Forecasting Depression in Bipolar Disorder Using Machine Learning Algorithms

Presented By:

Abdul Baqi Malik ***1886-2015***

Muhammad Anas Bin Munzir ***1614-2016***

Project Supervisor:

Sir Muhammad Shafiq

Department of Computing

Hamdard Institute of Engineering & Technology

Hamdard University Islamabad Campus

Overview of Presentation

- Introduction
- Aims & Objectives
- Problem Statement
- Literature Review
- Methodology
- Tools & Technology
- Executed work

i. Data Cleaning
ii. Exploratory Data Analysis
iii. Data Combination
iv. Algorithm Implementation
v. Results and Predictions

- Conclusion
- References

Introduction

- Mental illness is a type of health condition that changes a person's mind, emotions, or behavior.
- For understanding mental health conditions:
  - Provide better patient care.
  - Early detection of mental health problem.
  - Mental illness diagnoses based on an patient's self-report to specific questionnaires.
- Still a low level of overall recognition rate and treatment rate of the disease.
- Patient's spend average ten years between initial symptoms and the formal diagnosis.

Bipolar Disorder

Bipolar Disorder:

- A manic-depressive illness
- It is a brain disorder that causes unusual shifts in mood, energy, activity levels and the ability to carry out day-to-day tasks.

Symptoms:

- People with bipolar disorder experience unusually intense emotional states that occur in distinct periods called “mood episodes”.
  - ✓An overly joyful or excited state is called ‘manic episode’.
  - ✓An extremely sad or hopeless state is called a ‘depressive episode’.
  - ✓A mood episode includes symptoms of both mania and depression is called ‘a mixed state’.

Bipolar Disorder (cont…)

- The project was done to obtain:
  - ✓ For better understanding of forecasting in data science and their functions.
- It is designed to help for forecasting bipolar disorder using machine and deep learning.
- With the help of data science tools
  - Predict accuracies on different data sets.
  - Visualize through different mapping schemes like scatter plots, histograms and heat maps.
- The use of ML algorithms allows
  - ✓ The design of both classification and regression models.
  - ✓ Use for the diagnosis of different diseases, recommendation of drugs, automatic administration of drugs, etc.
- A predictive model is built using ML algorithms such as logistic regression, decision trees, random forest … etc. With the help of these algorithms:
  - Predict values based on different features.
  - Making these predictions better the more data they train on.

Bipolar Disorder (Cont…)

- Deep learning includes statistics and predictive modeling. DL beneficial to data scientists who are tasked with collecting, analyzing and interpreting large amount of data.
- This project will be done in two parts.

✓ The first part will be

- Data collection from various authentic resources.
- learning and understanding of data.
- Process of cleaning and visualizing patients data.
- It includes data analysis which contains code snippets and plots helps to understand the data better.
- Observe possible relationships and dependencies between them.

✓ Second part includes

- Predictive analysis of data.
- Compare different accuracies and select the algorithm that suit most to the problem.

Strong and Weak Points

- **Strong Points of Project:**

- Monitor patient's conditions.
- Set parameters, analyzed data.
- Project will be able to react on time according to patient's symptoms.
- Apply ML algorithms and predict condition.

- **Weak points of Project:**

- We have not to much huge size of dataset for finding better accuracy score.

- **Recommendation:**

- On huge dataset, apply new techniques and predict more accuracies.

Aims & Objectives

- **Aim:**

Aim of project is:

- Apply different ML algorithms on symptoms based patient's data.
- Crete a prediction model.
- Model would make it easier for psychiatrists to decide the state of patient's.
- State: Depression/mania episode/euthymic state.

- **Objective:**
- Develop a model using ML algorithms that will
  - analyze on patients data
  - predict their mental illness/mood swings.
- The purpose of this work
  - to improve the percentage of correctly predicted depressive disorder.

Problem Statement

- BpD identified by
  - Oscillations of patient changing mood.
  - Mood swings between two states i.e. depression and mania
  - Observed features like behavioral changes, mood swings, mental illness affects (i.e. confused thinking, tiredness, withdrawal activities .. Etc.)
- Late diagnosis in majority cases.
- Can not predict the state of the disorder at same time.
- BpD state vary from patient to patient.
- Base-paper problem:

  Bip4cast: - Predict crises in patient's with BpD.
  - Small size of data set.
  - Avoid previous history
  - Collect data through interviews only.
  - Required years for recovery due to late diagnosis
  - Predict low accuracies in many cases

Literature Review

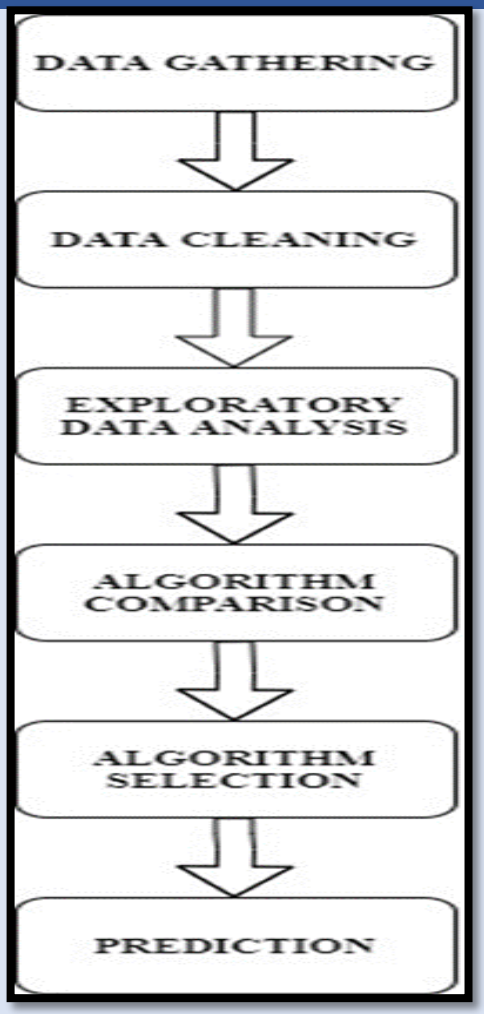
Aim of the paper	Author's	ML Approach	Data Source	Mental Health Type	Advantages	Disadvantages
Forecasting depression in BpD	P.J. Moore	CNN	SMS text messaging	Depression	- Using self rated mood data - Forecast next week mood ratings. - Conclude depression time series	- Limited data sets - Can't apply multiple approaches
Deep Learning for depression detection of twitter users	Ahmed Husseini Orabi, Prasadith Buddhitha, Mahmoud Husseini Orabi, Diana Inkpen	MLP, supervised learning CNN, RNN	Social Media	Depression	- Easily gathered data - Better accuracy than previous - Good results which are helpful for psychologists	- Official testing sets which are not available for us.
Implementing machine learning in bipolar diagnosis	Yantao Ma, Jun Ji, Yun Huang, Yue Zhu, Xin Yu…	BN, SMO, MLP, CNN, MCAR	Comprehensive Assessment, Questionnaire	Depression	- Individual observation - Use multiple approaches - Try to predict better accuracy	- Low accuracy percentage - Not feasible for practical purposes. - Add more complexity in patient's diagnosis
Review on Machine Learning Techniques to predict Bipolar Disorder	Nisha Agnihotri, Sanjeev Kumar Prasad	Decision Tree, Random Forest, SVM	Survey's, Questionnaire, Social Media Platform	Anxiety, Depression and other mental disorders	- Provides great insights in the fields such as disease diagnosis, personalized treatment. - Concludes prediction of stress and mental health conditions	- If we don't control emotional conditions, anxiety will become worse day by day. - Anxiety turns up to a pathological situations that is quite challenging to treat.

