aren't always willing to talk about their problems.

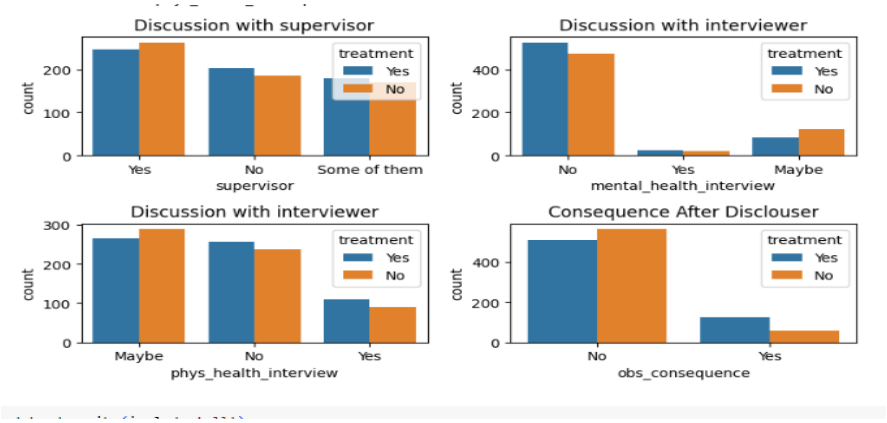
Machine learning is a branch of artificial intelligence that is mostly used nowadays. ML is becoming more capable for disease diagnosis and also provides a platform for doctors to analyze a large number of patient data and create personalized treatment according to the patient's medical situation.

In this, we are going to predict the mental health of Employees using various machine learning models.

Now let us look what features affect the mental health of employees and needs treatment.





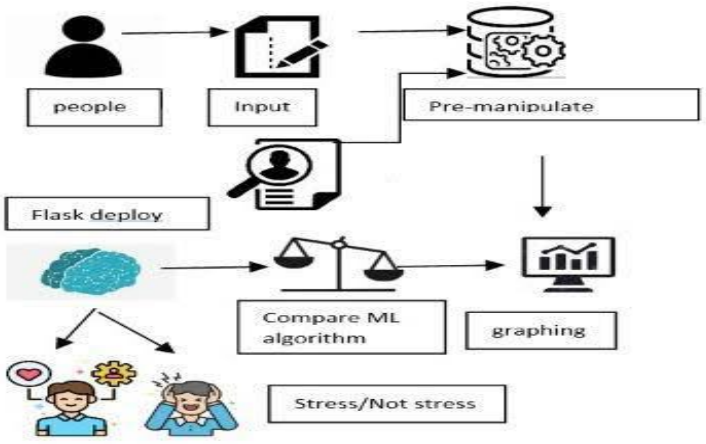


10. Concept Development

- Collect and prepare data
- Exploratory Data Analysis (EDA)
- Feature Engineering
- Split the Data
- Choose best Model
- Train and Evaluate the Model
- Hyperparameter Tuning
- Validate the Model
- Interpretability and Explainability
- Deploy the Model
- Monitor and Maintain

11. Final product prototype with schematic diagram:

The main scope of our project is to detect mental health Prediction and to design a classification model with a help of a machine learning algorithm.



