MENTAL HEALTH PREDICTION AT WORKPLACE USING MACHINE LEARNING

PRESENT BY

Shyam Lal Kafle	19691A05F0
Sandesh Pokhrel	19691A05J0
Jiwan Chaudhary Tharu	19691A05J4
Nabin Shahi	19691A05J6

Under the guidance
of
Mr. K H Shabbeer Basha, M.Tech
Assistant Professor

Department of Computer Science and Engineering
Madanapalle Institute of Technology and Science

March 10, 2023

OUTLINE

- Problem Formulation
- Feasibility
- Literature survey
- Existing Work
- Disadvantage of Existing Work
- Proposed Work
- Proposed Model Diagram
- Advantage of Proposed Work
- Requirement
- Uml Diagram
- Model Design
- Dataset
- Algorithm used
- Accuracy of ML algorithms
- Finding algorithm with high accuracy
- Performance measures analysis
- References

PROBLEM FORMULATION

- The problem of mental health prediction in the workplace using machine learning can be formulated as a multiple classification task
- Objective is to predict if an employee is likely to have a mental health issue based on their workplace-related factors.
- The key challenges in this problem include :
  - Data collection
  - Data pre processing
  - Feature selection
  - Model selection
  - Model Evaluation
  - Interpretability
  - Privacy and Ethics

FEASIBILITY

- **Technical Feasibility :**The technical requirement for the system is economic. It doesn’t use any hardware and software.

- **Operation Feasibility:** The operational feasibility includes User friendly, Reliability, Security, Potability, Availability and Maintainability of the software using project.

- **Economic Feasibility :** Analysis of the project cost and revenue in an effort to determine whether or not it is logical and possible to complete.

LITERATURE SURVEY : 1

Author Name and Publication	Title of paper	ML Technique used	Review of this paper
Gundeti Sai Manvith, Gunna Shivani Reddy,Bandi Lakshma Reddy. International Reaserch Journal of Modernization in Engineering Technology and Science,Vol.04/Issue.07. **(July 2022)**	Mental Health Prediction for an Individual using Machine Learning	Decision tree & Support Vector Machine	The Feasibility Investigation examine the longevity of a standard peak detection method which can be seen done using Support Vector Machine and Decision Tree which has higher Accuracy in prediction of mental health is 97%.

LITERATURE SURVEY : 2

Author Name and Publication	Title of paper	ML Technique used	Review of this paper
Jetli Chung ,Jason Teo Faculty of Computing and informatics, University Malaysia Sabah,Jalan UMS,88400 Kota Kinabalu, Sabah,Malaysia. (Jan 2022)	Mental Health Prediction using Machine Learning : Taxonomy Application Challenges	Bayesian Network, Naïve Bayes, Logistic Regression, Multiple Layer Perceptron, Sequential Minimal Optimization, K-star, Random Sub Space, J48, Random Forest, Random Tree	The main objective of this paper is to provide systematic literature review, critical review and summary of the machine learning technique that are being used to predict and diagnose and identify the mental health problem.

LITERATURE SURVEY : 3

Author Name and Publication	Title of paper	ML Technique used	Review of this paper
Dr.J.Arokia Renjit,Adlin Sajeesha M,J Sangavai V.D.Sree Devi D.SJETIR Journal of Emerging Technologies and innovative Reaserch, Volume 9, Issue 5. [May 2022]	Prediction of Mental Health using Machine Learning	Supervised Learning, Logistic Regression, Random Forest, Naïve Bayes, Decision Tree Classifier & Flask Deployment	This research paper suggest that machine learning technique could be used to predict mental health condition and identify at –risk individual in a cost effective and efficient manner.

LITERATURE SURVEY : 4

Author Name and Publication	Title of paper	ML Technique used	Review of this paper
Sofiantia Mutalib, Nor Sakifa Mohd Shafiee, Shuzlina Abdul-RahmanTurkish Journal of Computer and Mathematics Education Vol12.No.5(2021),1872-1972. [Jan 2021]	Mental Health Prediction Models using Machine Learning in Higher Education	Logistic Regression, Naïve Bayes, Decision Tree, Neural Network, Support Vector Machine	This paper compare the following algorithm with dataset and find out that among all this algorithm, Decision Tree, Support Vector Machine and Neural Network are the high accuracy and effective algorithm.

LITERATURE SURVEY : 5

Author Name and Publication	Title of paper	ML Technique used	Review of this paper
konda Vaishnavi et al J.Phys: Conf.Ser.2161 012021. [2022]	Predicting Mental Health Illness using Machine Learning Algorithms	Logistic Regression, Decision Tree Classifier, KNN Classifier, Random Forest Classifier, Stacking	This research paper shown that mental health prediction using machine learning algorithm using five difference algorithm, get highest accuracy of 81.75% and lowest accuracy of 79.63%.