Literature Review

Aim of the paper	Author's	ML Approach	Data Source	Mental Health Type	Advantages	Disadvantages
Classification of Schizophrenia and Bipolar Disorder by using ML algorithms	Cemil Cyan Saylan, Kaan Yilancioglu	Decision Tree, Naïve Bayes, K-NN	Clinical Information (text, speech ..)	Schizophrenia, depression and manic depressive disorder	- Data mining based investigations - Potential diagnostic biomarkers - Classify more than one mental illness disease - Predict disease state in test set with high accuracy	- Storage - Bias occurring in diagnostic biomarkers - High cost
Machine Learning and Big Data Analytics in Bipolar Disorder (Review)	Dr. Ives Cavalcante, Dr. Elisa Brietzke, Dr. Martin Alda, Prof. Anne Duffy ...	Random Forest, Naïve Bayes, SVM, RVM ..	Neuroimaging, blood biomarkers, clinical information	Borderline personality disorder, depression	- Classified majority of participants into the correct diagnostic group using standard approaches. - Provide an opportunity to more accurately detect those who are at risk.	- Several methodological challenges need to be addressed in order to translate research findings to clinical setting.

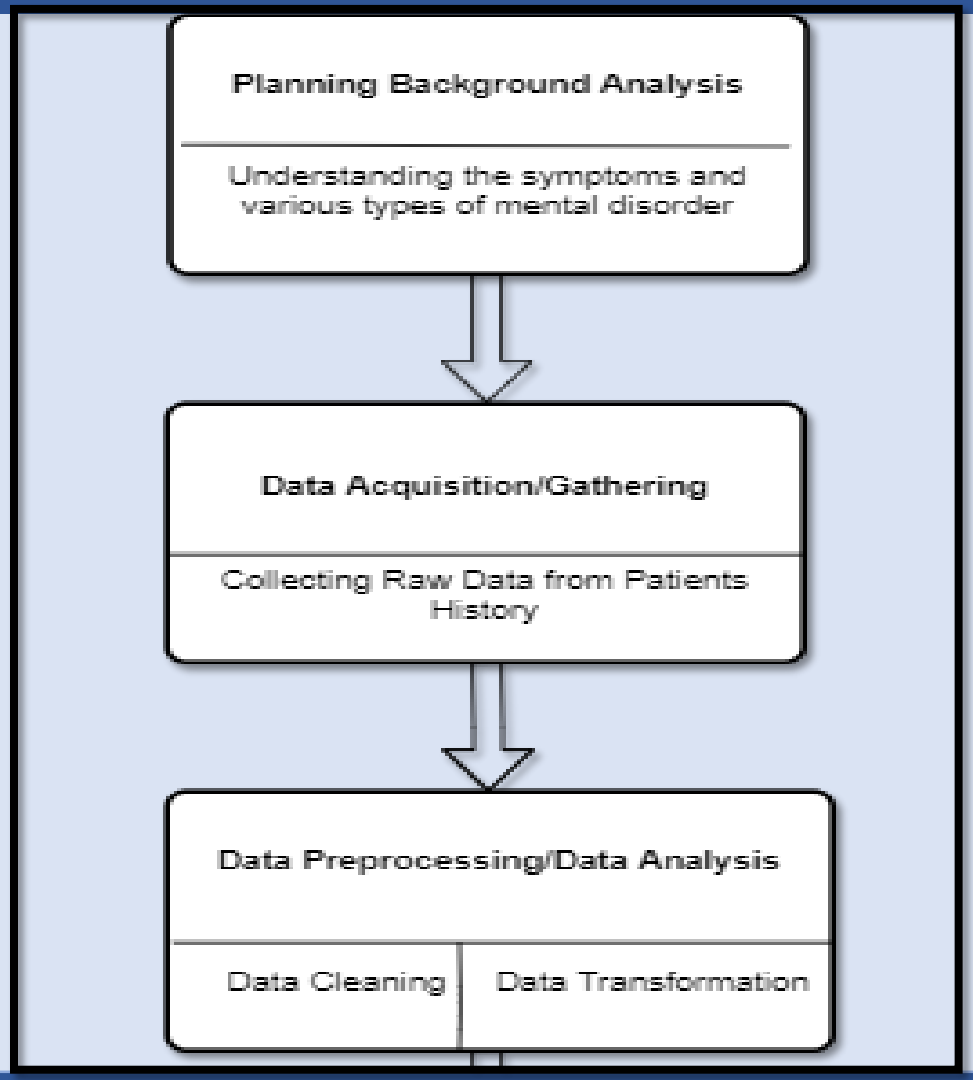
Proposed Method

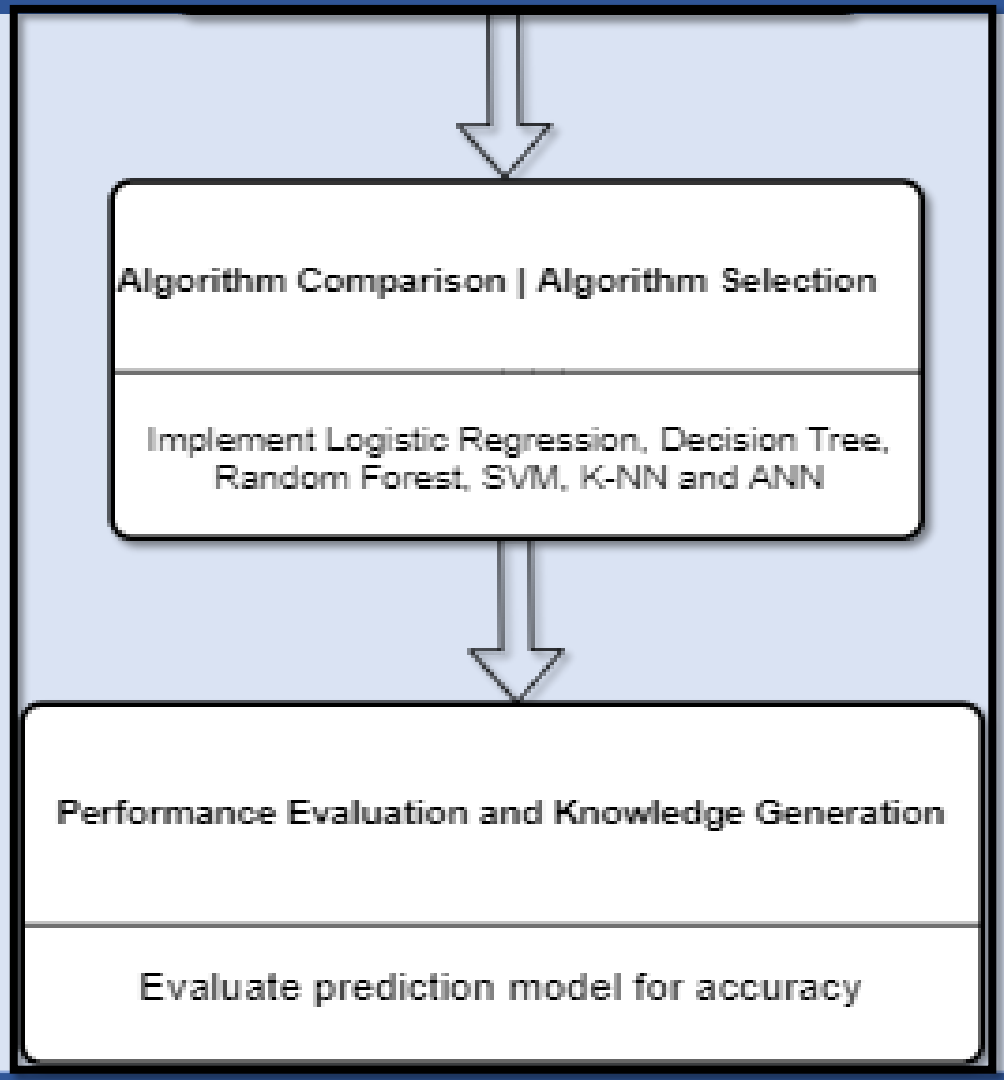
- Data Gathering
- Data Cleaning
- Exploratory Data Analysis
- Algorithm Comparison
- Algorithm Selection
- Prediction



Project Process Diagram

Work Process Diagram





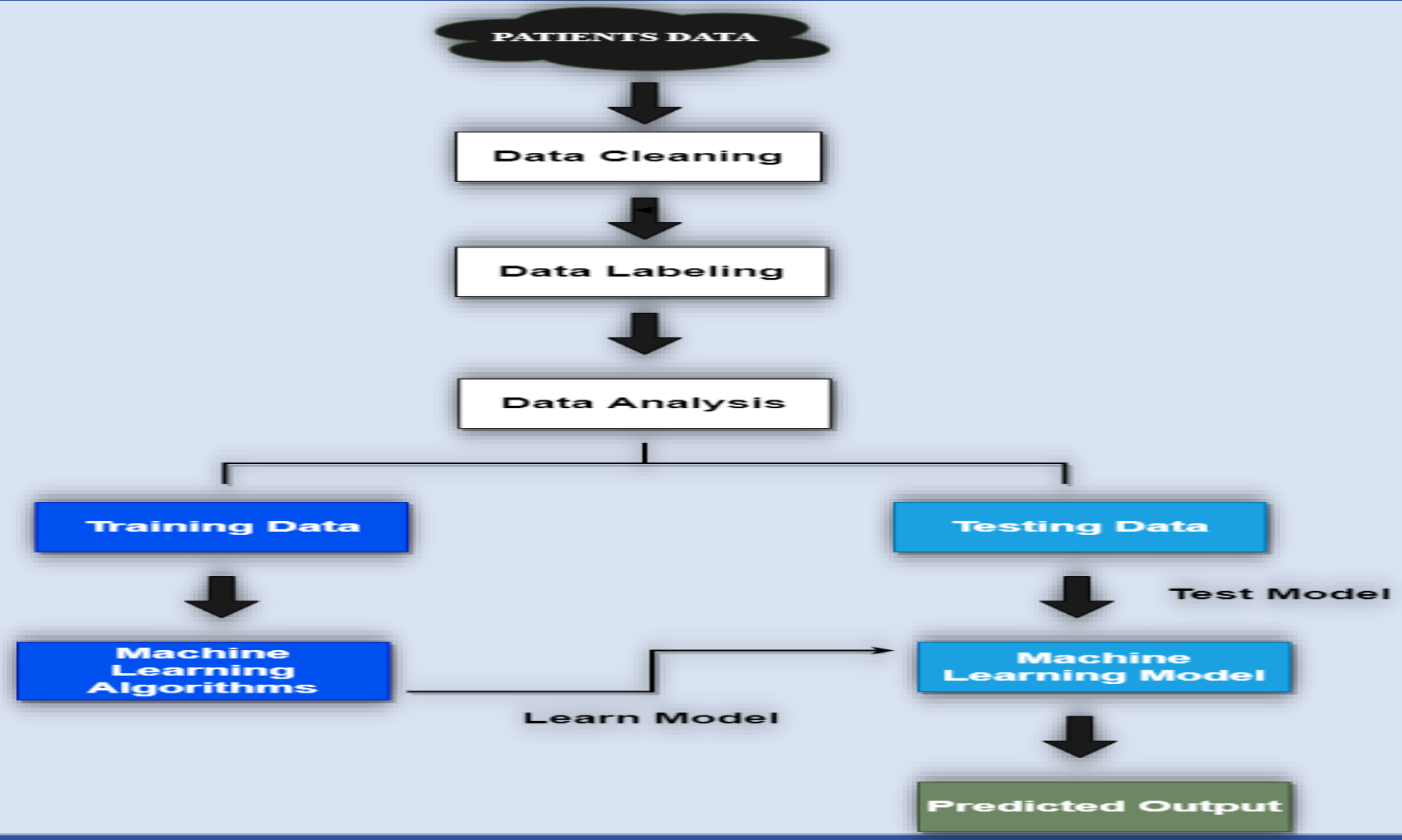
Tools/ Technology

- **Hardware Requirements**
  - ✓ Processor Core i3, 4th Gen
  - ✓ 4 GB RAM
  - ✓ 500 HDD