12. Product details

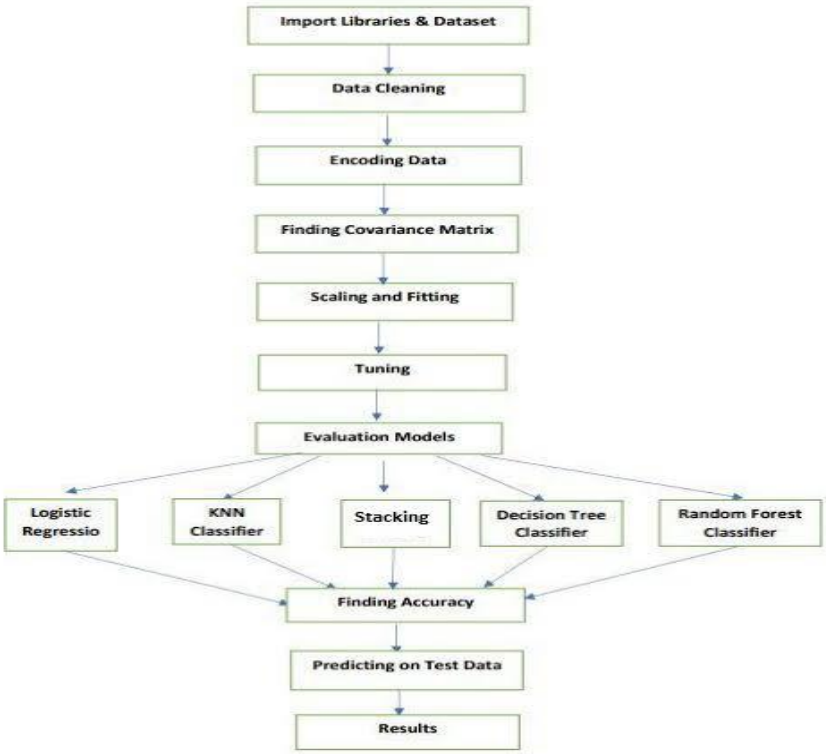
We have proposed an idea in which we collected datasets and preprocessed the data for our product and mapped the data into graphs to check the quality of the data then train the data with different algorithms to predict the output with higher accuracy.

This prediction also tells us that it is very important for IT employees to get a regular mental health check-up to track their health. The employers should have a medical service provided in their company and they should also give benefits to the affected employees There are many suggestions that employers and employees could keep in mind. Employers need to keep track of the number of their employees having mental disorders. Employers should allow a flexible work environment with flexible work scheduling and break timings. They should allow employees to work from home or have a flexible place of work.

12.1 How does it work

Includes Data Collection, Data Cleaning, encoding data, Finding Co-variance matrix, Scaling and Fitting, Tuning, Evaluation models, Finding Accuracy, Predicting Data and Results.

The accuracy of all the algorithm is compared and the best algorithm is deployed using flask into a webpage.



12.2 Data Sources

https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey

12.3 Algorithms, Frameworks, Software needed:

- This classification model is made up of four algorithms of Machine Learning such as Logistic Regression algorithm, Random Forest algorithm, Decision Tree Classifier algorithm, KNeighbors Classifier.
- Flask web framework written in Python.
- Jupyter notebooks is used as a software tool.

12.4 Team required to develop.

1. Machine learning engineering
2. Data scientists
3. Business analyst
4. Software developer
5. UI & UX developer
6. Test automation engineer

Step 2: Prototype Development

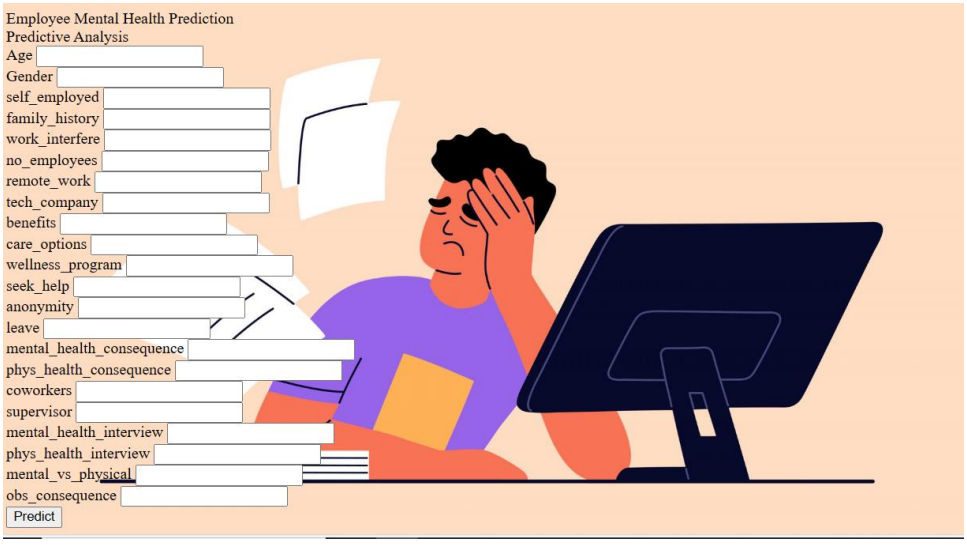
Code Implementation (Small Scale):

GitHub link: https://github.com/M-Amulya/Employee_Mental_Health_Prediction

Let's see how does it work

I have built a website which makes predictions on whether an employee needs a treatment or not due to work stress.

