Mental Health Prediction Using Machine Learning

Dr. S. N. ThirumalaRao,M.Tech,(Ph.D)[1], AdugulaChetanKumar[2], DuggiSharanKumar[3],ThotaNagaManiKanta[4]
[1] Professor, [2, 3 ,4]Student
[1]nagathirumalarao@gmail.com, [2]chetanxe@gmail.com, [3]duggesharankumar@gmail.com, [4]thotanagamanikanta01@gmail.com,
Department of Computer Science and Engineering,
Narasaraopeta Engineering College, Narasaraopet, Andhra Pradesh, India

Abstract: Mental illness is becoming more widely acknowledged as a significant health issue these days. Patients with mental health issues are becoming more prevalent. As a result, we must devise a strategy to address this issue. Creating machine learning models that can forecast users' mental illnesses—that is, those with poor mental health—is the primary objective of this research effort. Here, by using some of the machine learning models, we were able to provide some insight into this issue. We used datasets containing the individual mental health information to create our models. Pre-processing was also necessary because the dataset can include noisy or erroneous data. On the data, several machine learning methods are used.

KEYWORDS:**MentaHealth,SelfAnalysis,Machine Learning,LogisticRegression,DecisionTree,Radom ForestSupportVectorMachine,XGBoostClassifier, ADAboost, Gradient boost**

I. INTRODUCTION

Mental health is one of the leading causes to suicide. As most of the victims know a little about their mental health, it goes unseen. Every year more than 800,000 individuals die due to Depression, or some mental health disorder and this number is not a small Also, the one or which could be ignored [3]. Most of the individuals are between the age between 15-24.Which is mostly youth and have a lot of potential

Mostly youth and have a lot of potential to become something and give society a lot. To achieve success in such cases, the potential cases must be sorted before even they know. Mental health has a lot to do with the amount earned and the way it has been spent [1]. Our spending and liabilities tell a lot about our potential mental health. examining previous cases, in which it has considered how much individual do earn and there spending habits or where there most chunk of money goes, if they use tobacco, alcohol, medical expenses, social expenses, Education expenses, number of times they eat meat or fish, slept hungry or not, enough food to become something and give society a lot [4]. To achieve success in such cases, the potential cases must be sorted before even they know. Mental health has a lot to do with the amount earned and the way it has been spent [5]. Our spending and liabilities tell a lot about our potential mental health. examining previous cases, in which it has considered how much individual do earn and there spending habits or where there most chunk of money goes, if they use tobacco, alcohol, medical expenses, social expenses, Education expenses, number of times they eat meat or fish, slept hungry or not, enough food for tomorrow, Doctor consulting and other kind of liabilities.

Also, the family structure has been considered, in that how many members are there in family, age of the respondent, if the respondent is female, Marital status, number of children. Assets are also included, such as household size, number of farms, value of live stocks, value of durable goods, value of savings, food own production and many more. Being stress free is one of the key components for a healthy life [6].

Goal was to make a ML model, using which the future possibility of mental health can be identified. Using and analyzing all the data logical manner gave us insight that there are some similarities in all these cases [7]. Using which the mental health can be predicted and using the same model, the potential mental health victim can also be predicted.

II. LITERATURESURVEY

This study evaluated the practicality of many machine learning algorithms that classify the information into distinct mental health categories. This framework was developed to assess an individual's mental health status, and models for evaluation were constructed using this framework in mind.

Suresh, G et al. [8]A research investigation of machine learning and its potential connection to mental health difficulties has been prompted by the rise in mental health concerns and the need for effective medical care. This research offers a recent accurate as learning techniques for predicting mental health problems.

D. Graham, Sarah et at. [9] invented the Behavioral health disorders, specifically distress, are the kinds of health issues that many people are unaware of. One cannot possibly receive treatment for something they are unaware of. Therefore, identifying a person who may have a wellness problem is the first step in avoiding them. for that he used the advance machine learning models like Gradient Boosting Classifier (GBC), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), XGBoost (XGB)These studies demonstrate thevariety of methods and techniques used to predict an individual's mental illness. They also emphasize the importance of characteristics including age, gender, occupation, and social context. B this correctly projecting the current mental state.

Bhattacharyya et al. [10] tells that A person's enthusiastic, mental, and social well-being are all indicated by their mental health. It determines how a person thinks, feels, and responds to situations. A person's ability to work profitably and reach their maximum potential depends on their mental health. These remarks were made in 2018.