- **Software Requirements**
  - ✓ Python
  - ✓ Jupyter Notebook (Anaconda 2.7)
  - ✓ MS Word 2013, MS Excel 2013
  - ✓ Diagram Tool

  Libraries:
  - ➢ Pandas
  - ➢ Numpy
  - ➢ Seaborn
  - ➢ Matplotlib
  - ➢ Ipython
  - ➢ Pydotplus

Methodology



Interviews Data Set

- **Interviews Data Set: (as sample)**

➢ Interviews performed by psychologists on different medical appointments with the patients. It contains both psychological and physical substances. Psychological substances represents subjective data like anxiety, irritability or concentration problems. Physical substances represents objective data like no of cigarettes that the patient has smoked in a day and the time in which the patient wake up or went to bed.

Features:

- Mood
- Motivation
- Concentration problems
- Irritability
- Anxiety
- Sleep quality
- No of Cigarettes
- Other drugs
- Wakeup time
- Going to bed time
- Patient's code
- Date
- Alcohol
- Caffeine

Data Cleaning

By applying different algorithms to create predictive models, data cleaning is essential for all data sets. Some data sets having small amount of data. Noise and outliers in data sets can affect the accuracy of our models.

- Renaming the columns to a more recognizable set of labels
- Dropping unnecessary columns in a DataFrame if all values are ‘NaN’ or ‘NULL’
- Changing the index of a DataFrame
- Skipping unnecessary rows in a CSV file

Gathering the Data:

- **Episode Data Set:** Data Set which signified different episode periods of the patients like depression or mania.

	patient	start	end	episode
0	D	2017-07-01 00:00:00	2017-07-24 00:00:00	D
1	D	2017-08-15 00:00:00	2917-09-11 00:00:00	D
2	G	2017-07-24 00:00:00	2017-08-07 00:00:00	D
3	G	2017-09-04 00:00:00	2017-11-01 00:00:00	M
5	M	2017-06-07 00:00:00	2017-07-01 00:00:00	M
6	M	2017-07-14 00:00:00	2017-07-30 00:00:00	D
7	M	2017-09-25 00:00:00	2017-10-10 00:00:00	D

Data Cleaning

- **Interviews Data Set Table: (sample)**

	mood	motivation	attention	irritability	anxiety	sleep_quality	nr_cigarettes
274	0	1	1	1	1	2	0
230	-1	-1	2	1	1	3	24
241	0	-1	2	1	1	1	24
0	2	2	3	3	3	3	34
1	2	2	3	3	3	3	38

caffeine	alcohol	other_drugs	patient	date	active_time
90	0	0	G	2007-07-17 00:00:00	1470
120	0	0	D	2017-01-02 00:00:00	1790
90	0	0	D	2017-01-15 00:00:00	1545
150	0	0	D	2017-06-01 00:00:00	1710
150	0	0	D	2017-06-02 00:00:00	1770

Data Cleaning

Preparing the Data

Initiate by giving the columns proper name. CNT of episode data set:

```
episodes.columns = ['patient', 'start', 'end', 'episode']
```

Interview Data Set:

- *Rename the Columns*
- *Absolute value mapping*
- *Calculate active amount time*
- *Concluding missing values*

```
interviews.columns = ['mood', 'motivation', 'attention', 'irritability', 'anxiety',
                      'sleep_quality', 'menstrual_cycle', 'nr_cigarettes', 'caffeine',
                      'alcohol', 'other_drugs', 'wake up time', 'going to bed time',
                      'patient', 'date']
```

Exploratory Data Analysis

- Exploratory Data Analysis (EDA) is an approach to analyzing datasets to summarize their main characteristics, often with visual methods. It is not easy to look at a column of numbers or a whole spreadsheet and determine important characteristics of the data.
- With the support of different charts and maps, we can get a well understanding of all the data sets as well as the changed combinations of them.
- We plotted each dataset separately. Before plotting, we removed those columns which could not be plotted. It was important to know which variables had more than one value in the dataset because singular matrices cannot be plotted in a dataset.
- EDA classified into two ways. Graphical or non graphical and uni-variate or multi-variate.
- **Visualization Techniques:**

➢ *Histograms.*

➢ *Heatmaps.*

➢ *Scatterplots.*

Visualization

- **Interview dataset as sample:**

- Histogram:

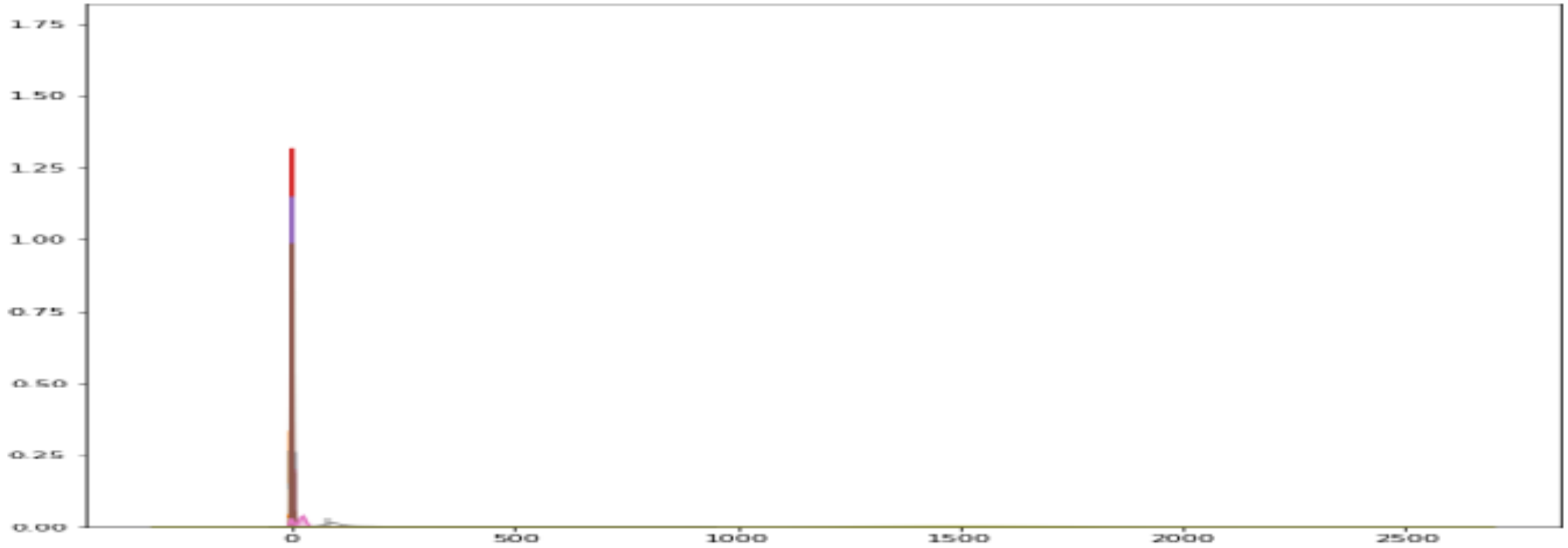


Figure 5.12. Feature distribution in the Interview data set

Visualization

- **Heatmap:**

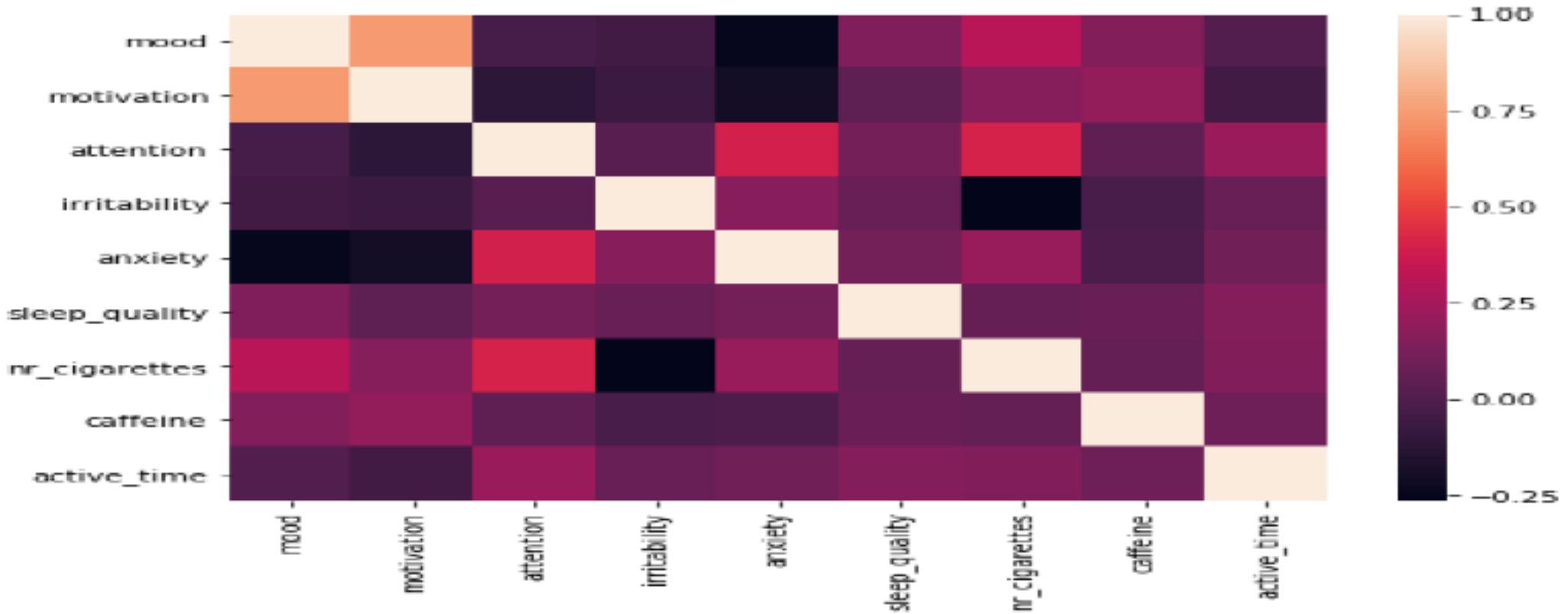


Figure 5.20. Correlation heatmap of Interview data set

Data Combination

Combining the Data Sets:

- Combining the data in different ways to make new datasets with different combination of features.
- The goal was to find dataset that had enough data for the algorithm to process and will return highest accuracy.
- Plotted some correlated features that found in data visualization, which would help us see how the data was distributed depending on which state the patient was in.

❖ Episode and YMRS dataset

❖ Episode and HDRS dataset

❖ Episode and Interviews dataset

❖ Episode and Interventions dataset

❖ Young and Hamilton datasets

❖ Interviews and Interventions datasets

Data Combination

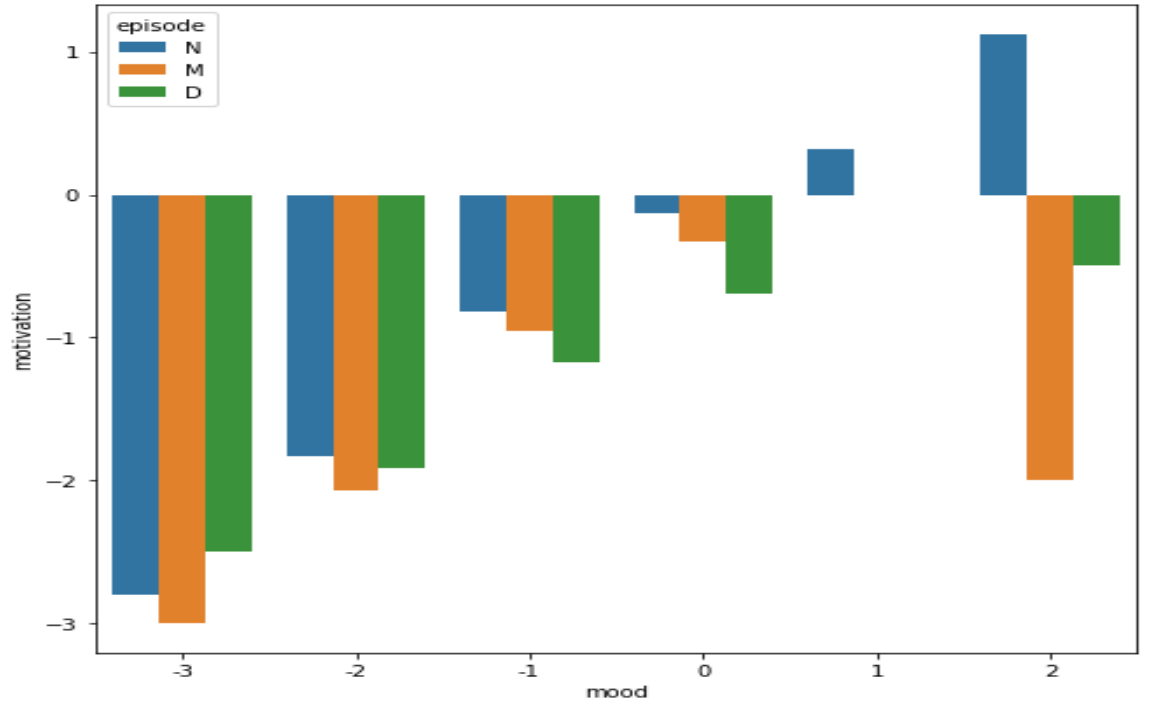
Episodes and Interviews Data Sets: (sample)

