



UTILIZING MACHINE LEARNING ALGORITHMS TO PREDICT THE SEVERITY OF MAJOR DEPRESSIVE DISORDER

Monal H. Jain

Research Scholar: Computer Engineering Department
LDRP-ITR, KSV University
Gandhinagar, India

Mehul P. Barot

Assistant Professor: Computer Engineering Department
LDRP-ITR, KSV University
Gandhinagar, India

Abstract- Health care area is rising immensely from last few decades. It creates large amount of records about patients, clinical sources, infection analysis, electronic patient records, curative trials etc. Huge size records are a key source which needs preprocessing and evaluation for information mining that assists backing for cost-savings and verdict constructing. When medical organizations apply data mining on their prevailing data, they can realize new, beneficial and possibly life-saving information. Heterogeneity of major depressive disorder (MDD) disease progression muddles medical management. It can pretense dares for clinicians concerning correct analysis and actual well-timed cure. These challenges have provoked the increase of many machine learning techniques to aid advance the controlling of this disease. Machine learning (ML) is an application of artificial intelligence (AI) that offers structures the capability to mechanically acquire and improve from experience minus being explicitly programmed. Far-reaching data mining practices utilizing machine learning procedures have potential in the study of big epidemiological datasets. The following proposed research illustrates the idea of machine learning algorithms to analyze the depression epidemiological datasets in terms of its severity. Here, mean value substitution is performed on the data to preprocess it and then multiclass classifier machine learning algorithms are practiced to know the severity of the depression among the people at an early stage so that proper and correct timely treatment can be provided which will result in cost saving and faster improvement of health conditions of patients.

Keywords - Machine Learning, Major Depressive Disorder, severity, data preprocessing, multiclass classification.

I. INTRODUCTION

Major Depressive Disorder (MDD) is a psychological health syndrome defined by insistently depressed mood or loss of interest in activities, creating main impairment in everyday life. Machine learning (ML) is an application of artificial intelligence (AI) that offers structures the capability to mechanically acquire and improve from experience minus being explicitly programmed. Machine learning algorithms can be utilized to evaluate the depression epidemiological datasets in the usage of self-reports done by the patients so that the severity of the MDD is recognized which lets us to take judgments of choosing the type of specific cure to be delivered to the patients resulting in the initial accurate treatment and price saving to the patients.

Depression is a severe health syndrome, with the World Health Organization approaching 350 million individuals universally under the influence of depression, also standing second in positions of global disability problem and depression probable to be the figure one health worry in equally developed and developing nations by 2020 [1]. Major depressive disorder has an assessed life span occurrence of about 17%, and has notable effects on class of life, comorbid health disease, suicide likelihood, and healthcare job.

Depression is an intricate medical thing that can produce tasks for clinicians regarding equally precise analysis and in effect well-timed cure. These tasks have aggravated the advancement of many machine learning techniques to aid the controlling of this disease.

In the Major Depressive Disorder (MDD) illness, much research has been carried out after the MDD is

identified and originates severe which creates the necessity to provide the cure to the specific. Generally, the outcomes are examined and anticipated on the grounds of fMRI done on the enduring individual. Not much research has been carried out on the preliminary phase of MDD where specific reports himself. Research grounded on the self-reports by applying machine learning techniques can let us to predict the MDD at primary stage. Also because of this one can evaluate and predict the severity of MDD existence in the specific who benefits the specific to take precautions and timely treatment to get cured early.

II. MAJOR DEPRESSIVE DISORDER

Major Depressive Disorder (MDD) is a psychological health syndrome described by persistently depressed mood or loss of interest in activities, producing major impairment in day-to-day life. MDD is a severe psychological state that along with the suffering immense sensitive agony, also furthermore eventually can lead to demise by suicide.