Shatter et al. [11] Predicting which kids will have detrimental effects on their mental health as teenagers is essential for early intervention and foreseeing severe consequences down the road. Despite the fact that many perspectives on a child's life, identity, and side effects have been praised as helpful, no show has yet been developed to test the general public for the risk of developing mental health difficulties as a result of these things.

III. PROPOSEDSYSTEM

Our Model is proposed based on certain criteria as follows.

A. System Setup
B. Dataset Collection
C. Pre-processing
D. Model Architecture
E. Data Visualization

A. System Setup

The system configuration consists of building and evaluating machine learning algorithms using Python 3.11 and the Scikit-learn module. Google Collaborator is the platform used for the training and testing phases. Scalability and seamless workflow integration are made possible for machine learning applications through the usage of Python, Scikit-learn, and Visual studio code.

B. Dataset Collection

```
Index(['age', 'gender', 'self_employed', 'family_history', 'treatment',
       'work_interfere', 'no_employees', 'remote_work', 'tech_company',
       'benefits', 'care_options', 'wellness_program', 'seek_help',
       'anonymity', 'leave', 'mental_health_consequence',
       'phys_health_consequence', 'coworkers', 'supervisor',
       'mental_health_interview', 'phys_health_interview',
       'mental_vs_physical', 'obs_consequence'],
      dtype='object')
```

Fig. 1: Dataset fields

We are taken the datasets from kagglewebsite[15] from internet, to make any prediction we need the previous datasets of the Mantel Health. We have collected dataset which is survey.csv (2014-2016) The dataset includes age, gender, self-employed, family-history, treatment, work-interface, no-employee, remote-work, tech-company, benefits', care-options, abnormality, leave, mental and physical consequences, co-workers and etc. Which is related to wellbeing's environment.

C. Pre-processing

In the pre-processing steps we are going to do some date manipulation operations on the taken dataset. Initially in data pre-processing we are going describe the dataset.
df. describe (): here we can see all statistical values.

	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

Fig.2: description

df.info (): Which is used to the see the attribute values are null or not null.

```
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
```

Fig.3: dataset info

After the Categorical string values will be transform into the numerical values the correlation of the each attribute will be represent in the correlation heatmap.