EXISTING SYSTEM

- Predicting Depression
- Mental Health Assessment
- Anxiety and stress prediction
- Suicide Risk Assessment

Note:

They use Decision tree, Support Vector Machine,Bayesian Network,Naïve Bayes,Logistic Regression,Multiple Layer Perceptron,Sequential Minimal Optimization Algorithm for predicting above following mental health related problems.

DISADVANTAGE OF EXISTING SYSTEM

- Data Quality and Bias
- Lack of interpretability
- Privacy concern
- Ethical Consideration
- Limitation of self reported data
- Limitation of physiological sign
- Lack of clinical validation

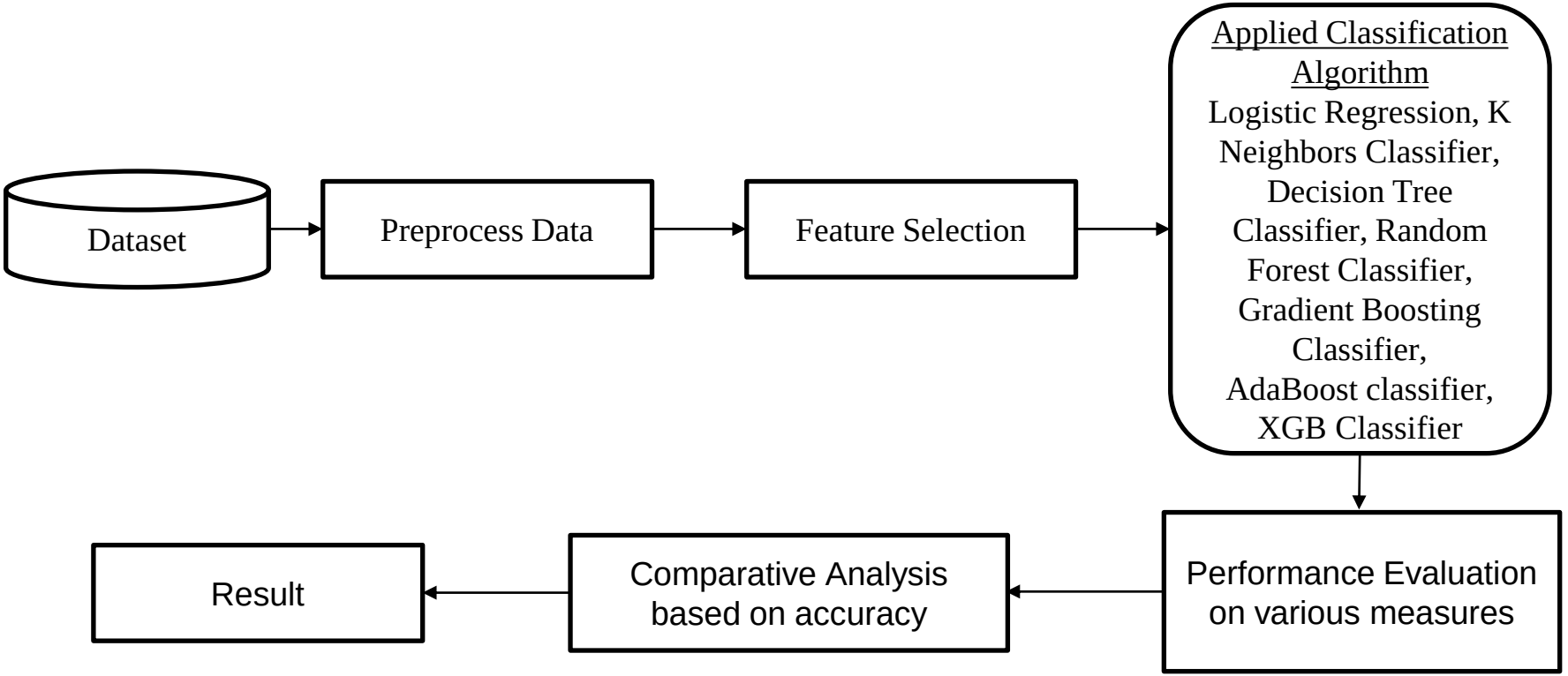
PROPOSED WORK

- In this project we try to predict employee stress prediction, work place burnout prediction, mental health support recommendation using different algorithms: **Logistic Regression, K Neighbors Classifier, Decision Tree Classifier, Random Forest Classifier, Gradient Boosting Classifier, XGB Classifier** with large number of dataset.

- After performing, we analysed the performance of Accuracy, Confusion Matrix, F1 score, Precision, ROC Curve by using above algorithms.

- After analyzing the performance we compare the algorithm and choose the best algorithm which has high accuracy and efficiency to predict the employee attrition using Machine Learning Technique.