	mood	motivation	attention	irritability	anxiety	sleep_quality	nr_cigarettes	caffeine	alcohol	other_drugs	active_time	episode
274	0	1	1	1	1	2	0	90	0	0	1470	N
230	-1	-1	2	1	1	3	24	120	0	0	1790	N
241	0	-1	2	1	1	1	24	90	0	0	1545	N
0	2	2	3	3	3	3	34	150	0	0	1710	N
1	2	2	3	3	3	3	38	150	0	0	1770	N

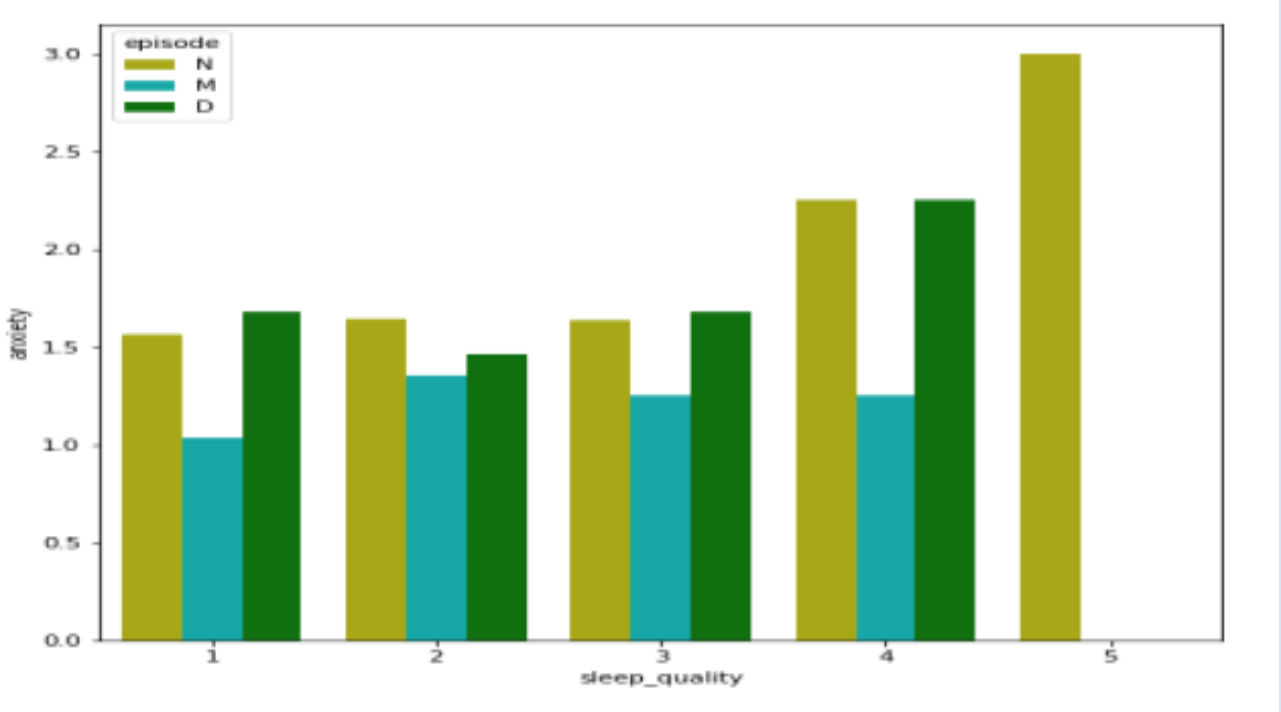
Combining Interview dataset with episode dataset