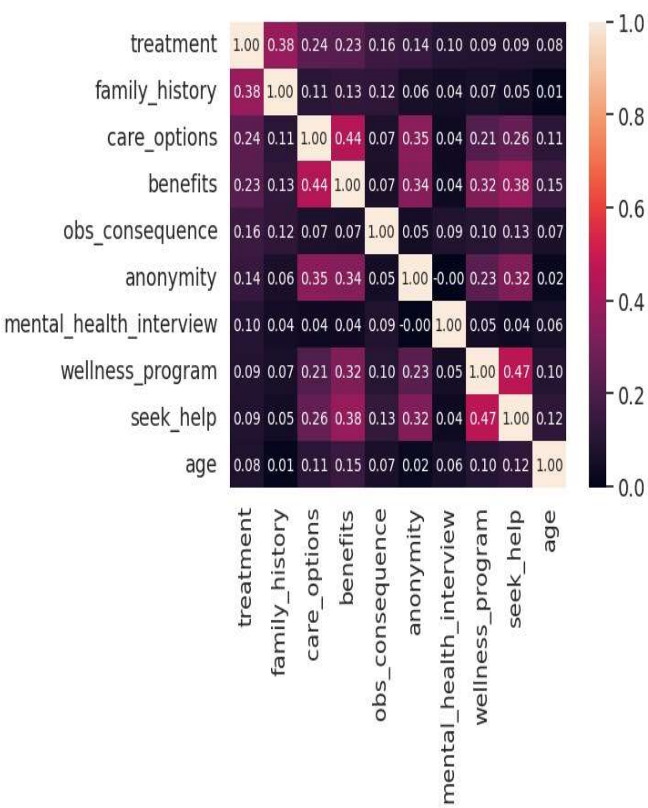

Fig.4: Correlation

D. Model Architecture

This phase involves selecting a suitable machine learning algorithm and training it on the preprocessed data. By fine-tuning its parameters to minimize the discrepancy between the expected and actual outputs in the training set, the model is trained. The model is tested on a different approval dataset to gauge its performance after training. The selection of evaluation features is contingent upon the type of conflicts and the performance standards. Accuracy, precision, cross-validation, and F1score are examples of common evaluation measures. By changing its settings or using various algorithms, the model can be further refined in light of the evaluation results. To improve the model's performance on fresh, untested data, this step is crucial. The model can be used to make predictions or choices in a production setting after it has been trained and assessed. This phase involves selecting a suitable machine learning algorithm and training it on the preprocessed data. By fine-tuning its parameters to minimize the discrepancy between the expected and actual outputs in the training set, the model is trained. The model is tested on a different approval dataset to gauge its performance after training. The selection of evaluation features is contingent upon the type of conflicts and the performance standards. Accuracy, precision, cross-validation, and F1score are examples of common evaluation measures. By changing its settings or using various algorithms, the model can be further refined in light of the evaluation results. To improve the model's performance on fresh, untested data, this step is crucial. The model can be used to make predictions or choices in a production setting after it has been trained and assessed.

DATA VISUALIZATION

A graphical representation, or visual representation of the dataset, and making diagnostics through this make east to the users. is created to facilitate a quick comprehension and visualization of the characteristics found in the dataset that were specifically utilized to support human mental health.

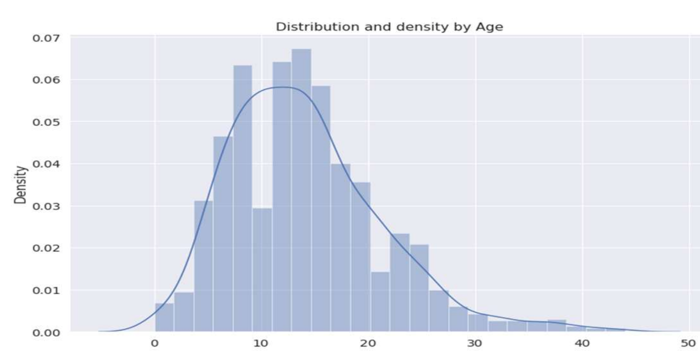

Fig. 5: density graph

The Age column in relation to density is displayed in the plot above. In our dataset, we can see that the density increases between the ages of 10 and 20.

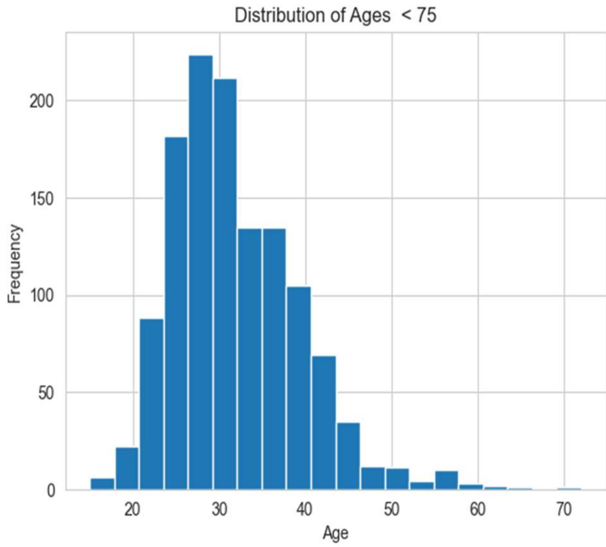

Fig. 6: Bar graph

Bar graphs for data visualization are essential to the "mental health prediction using Machine Learning Algorithms" project because they provide various aspects of the dataset and model performance.

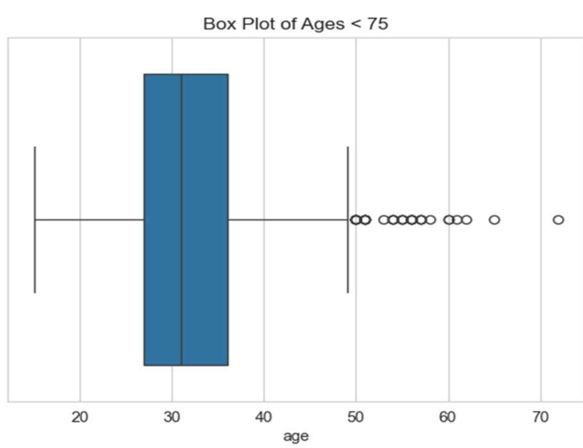

Fig. 7: box plot

Box plots are a useful tool for displaying the collection of numerical data values in between the different age groups, particularly when compare them across several groups.

sns. heatmap: makes a heatmap to show where in the dataset there are missing values. It is simple to identify patterns of missingness across several attributes with this display.