PROPOSED MODEL DIAGRAM



ADVANTAGE OF PROPOSED WORK

- Improving predictive accuracy
- Incorporating additional factors
- Personalized interventions
- Integration with existing HR system
- Ethical and legal considerations
- Address the employee issue

REQUIREMENTS

Hardware Requirement :

- System : Pentium i3 processor
- Hard disk : 256 GB
- Monitor : 14“ LED
- Input Devices : Keyboard, Mouse
- RAM : 4 GB

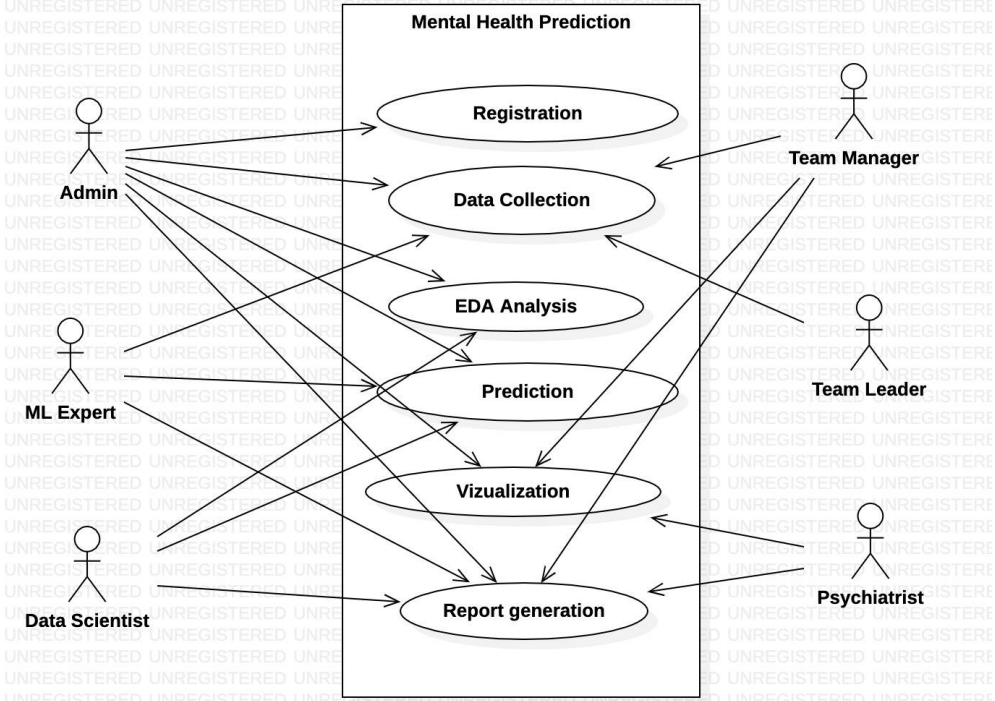
Software Requirement :

- Operating System : Window 10,11
- Coding Language : Python
- Coding Platform : Juypter Notebook, Google Colaboratory
- UML Diagram: Star UML