Data Combination

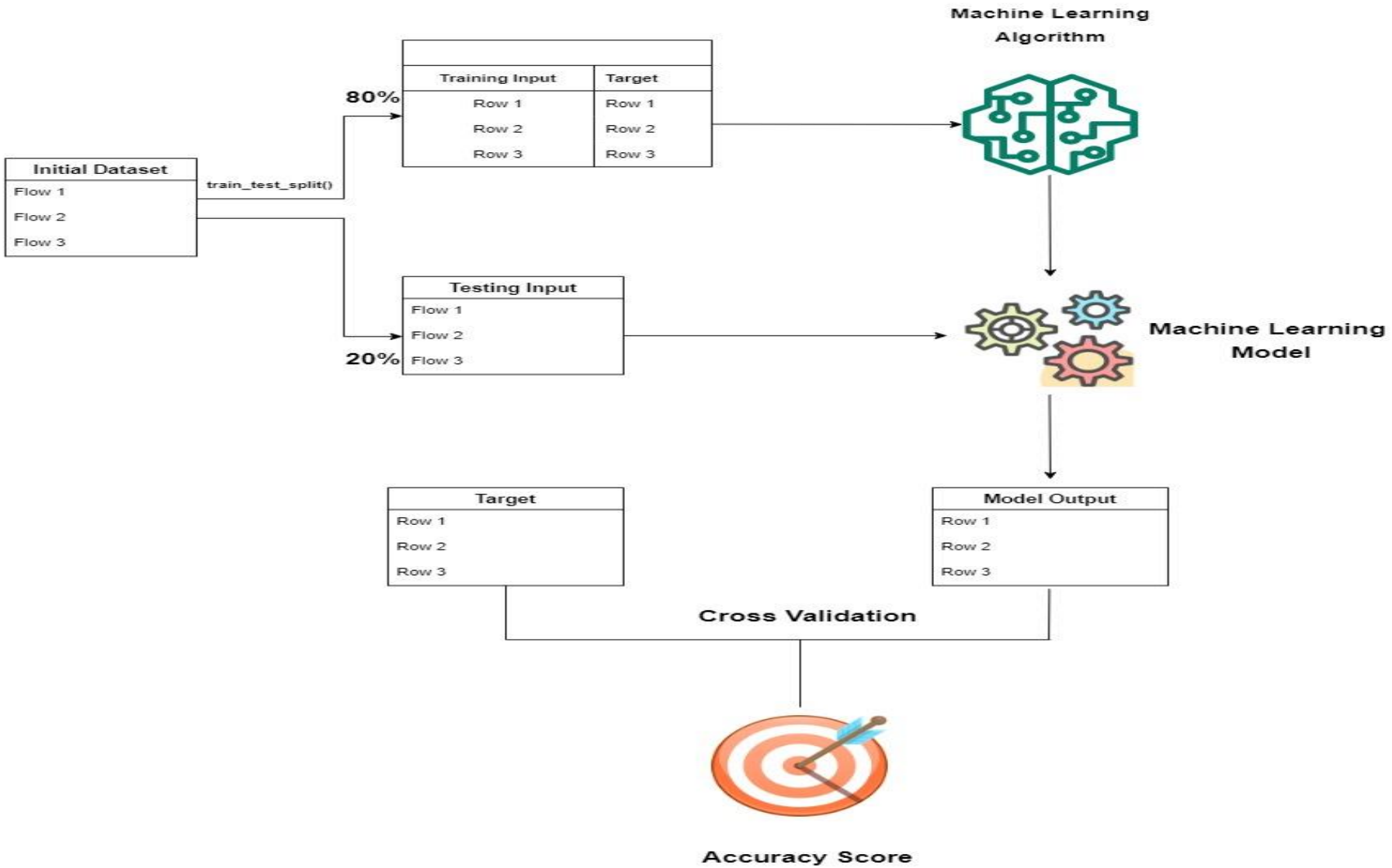
Distribution of mood and motivation on different patient states in Interview dataset



Bar plot of anxiety and sleep quality with patient states in interview dataset



Application Process Diagram



Logistic Regression Algorithm

- Logistic Regression model is the natural logarithm of an odds ratio between two variables called logit. It explained with the logistic function, which takes an input and outputs a number between 0 and 1.

$$\sigma(t) = \frac{e^t}{e^t + 1}$$

- t is a linear function that can be represented as $t = \beta 0 + \beta 1 x$

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{e^{\beta_0 + \beta_1 x} + 1}$$

- The inverse of the Logistic Function:

$$g(x) = ln\left[\frac{\pi(x)}{1 - \pi(x)}\right] = \beta_0 + \beta_1 x$$

Decision Tree Algorithm

- The Decision Tree algorithm was implemented with the *DecisionTreeClassifier*. It uses certain hyper-parameters which can be tuned depending on the size and number of features of a data set.
- Hyper-parameters are:
  - ✓ Max_depth
  - ✓ Min_samples_leaf
- Divide dataset into training and testing test, model training with decision tree classifier:

```
clf = DecisionTreeClassifier(max_depth=5)
clf.fit(X_train, y_train)
```

- Calculate model accuracy score with validation function:

```
Scores = cross_val_score (clf, X_test, y_test)
print "Model accuracy: ",scores.mean()
```

Random Forest Algorithm

- Random Forest Algorithm consists an array of tree predictors in which every tree is depending on value of random vector that is sampled with same distribution for all the trees.
- Implemented with *RandomForestClassifier.* Model training with RM classifier:

```
clf = RandomForestClassifier(n_jobs=-1)
clf.fit(X_train, y_train)
```

- The n_jobs hyper-parameter represents the number of jobs running in parallel for both fitting and predicting.
- The -1 value indicates that the number of jobs is equal to the number of cores of the processor.

Tests on Real Data

- Find state of patient using different indicators:

State:

- ✓ Mania Episode
- ✓ Depression Episode
- ✓ Euthymic Episode

Indicators:

- ✓ Mood
- ✓ Motivation
- ✓ Attention
- ✓ Irritability
- ✓ Anxiety
- ✓ Sleep quality
- ✓ No of cigarettes smoked?
- ✓ Amount of caffeine ingested?
- ✓ When did the patient wakeup?
- ✓ When did the patient go to bed?

```
(DS27) C:\Users\abdul\FBPD - master>python state_prediction.py


-----------------------------------------
  Forecasting BiPolar Disorder           |
-----------------------------------------
         Final Year Project              |
           BS - Computer Sciences        |
-----------------------------------------
 Author : Abdul Baqi, Anas Munzir        |
-----------------------------------------


Please indicate each of the values gathered during the interview:
 Mood (-3 to 3)? 1
 Motivation (-3 to 3)? 3
 Attention (0 to 4)? 2
 Irritability (0 to 4)? 1
 Anxiety (0 to 4)? 4
 Sleep quality (0 to 4)? 2
 Number of cigarettes smoked? 5
 Amount of caffeine ingested? 100
 When did the patient wake up (hh:mm)? 09:30
 When did the patient go to bed (hh:mm)? 23:40


[33mThe patient could be tending towards a DEPRESSION episode[0m
```

```
(DS27) C:\Users\abdul\FBPD - master>python state_prediction.py


----------------------------------------
|  Forecasting BiPolar Disorder         |
----------------------------------------
|         Final Year Project            |
|           BS - Computer Sciences      |
----------------------------------------
| Author : Abdul Baqi, Anas Munzir      |
----------------------------------------


Please indicate each of the values gathered during the interview:
* Mood (-3 to 3)? 2
* Motivation (-3 to 3)? -2
* Attention (0 to 4)? 1
* Irritability (0 to 4)? 2
* Anxiety (0 to 4)? 1
* Sleep quality (0 to 4)? 3
* Number of cigarettes smoked? 20
* Amount of caffeine ingested? 0
* When did the patient wake up (hh:mm)? 10:00
* When did the patient go to bed (hh:mm)? 01:00


[33mThe patient could be tending towards a MANIA episode[0m
```

```
(DS27) C:\Users\abdul\FBPD - master>python state_prediction.py


-----------------------------------------
|   Forecasting BiPolar Disorder          |
-----------------------------------------
|          Final Year Project             |
|            BS - Computer Sciences       |
-----------------------------------------
| Author : Abdul Baqi, Anas Munzir        |
-----------------------------------------


Please indicate each of the values gathered during the interview:
* Mood (-3 to 3)? -2
* Motivation (-3 to 3)? -2
* Attention (0 to 4)? 4
* Irritability (0 to 4)? 3
* Anxiety (0 to 4)? 3
* Sleep quality (0 to 4)? 0
* Number of cigarettes smoked? 30
* Amount of caffeine ingested? 200
* When did the patient wake up (hh:mm)? 05:00
* When did the patient go to bed (hh:mm)? 20:30


[36mThe patient is staying in a EUTHYMIC state[0m
```