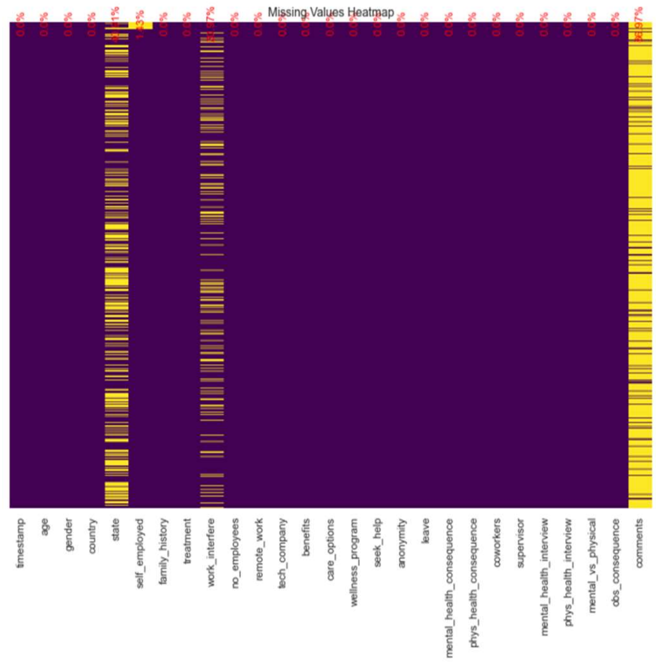

Fig 8: Heat map

IV. RESULT AND ANALYSIS

Deployment

The completed model is combined with the Python AWS module and Dockers to be integrated into an application. The program only takes in input and outputs it.

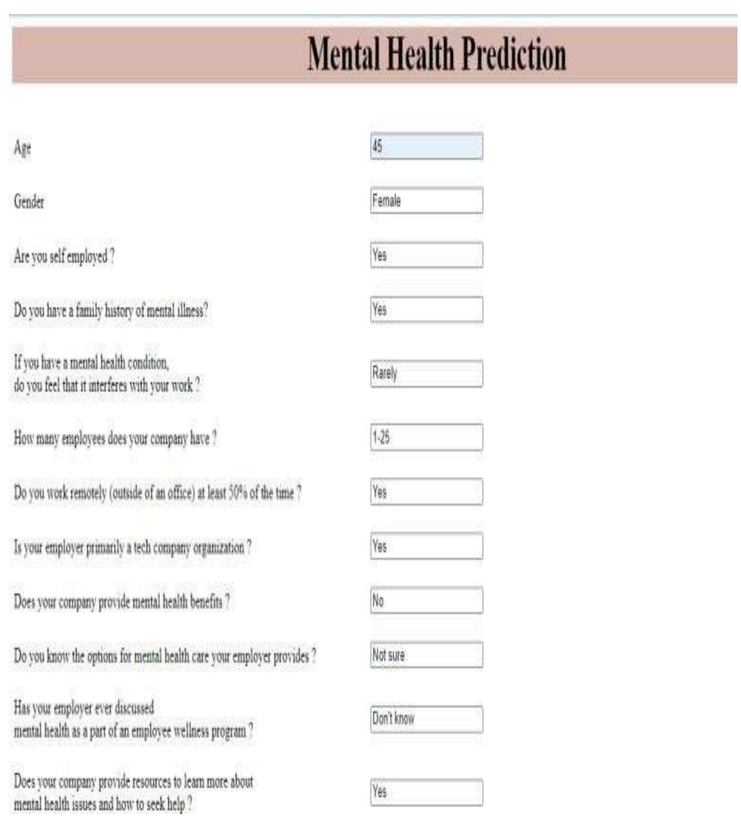

Fig.9: output screen

Comparison

TABLE I Accuracy of Existing Paper

Existing Paper	Accuracy
Decision Tree	0.78
Gradient boost	0.84
Logistic Regression	0.85
Random Forest	0.87
XGBoost	0.84
Support vector	0.87

TABLE II Accuracy of Research Paper

Research Paper	Accuracy
Logistic Regression	0.94
Support Vector	0.90
Decision tree	0.94
ADABoost	0.95
Gradient Boost	0.95
XGBoost	0.97
Random Forest	0.97

In our project we considered number of machine learning algorithms to predict the results and we found large variation between the existing paper results and our project results.