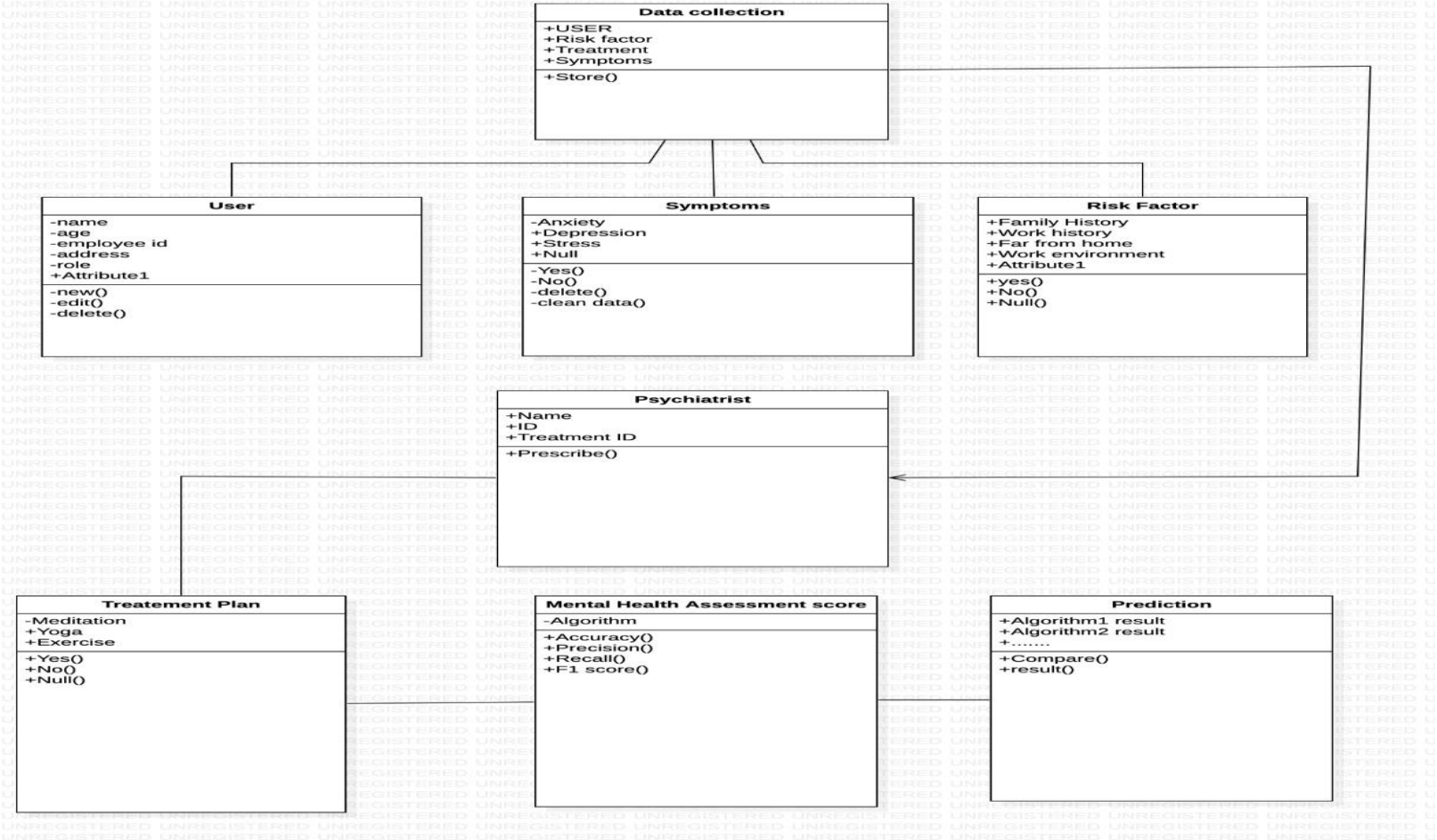
UML DIAGRAM

- **Use Case Diagram**
- **Class Diagram**
- **Sequence Diagram**

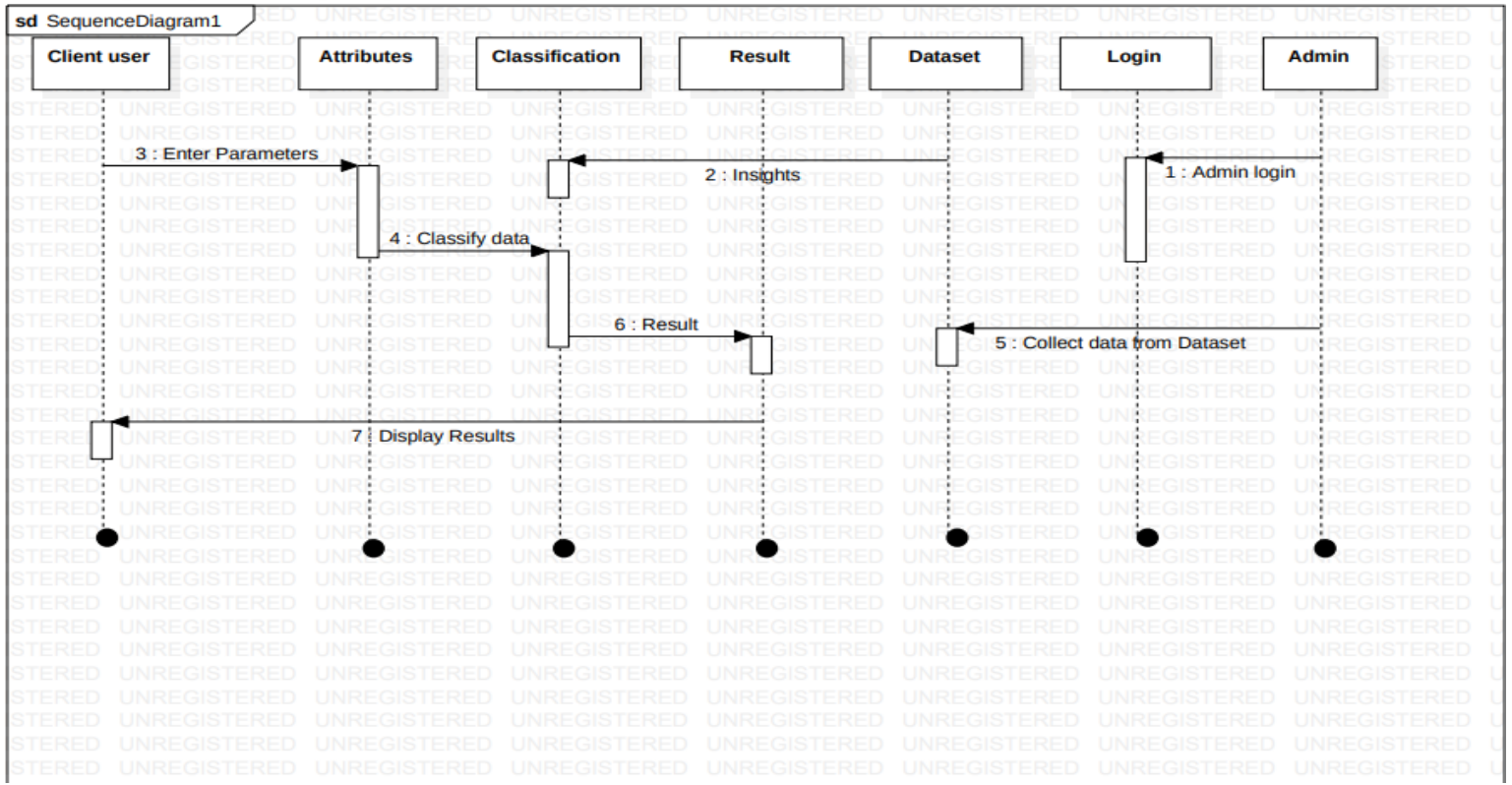
USECASE DIAGRAM



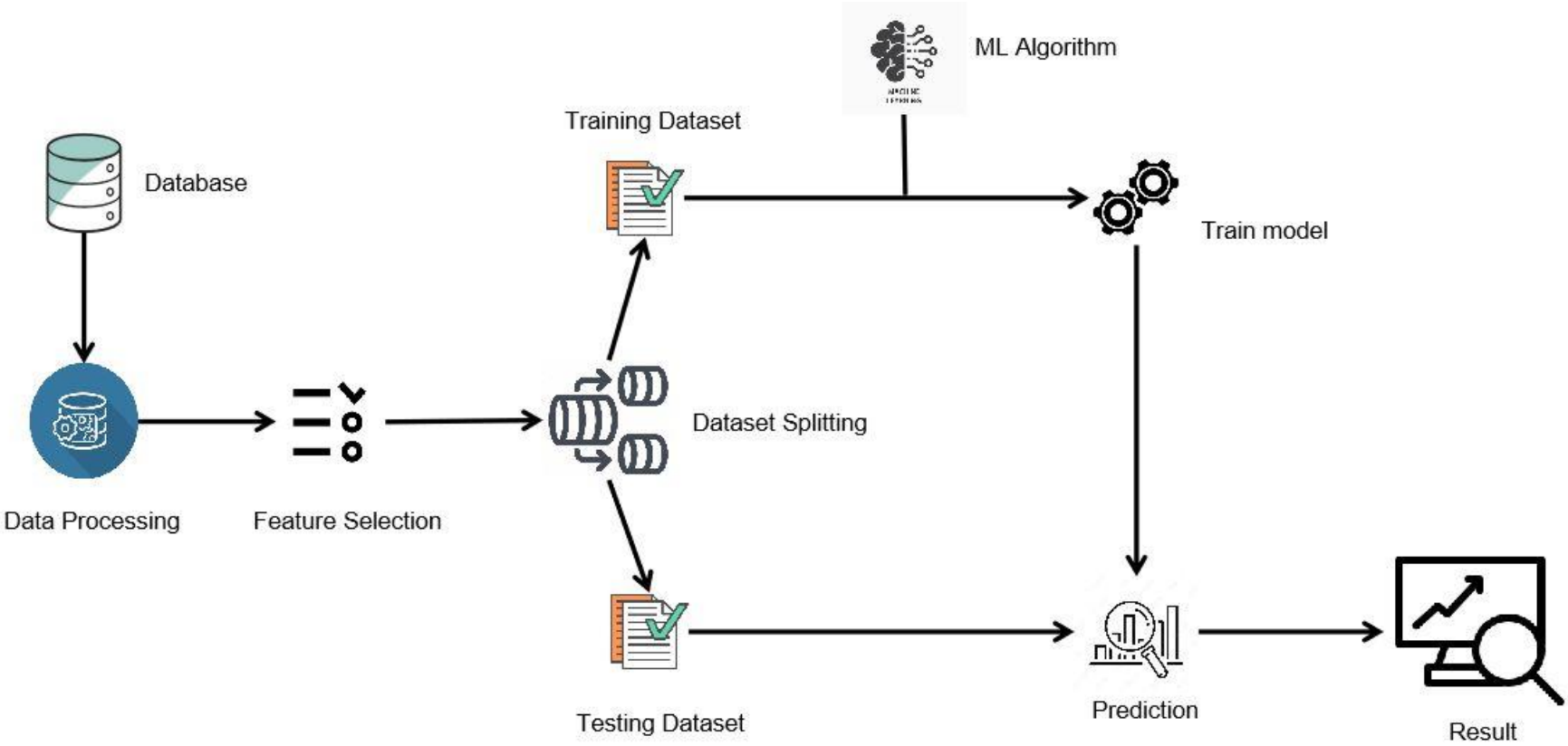
CLASS DIAGRAM

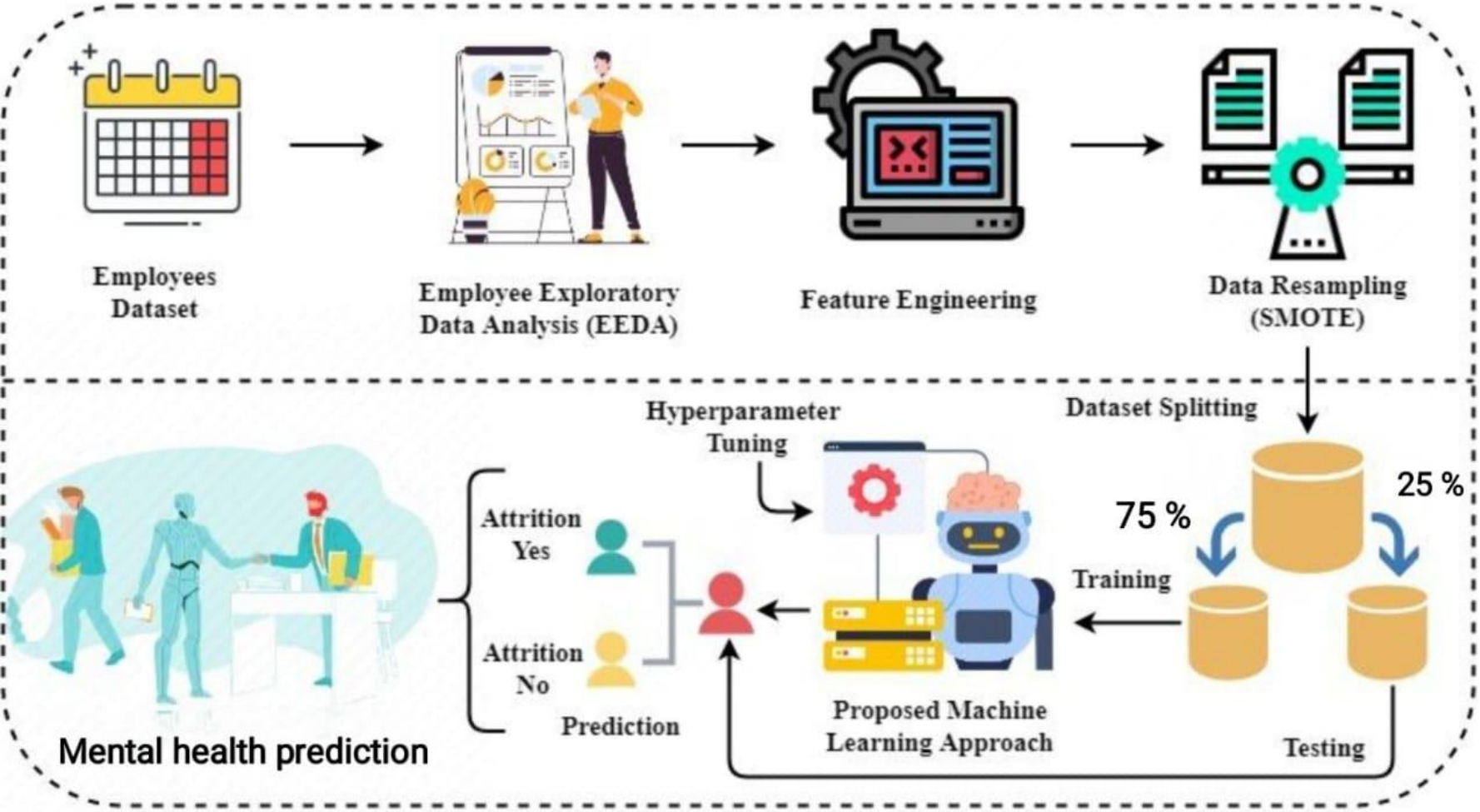


SEQUENCE DIAGRAM

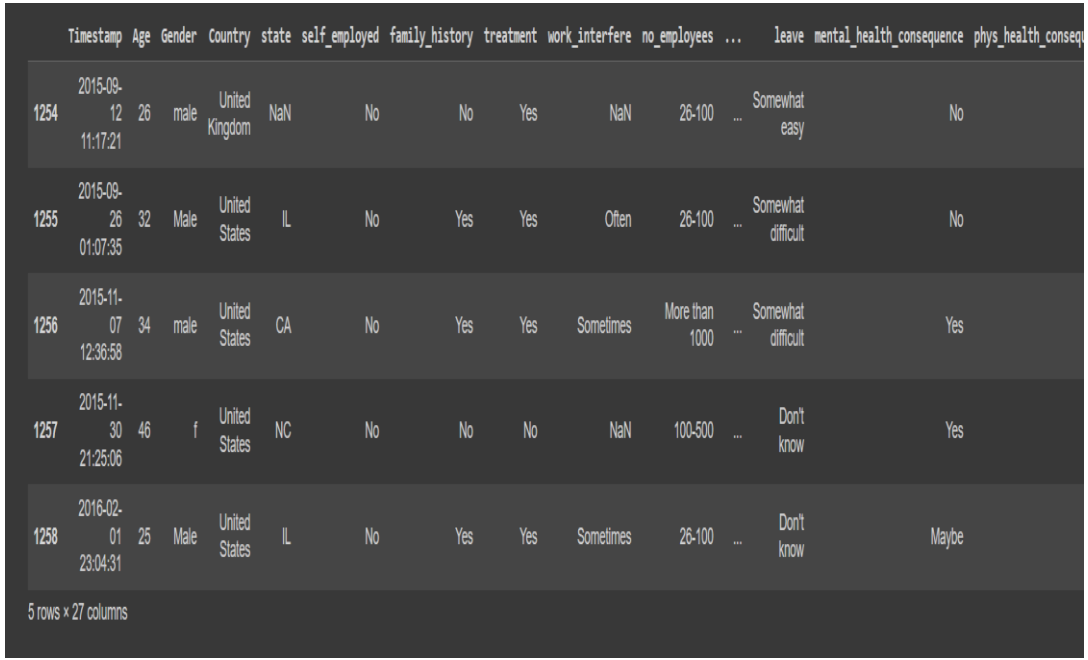


MODEL DESIGN




Employees Dataset
Employee Exploratory Data Analysis (EEDA)
Feature Engineering
Data Resampling (SMOTE)
Dataset Splitting
25 %
75 %
Training
Testing
Hyperparameter Tuning
Proposed Machine Learning Approach
Attrition Yes
Attrition No
Prediction
Mental health prediction

DATASET