Major Depressive Disorder (MDD) is a worldwide problem. Agreeing to the World Health Organization (2004), unipolar depression is on position three globally on the encumbrance list of malady. It orders eighth among less developed nations and first in developing and highly developed nations. Depression orders upper than ischemic heart illness, violence, strokes, and transportation accidents. Even though MDD is frequently intermittent and lasting, persons may only search for aid while in the middle of a depressive incident. After recovering up to their level of functionality, they may put an end to cure until the cycle begins another time.

The DSM-IV TR (American Psychological Association, 2000) finds and talks over the analytical conditions. To have the conditions for MDD, minimum five of the signs enumerated beneath need to exist in the similar 2-week duration, and must characterize a transformation in the distinct. The signs should also cause major sorrow or loss in public, profession, or other fields [2]. The signs shouldn't be the reason of utilization of any kind of stuff, a health state, another psychological syndrome, or grief [2]. Also, the specific can't have hyper in the past, hypomanic, or mixed incidence [2]. It is furthermore essential to annotate that those with MDD are at a greater than before threat for evolving co-morbid mental disorder such as nervousness, frights, whim control, or substance misuse.

▪ **Symptoms[2]**

Onset of a Major Depressive Episode can be anywhere from days to a few weeks. Symptoms must be present for most of the day, nearly every day, for at least 2-weeks and include:

- Depressed Mood
- Diminished interest/pleasure
- Significant weight loss or gain
- Insomnia or hypersomnia
- Psychomotor agitation or retardation
- Fatigue
- Feelings of worthlessness or guilt
- Difficulty concentrating
- Recurrent thoughts of death/suicidal ideation

Depressed mood: Depressed mood is considered as the promising indicator of MDD. It is defined as sense of unhappiness, downhearted, or discouraged. Few individuals don't have the ability to describe their emotions while other few individuals don't want to agree of sensing depression. Few individuals might agree of having bad temper along with or instead of having depressed mood. It is significant to note the effect of the person, giving close consideration to the reactions on face, stance, and tendency of speech. This is specifically significant if the individual is in rejection about his or her mental state.

Loss of interest in activities: Depressed persons will frequently observe that they are not finding things interesting which they used to enjoy earlier. Some call it as not viewing frontward to things, or being incapable to feel pleasure. It might come into notice of some people that the affected individual is isolating himself from his family and friends. A loss of sexual desire may also be noticed.

Weight changes: Hungriness varies bring about in noteworthy, unplanned variation in weight are a lot observed in MDD. This might be seen as craving for fat and carbohydrate foods in some beings while appetite loss in other beings.

Sleep changes: Sleeplessness is mostly observed in MDD. Beings may discover themselves awake in the midnight and are not capable to again asleep. Others may lie wakeful, unable to initiate forty winks. In some occurrences, hypersomnia happens.

Psychomotor changes: Psychomotor changes should be observable by others and take account of agitation causing in walking, leg rebounding, or twitching. Also, retardation may mark as speaking and moving slowly or delay in reacting to questions.

Fatigue: Too much fatigue is often a indication that really influences the subject. He or she may have non existence of the liveliness to do the day-to-day chores of existing. Weariness, even after a whole dusk's nap, is common.

Feelings of worthlessness or excessive or inappropriate guilt: Those be subjected to a major depressive episode repeatedly dock strong feelings of insignificance and culpability. Individual may sense worthless of the belongings in their existence. They may preoccupy and know-how strong culpability over former or existent happenings. Individual also might adversely misjudge things spoken or done by others. This continues the culpability and feelings of contemptibility.

Indecisiveness and concentration problems: Individuals having MDD repeatedly experience trouble focusing on responsibilities. This must be an alteration from usual working. Concentrating problems also ascend, and can consequence in professional, scholastic, and association difficulties. Creating choices can unexpectedly seem mammoth to a blue individual.

Recurrent thoughts of death and/or suicide: The major fear with Major Depressive Disorder is that of suicide. Considerations of death are a common incidence in the distressed. These views may differ liable on the sternness of the depression. It is further vis-à-vis if the person has made a strategy of how he or she would do suicide. Regrettably, some follows the strategy and take their own breathes.