V.CLASSIFICATION REPORT

A classification report is a crucial tool in machine learning that is used to assess the effectiveness of a classification model. It offers a comprehensive overview of all the various performance metrics for each class in a classification task.

```
Cross-validation scores: [0.80934579 0.8317757 ]
Mean accuracy: 0.8205607476635515
```

```
Accuracy 0.985
Precission 0.9882352941176471
Sensitivity (True Positive Rate): 0.9767441860465116
Specificity (True Negative Rate): 0.9912280701754386
False Negative Rate: 0.023255813953488372
False Positive Rate: 0.008771929824561403
False Omission Rate: 0.017391304347826087
False Discovery Rate: 0.011764705882352941
False Predictive Rate: 0.017391304347826087
```

Fig. 10: Metrics using Logistic Regression

classification report for a Logistic Regression model:
Precision: It gauges how well optimistic forecasts come to pass. For instance, precision is 0.98, meaning that 98% of recommendations were accurate.
Accuracy: The percentage of accurate predictions; in this case, the model's accuracy is 0.87, meaning 87% of the classifications were correct.

Sensitivity: Sensitivity quantifies the ability of a machine learning model to identify positive examples. And here the sensitivity is 0.97%

Specificity: discusses how many negative records were accurately anticipated. Here the specificity is 0.99%

False negative rate: True but predicting as false 0.02%

False positive rate:0.008% stating false as true.

False omission rate: The value is true but predicting as negative of 0.01%

False discovery rate: When carrying out several comparisons in null hypothesis, it is employed to conceptualize type 1 errors in the 0.01%

False predictive rate: The percentage of false rate predicting by the model is 0.1%

After k-folds cross validation and mean accuracy is 0.89% and 0.83%

```
Cross-validation scores: [0.81308411 0.79439252 0.85514019 0.84579439 0.80841121]
Mean accuracy: 0.8233644859813085
```

```
0.965
0.9479166666666666
Sensitivity (True Positive Rate): 0.978494623655914
Specificity (True Negative Rate): 0.9532710280373832
False Negative Rate: 0.021505376344086023
False Positive Rate: 0.04672897196261682
False Omission Rate: 0.019230769230769232
False Discovery Rate: 0.052083333333333336
False Predictive Rate: 0.019230769230769232
```

Fig. 11: Metrics using SVM.

classification report for a Support vector machine Precision: It gauges how well optimistic forecasts come to pass. For instance, precision is 0.94, meaning that 94% of recommendations were accurate.

Accuracy: The percentage of accurate predictions; in this case, the model's accuracy is 0.96, meaning 96% of the classifications were correct.

Sensitivity: Sensitivity quantifies the ability of a machine learning model to identify positive examples. And here the sensitivity is 0.97%

Specificity: discusses how many negative records were accurately anticipated. Here the specificity is 0.95%

False negative rate: True but predicting as false 0.021%

False positive rate:0.046% stating false as true.

False omission rate: The value is true but predicting as negative of 0.01%

False discovery rate: When carrying out several comparisons in null hypothesis, it is employed to conceptualize type 1 errors in the 0.052%

False predictive rate: The percentage of false rate predicting by the model is 0.01%

After k-folds cross validation and mean accuracy is 0.81% and 0.82%

```
Cross-validation scores: [0.81308411 0.8364486  0.8364486  0.85981308 0.81775701]
Mean accuracy: 0.8327102803738319
```

```
0.96
0.9722222222222222
Sensitivity (True Positive Rate): 0.9545454545454546
Specificity (True Negative Rate): 0.9666666666666667
False Negative Rate: 0.045454545454545456
False Positive Rate: 0.03333333333333333
False Omission Rate: 0.05434782608695652
False Discovery Rate: 0.027777777777777776
False Predictive Rate: 0.05434782608695652
```

Fig. 12: Metrics using Decision Tree

classification report for the Decision Tree: Precision: It gauges how well optimistic forecasts come to pass. For instance, precision is 0.97, meaning that 97% of recommendations were accurate.

Accuracy: The percentage of accurate predictions; in this case, the model's accuracy is 0.96, meaning 96% of the classifications were correct.