	Timestamp	Age	Gender	Country	state	self_employed	family_history	treatment	work_interfere	no_employees	...	leave	mental_health_consequence	phys_health_consequ
1254	2015-09-12 11:17:21	26	male	United Kingdom	NaN	No	No	Yes	NaN	26-100	...	Somewhat easy	No	
1255	2015-09-26 01:07:35	32	Male	United States	IL	No	Yes	Yes	Often	26-100	...	Somewhat difficult	No	
1256	2015-11-07 12:36:58	34	male	United States	CA	No	Yes	Yes	Sometimes	More than 1000	...	Somewhat difficult	Yes	
1257	2015-11-30 21:25:06	46	f	United States	NC	No	No	No	NaN	100-500	...	Don't know	Yes	
1258	2016-02-01 23:04:31	25	Male	United States	IL	No	Yes	Yes	Sometimes	26-100	...	Don't know	Maybe	

5 rows × 27 columns

Description of data set

Dataset used : Survey.csv

Source:[Mental Health in Tech Survey | Kaggle]

Dataset Size:[1259 rows * 25 columns]

Training dataset : 75%

Testing Dataset: 25%

The dataset consist of several type of parameters

1. Age
2. Gender
3. Country
4. state
5. self-employed
6. family history
7. treatment
8. work interferes
9. no employees'
10. remote work
11. tech_company
12. benefits
13. care options
14. wellness program
15. seek help
16. anonymity
17. leave
18. mental_health_consequence
19. phys_health_consequence
20. coworkers
21. supervisor
22. mental_health_interview
23. phys_health_interview
24. mental_vs_physical
25. obs_consequence

ALGORITHM USED

1. Logistic Regression
2. K-Nearest Neighbor’s
3. Decision Tree classifier
4. Random Forest Classifier
5. Gradient Boosting classifier
6. Ada Boost classifier
7. XGB classifier

ACCURACY OF ML ALGORITHMS

```
[ ]  predicted=[]
     for name,algo in models.items():
      model=algo
      model.fit(X_train,y_train)
      predict = model.predict(X_test)
      acc = round(accuracy_score(y_test, predict),3)
      #acc = (accuracy_score(y_test, predict))
      predicted.append(acc)
      print(name,acc)
```

```
LogisticRegression 0.79
KNeighborsClassifier 0.697
DecisionTreeClassifier 0.748
RandomForestClassifier 0.803
GradientBoostingClassifier 0.806
AdaBoostClassifier 0.787
XGBClassifier 0.796
```