Algorithm Performance Table

Algorithms/Datasets	Logistic Regression	Decision Tree	Random Forest	SVM	KNN	Average
YMRS	83%	53%	52%	60%	64%	**63%**
HDRS	70%	45%	45%	51%	57%	**54%**
Interviews	73%	68%	69%	69%	59%	**68%**
Interventions	67%	33%	50%	45%	40%	**47%**
YMRS-HDRS	61%	70%	63%	63%	38%	**60%**
Interviews-Interventions	100%	58%	41%	42%	66%	**62%**
Average	**76%**	**55%**	**54%**	**55%**	**54%**	

Conclusion

- Having understanding of data we are working with ML, where knowing behaviors are normal and abnormal in the patients can help us create much more precise prediction model.
- With different parameters or indicators easily identify actual condition of patient and it’s helpful for psychiatrists to diagnose actual state of patient either depression or mania or euthymic state.
- Proposed techniques are helpful to get good accuracy score from datasets in this project.
- In future, train the same algorithms with large amount of data, to see if they perform in similar way.
- In future, gathering objective data from patients through devices like mobile phones, wrist bands. The purpose of this recommendation would be compare the performance of algorithms on objective data gathered from devices with the performance results obtained on subjective data used in this project.

References

1. M. Bonsall et al., Bipolar disorder dynamics: affective instabilities, relaxation oscillations and noise, vol. 12, Journal of The Royal Society Interface, The Royal Society, 2015.

2. Kraepelin E. Manic-depressive insanity and paranoia. New York: E & S Livingstone; 1921

3. https://searchenterpriseai.techtarget.com/definition/machine-learning-ML

4. https://searchenterpriseai.techtarget.com/definition/deep-learning-deep-neural-network

5. www.bip4cast.org (Last accessed: 23/05/2018).

6. J. Anchiraico, Diseño de una Arquitectura Big Data para la Predicción de Crisis en el Trastorno Bipolar, Trabajo de Fin de Master, E-Prints Universidad Complutense de Madrid, 2016.

7. A signature-based machine learning model for distinguishing bipolar disorder and borderline personality disorder' published by National Center for Biotechnology Information. link: https://pubmed.ncbi.nlm.nih.gov/30546013/

References

8. Design of BD Architecture for Prediction of Crises in Bipolar Disorder' link: https://www.mdpi.com/2227-7390/9/11/1174/pdf

9. A. Martínez, Introduction to Big Data and First Steps in a Big Data Project, Trabajo de Fin de Grado, Madrid: E-Prints Universidad Complutense de Madrid, 2016.

10. Python Software Foundation, Python Language Reference, version 2.7 (http://www.python.org).

11. F. Pérez and B.E. Granger, IPython: A System for Interactive Scientific Computing, Computing in Science & Engineering, 2007.

12. W. McKinney, Data Structures for Statistical Computing in Python, Proceedings of the 9th Python in Science Conference, 2010.

13. J. D. Hunter, Matplotlib: A 2D Graphics Environment, Computing in Science & Engineering, 2007.

14. F. Pedregosa et al., Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 2011.

15. Research paper by P.J Moore published by National Library of Medicine. Link: https://pubmed.ncbi.nlm.nih.gov/22855220/

References

16. DL by Depression Detection. Link: https://aclanthology.org/W18-0609.pdf

17. https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0148195&type=printa%5b7%5dble

18. An Item Response Theory Evaluation of the Young Mania Rating Scale in the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). link: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5048523/ . https://www.semanticscholar.org/paper/DeepMood%3A-Forecasting-Depressed-Mood-Based-on-via-Suhara-Xu/a656cab6ec7f825eb54de24f8224246e3acef31b?p2df.

19. https://www.researchgate.net/publication/350639348_Review_on_Machine_Learning_Techniques_to_predict_Bipolar_Disorder

20. https://www.researchgate.net/publication/311336418_Classification_of_Schizophrenia_and_Bipolar_Disorder_by_Using_Machine_Learning_Algorithms?enrichId=rgreq-7be31377809c41f3a35a86e966fbbe1e-XXX&enrichSource=Y292ZXJQYWdlOzMxMTMzNjQxODtBUzo1MDg0MjA2Nzc5MzEwMDhAMTQ5ODIyODM0MzAyMg%3D%3D&el=1_x_2&_esc=publicationCoverPdf

21. Python string to datetime - strptime() – JournalDev

References

22. https://analyticsindiamag.com/5-ways-handle-missing-values-machine-learning-datasets/

23. Detection of Bipolar Disorder Using Machine Learning with MRI. URL: http://ceur-ws.org/Vol-2786/Paper53.pdf

24. https://www.slideshare.net/MohammedARajab/bipolar-disorder-59525559

25. https://www.slideshare.net/vaishnavisnair/manic-depressive-psychosis-bipolar-disease

26. Implementing machine learning in bipolar diagnosis. URL: https://pubmed.ncbi.nlm.nih.gov/31740657/

27. Garima, H. Gulati and P. K. Singh, Clustering techniques in data mining: A comparison, New Delhi: 2015 2nd International Conference on Computing for Sustainable Global Development, 2015, pp. 410-415.

28. J. R. Quinlan, Induction of Decision Trees, vol. 1, Hingham, Massachusetts: Kluwer Academic Publishers, 1986, pp. 81-106.

29. L. Breiman, Random Forests, vol. 45, Berkeley, California: Machine Learning, 2001, pp. 5-32.

30. C. Hsu, C. Chang and C. Lin, A Practical Guide to Support Vector Classification, Taipei: National Taiwan University, 2003.

Thank You!