Sensitivity: Sensitivity quantifies the ability of a machine learning model to identify positive examples. And here the sensitivity is 0.95%.

Specificity: discusses how many negative records were accurately anticipated. Here the specificity is 0.96%.

False negative rate: True but predicting as false 0.04%.

False positive rate:0.03% stating false as true.

False omission rate: The value is true but predicting as negative of 0.054%

False discovery rate: When carrying out several comparisons in null hypothesis, it is employed to conceptualize type 1 errors in the 0.027%

False predictive rate: The percentage of false rate predicting by the model is 0.054%

After k-folds cross validation and mean accuracy is 0.81% and 0.83%

```
Cross-validation scores: [0.80373832 0.8411215  0.86448598 0.84579439 0.80841121]
Mean accuracy: 0.8327102803738319
```

```
0.86
0.9024390243902439
Sensitivity (True Positive Rate): 0.7872340425531915
Specificity (True Negative Rate): 0.9245283018867925
False Negative Rate: 0.2127659574468085
False Positive Rate: 0.07547169811320754
False Omission Rate: 0.1694915254237288
False Discovery Rate: 0.0975609756097561
False Predictive Rate: 0.1694915254237288
```

Fig. 13: Metrics using Random Forest

classification report for the Random Forest: Precision: It gauges how well optimistic forecasts come to pass. For instance, precision is 0.90, meaning that 97% of recommendations were accurate.

Accuracy: The percentage of accurate predictions; in this case, the model's accuracy is 0.86, meaning 86% of the classifications were correct.

Sensitivity: Sensitivity quantifies the ability of a machine learning model to identify positive examples. And here the sensitivity is 0.78%

Specificity: discusses how many negative records were accurately anticipated. Here the specificity is 0.92%

False negative rate: True but predicting as false 0.2%

False positive rate:0.07% stating false as true.

False omission rate: The value is true but predicting as negative of 0.016%

False discovery rate: When carrying out several comparisons in null hypothesis, it is employed to conceptualize type 1 errors in the 0.09%

False predictive rate: The percentage of false rate predicting by the model is 0.16%

After k-folds cross validation and mean accuracy is 0.80% and 0.83

```
Cross-validation scores: [0.70560748 0.78504673 0.78504673 0.77570093 0.73831776]
Mean accuracy: 0.7579439252336448
```

```
0.955
0.9405940594059405
Sensitivity (True Positive Rate): 0.9693877551020408
Specificity (True Negative Rate): 0.9411764705882353
False Negative Rate: 0.030612244897959183
False Positive Rate: 0.058823529411764705
False Omission Rate: 0.030303030303030304
False Discovery Rate: 0.0594059405940594
False Predictive Rate: 0.030303030303030304
```

Fig. 14: Metrics using ADABoost.

Classification report for a ADABoost classifier:

Precision: It gauges how well optimistic forecasts come to pass. For instance, precision is 0.94, meaning that 94% of recommendations were accurate.

Accuracy: The percentage of accurate predictions; in this case, the model's accuracy is 0.95, meaning 95% of the classifications were correct.

Sensitivity: Sensitivity quantifies the ability of a machine learning model to identify positive examples. And here the sensitivity is 0.96%

Specificity: discusses how many negative records were accurately anticipated. Here the specificity is 0.94%

False negative rate: True but predicting as false 0.03%

False positive rate:0.05% stating false as true.

False omission rate: The value is true but predicting as negative of 0.03%

False discovery rate: When carrying out several comparisons In null hypothesis, it is employed to conceptualize type 1 errors in the 0.05%

False predictive rate: The percentage of false rate predicting by the model is 0.03%

After k-folds cross validation and mean accuracy is0.70% and 0.75%

```
Cross-validation scores: [0.81775701 0.8364486  0.86448598 0.85514019 0.80841121]
Mean accuracy: 0.836448598130841
```

```
0.985
1.0
Sensitivity (True Positive Rate): 0.9705882352941176
Specificity (True Negative Rate): 1.0
False Negative Rate: 0.029411764705882353
False Positive Rate: 0.0
False Omission Rate: 0.0297029702970297
False Discovery Rate: 0.0
False Predictive Rate: 0.0297029702970297
```

Fig. 15: Metrics using Gradient Boost

Classification report for a Gradient Boost classifier

Precision: It gauges how well optimistic forecasts come to pass. For instance, precision is 0.1, meaning that 100% of recommendation were accurate.