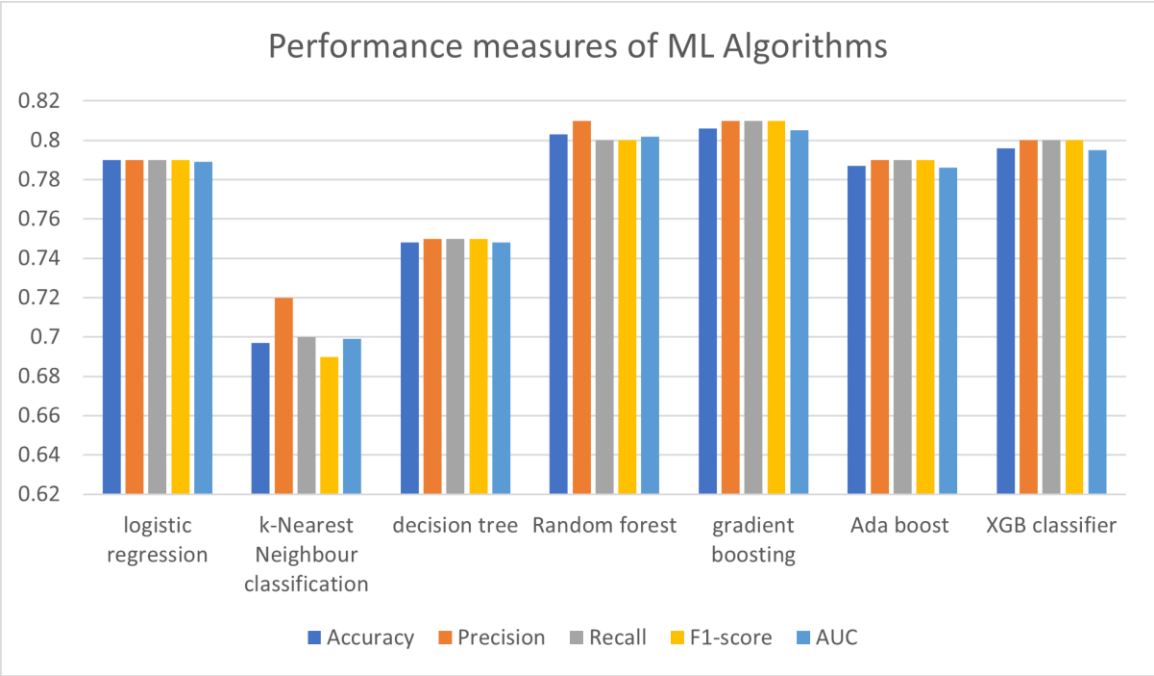
FINDING ALGORITHM WITH HIGH ACCURACY

```
m=max(list(predd.values()))
print("The Algorithm which gives maximum accuracy...")
for i in predd:
  if predd[i]==m:
    print("{} : {}".format(i,predd[i]))
```

```
The Algorithm which gives maximum accuracy...
GradientBoostingClassifier : 0.8057324840764332
```

We found that “ Gradient Boosting Classifier Algorithm” is achieving the highest accuracy among all the algorithm.

PERFORMANCE MEASURES ANALYSIS



REFERENCES

[1] Gundeti Sai Manvith, Gunna Shivani Reddy,Bandi Lakshma Reddy, *"Mental Health Prediction for an Individual using Machine Learning"* IRJMETS International Reaserch Journal of Modernization in Engineering Technology and Science,Vol.04/Issue.07. [July 2022]

[2] Jetli Chung ,Jason Teo, *"Mental Health Prediction using Machine Learning : Taxonomy Application Challenges"* Faculty of Computing and informatics, University Malaysia Sabah,Jalan UMS,88400 Kota Kinabalu, Sabah,Malaysia. [Jan 2022]

[3] Dr.J.Arokia Renjit,Adlin Sajeesha M,J Sangavai V.D.Sree Devi D.S, *"Prediction of Mental Health using Machine Learning"* JETIR Journal of Emerging Technologies and innovative Reaserch, Volume 9, Issue 5. [May 2022]

[4] Sofiantia Mutalib, Nor Sakifa Mohd Shafiee, Shuzlina Abdul-Rahman, *"Mental Health Prediction Models using Machine Learning in Higher Education"* Turkish Journal of Computer and Mathematics Education Vol12.No.5(2021),1872-1972. [Jan 2021]

[5] konda Vaishnavi et al, *"Predicting Mental Health Illness using Machine Learning Algorithms "* J.Phys: Conf.Ser.2161 012021. [2022]

THANK YOU