Fill the below form and click on predict button, our model makes predictions successfully.



Step 3: Business Modeling

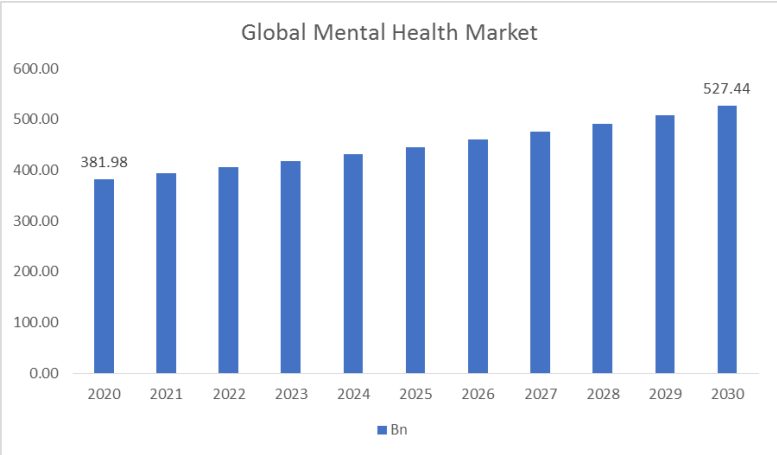
Employees are in need and are demanding better coverage. Here's how leading companies get mental health coverage right.

Business model for profit- making

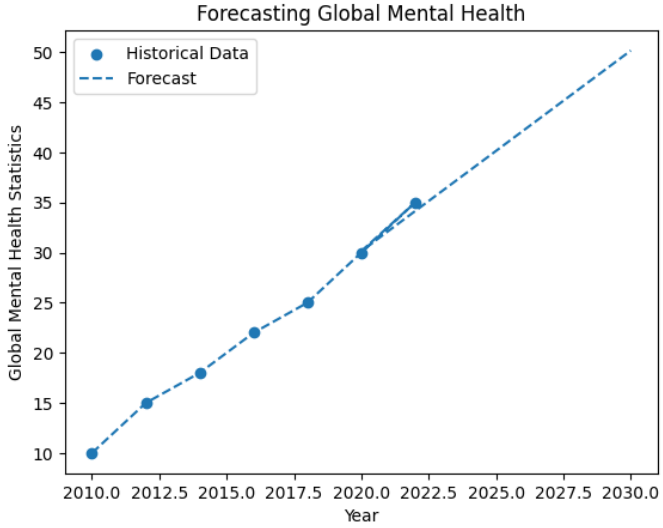
Employer/Corporate Wellness Model:

- Partnerships with employers to offer mental health prediction and support services as part of corporate wellness programs.
- Can include bulk subscriptions or enterprise-level agreements

Step 4: Financial Modeling



Forecasting with sample data



Global Mental health statistics = m * Year + b

In the code model.coef_[0] correspond to slope(m) and model.intercept_ correspond to y-intercept(b).

We got m=1.99, b=-4009.85

Global Mental health statistics = 1.99 * Year – 4009.85

This means that according to the linear regression model, for every additional year, the global mental health statistics are expected to increase by 1.99 units.

13. Conclusion:

There are many suggestions that employers and employees could keep in mind. Employers need to keep track of number of their employees having mental disorder. Employers should allow flexible work environment with flexible work scheduling and break timings. They should allow employees to work from home or have flexible place of work. They should give day-to-day feedback and guidance for nurturing employees' health. This type of model could be used to detect metal health progress among employees and also could lead to policy changes. Employees could talk to colleagues and their managers about their problem freely. Hence upper management could help them to get correct aid with beneficiaries like work from home, flexible timings, more leaves, many more. Employees should know health benefits provided by their organization participate in any wellness programs. Proper feedback should be provided to employee when they resign from their job. This could help them to improve their health.