Accuracy: The percentage of accurate predictions; in this case, the model's accuracy is 0.98, meaning 98% of the classifications were correct.

Sensitivity: Sensitivity quantifies the ability of a machine learning model to identify positive examples. And here the sensitivity is 0.97%

Specificity: discusses how many negative records were accurately anticipated. Here the specificity is 100%

False negative rate: True but predicting as false 0.02%

False positive rate:0.0% stating false as true.

False omission rate: The value is true but predicting as negative of 0.02%

False discovery rate: When carrying out several comparisons In null hypothesis, it is employed to conceptualize type 1 errors in the 0.0%

False predictive rate: The percentage of false rate predicting by the model is 0.02%

After k-folds cross validation and mean accuracy is 0.81% and 0.83 %

```
Cross-validation scores: [0.81775701 0.8364486  0.86448598 0.85514019 0.80841121]
Mean accuracy: 0.836448598130841
```

```
1.0
1.0
Sensitivity (True Positive Rate): 1.0
Specificity (True Negative Rate): 1.0
False Negative Rate: 0.0
False Positive Rate: 0.0
False Omission Rate: 0.0
False Discovery Rate: 0.0
False Predictive Rate: 0.0
```

Fig.16: Metrics using XGBoost.

Classification report for a XGBoost classifier.

Precision: It gauges how well optimistic forecasts come to pass. For instance, precision is 0.1, meaning that 100% of recommendations were accurate.

Accuracy: The percentage of accurate predictions; in this case, the model's accuracy is 0.1, meaning 100% of the classifications were correct.

Sensitivity: Sensitivity quantifies the ability of a machine learning model to identify positive examples. And here the sensitivity is 0.1%

Specificity: discusses how many negative records were accurately anticipated. Here the specificity is 0.1%

False negative rate: True but predicting as false 0.0%

False positive rate:0.0% stating false as true.

False omission rate: The value is true but predicting as negative of 0.0%

False discovery rate: When carrying out several comparisons in null hypothesis, it is employed to conceptualize type 1 errors in the 0.0%

False predictive rate: The percentage of false rate predicting by the model is 0.0%

After k-folds cross validation and mean accuracy is 0.80% and 0.0 %

V. CONCLUSIONANDFUTURESCOPE

To sum up, this study investigated the use of popular machine learning algorithms, such as ADABoost, GRADIENT Boost, SVM, Random Forest, Decision Tree, Logistic Regression, and XGBoost, to calculate the likelihood that a child will experience mental illness. The acquired findings show that the algorithms can produce accurate predictions, with XGBoost and Random Forest surpassing all other models. These two models function in terms of productivity and accuracy. The factors taken into account in this predictive model include age, self-employment, work-life balance, family history, coworkers, and the effects on one's physical and mental health. All things considered, this work demonstrates the promise of machine learning algorithms and how they might improve forecast accuracy and dependability. By adding more intricate characteristics, investigating novel and alternate techniques, and assessing the model's performance on sizable datasets, future research can build on this work.

Accuracy comparison between the taken existed base paper and present base paper.

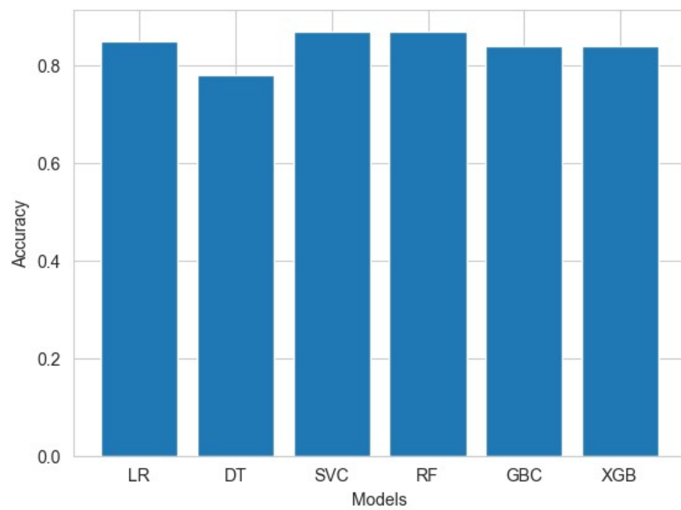

Fig.17: Accuracy of Existed paper

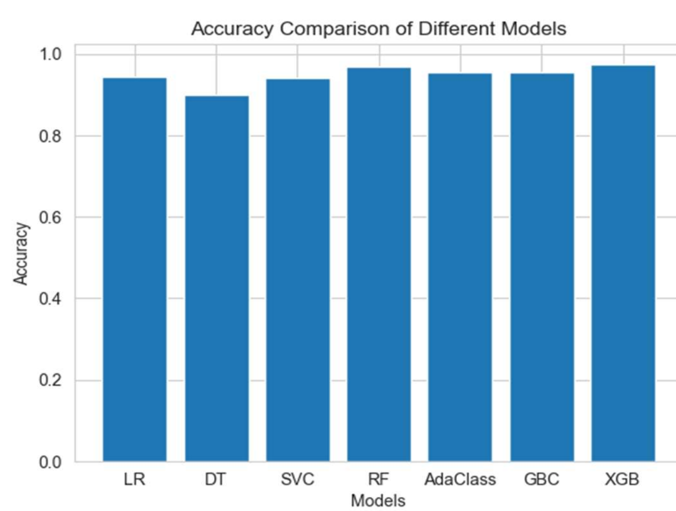

Fig.18: Accuracy comparison graph

VI. REFERENCES

[1] Stevie Chancellor, Eric PS Baumer, and Munmun De Choudhury. Who is the" human" in human-centered machine learning: The case of predicting mental health from social media. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW):1–32, 2019.

[2] Sarah Graham, Colin Depp, Ellen E Lee, Camille Nebeker, Xin Tu, Ho-Cheol Kim, and Dilip V Jeste. Artificial intelligence for mental health and mental illnesses: an overview. Current psychiatry reports, 21(11):1–18, 2019

[3] Theodoros Iliou, Mandani Ntekouli, Christina Iymperopoulou, Konstantin's Assimakopoulos, Dimities Galiatsatos, and George Anastassopoulos. Iliou machine learning preprocessing method for depression type prediction. Evolving Systems, 10(1):29–39, 2019 .

[4] T Nagar. Prediction of mental health problems among children using machine learning techniques

[5] U. Srinivasulu Reddy, Aditya Vivek Thota and A. Dharun, "Machine learning techniques for stress prediction in working employees", 2018 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2018.

[6] Suresh, A Stephen M. Schueller, Adrian Aguilera and David C. Mohr, "Ecological momentary interventions for depression and anxiety", Depression and Anxiety, vol. 34, no. 6, pp. 540-545, 2017 .

[7] U. Srinivasulu Reddy, Aditya Vivek Thota and A. Dharun, "Machine learning techniques for stress prediction in working employees", *2018 IEEE* International Conference on Computational Intelligence and Computing Research (ICCIC), 2018.

[8] Suresh, G., A. Senthil Kumar, S. Lekashri, and R. Manikandan. "Efficient Crop Yield Recommendation System Using Machine Learning for Mental Health prediction." International Journal of Mental health 10, no. 1 (2021): 906-914.

[9] D. Graham, Sarah, "Artificial intelligence for mental health and mental illnesses: an overview." Current psychiatry reports 21.11 (2019): 1-18.

[10] Bhattacharyya, R., & Basu, S. (2018). India Inc "looks to deal with rising stress in employees". Retrieved from 'The Economic Times'

[11] Shatter, Adrian BR, Delyse M. Hutchinson, and Samantha J. Teague. "Machine learning in mental health" a scoping review of methods and applications. Psychological medicine (2019)

[12] M. Hamilton, "Development of a rating scale for primary depressive illness," *British Journal of Social and Clinical Psychology*, vol. 6, no. 4, pp. 278–296, 1967.

[13] G. Cho, J. Yim, Y. Choi, J. Ko, and S.-H. Lee, "Review of machine learning algorithms for diagnosing mental illness," *Psychiatry Investigation*, vol. 16, no. 4, pp. 262–269, 2019.

[14] M. Hamilton, "Development of a rating scale for primary depressive illness," *British Journal of Social and Clinical Psychology*, vol. 6, no. 4, pp. 278–296, 1967.

[15] Dataset link https://www.kaggle.com/code/kairosart/machine-learning-for-mental-health-1/input