A. PHQ-9 as a Diagnostic Measure

At 9 objects, the PHQ (Patient Health Questionnaire) depression rule (which we call the PHQ-9) is depression methods, has sensitivity and specificity, and comprises of the genuinenine criteria on which the analysis of DSM-IV depressive disorders is grounded. The PHQ-9 is a two-purpose implement that, with the similar nine objects, can form transitory depressive disorder analyzes as well as score depressive indicator rigorousness. The PHQ-9 symptom checklist is shown in Figure 1.

Nine Symptom Checklist

Over the last 2 weeks, how often have you been bothered by any of the following problems?

	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things.....	0	1	2	3
2. Feeling down, depressed, or hopeless.....	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much.....	0	1	2	3
4. Feeling tired or having little energy.....	0	1	2	3
5. Poor appetite or overeating.....	0	1	2	3
6. Feeling bad about yourself - or that you are a failure or have let yourself or your family down.....	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television.....	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed? Or the opposite - being so fidgety or restless that you have been moving around a lot more than usual.....	0	1	2	3
9. Thoughts that you would be better off dead or of hurting yourself in some way.....	0	1	2	3

(For office coding: Total Score ____ = ____ + ____ + ____)

Figure 1: PHQ-9 Symptom Checklist[3]

As a severity degree, the PHQ-9 markarrays from 0 to 27, asevery of the 9 items can be counted from 0 (“not at all”) to 3 (“nearly every day”). Easy-to-remember change points of 5, 10, 15, and 20 signify the onsets for mild, moderate, moderately severe, and severe depression, respectively [3]. Advisedcureactivities in reaction to these differentstages of PHQ-9 depression severity are shown in Table1. Scores less than 10 seldom occur in individuals with major depression whereas scores of 15 or greater usually signify the presence of major depression.

TABLE 1: PHQ-9 MARKS&RECOMMENDEDCURE [3]

PHQ-9 Score	Depression Severity	Proposed Treatment
1 to 4	None	None
5 to 9	Mild	On the lookout coming up, PHQ-9 follow up
10 to 14	Moderate	Strategize cure, Think throughtherapy
15 to 19	Moderately Severe	Abruptbeginning of pharmacotherapy/psychothe rapy
20 to 27	Severe	If severe diminishing or lowlyreaction to psychotherapy, talk about to psychological health professional for psychiatric therapy and combinedmanaging

III. MACHINE LEARNING

One of the subareas of artificial intelligence (AI) is Machine Learning[4]. The objective of machine learning commonly is to realize the structure of data and fit that data into models that can be understood and exploited by people. Machine learning systems in its place let computers to train on data inputs and practice statistical analysis in order to output values that fall inside a particular range. As of this, machine learning enables computers in building models from sample data in order to mechanize decision-making procedures centered on data inputs. Two of the most broadly accepted machine learning approaches is **supervised learning** which trains algorithms based on example input and output data that is labeled by humans, and **unsupervised learning** which provides the algorithm with no labeled data in order to allow it to find structure within its input data.

- **Supervised Learning**

In supervised learning, the computer is delivered with sample inputs that are characterized with their preferred outputs. The determination of this technique is for the algorithm to be capable to “learn” by equaling its real output with the “taught” outputs to discover errors, and transform the model accordingly. Supervised learning therefore uses patterns to guess label values on additional unlabeled data. A common use case of supervised learning is to utilize past data to predict statistically probable future happenings.

- **Unsupervised Learning**

In unsupervised learning, data is unlabeled, so the learning algorithm is left to discover commonalities among its input data. As unlabeled data are more ample than labeled data, machine learning methods that assist unsupervised learning are principally valuable. The objective of unsupervised learning may be as direct as determining hidden patterns inside a dataset, but it may also have a objective of feature learning, which lets the computational machine to mechanically learn the representations that are required to categorize raw data.

Learning approaches can be classified into linear and nonlinear approaches. Linear approaches are simpler, while nonlinear approaches are more flexible in behavior. For supervised learning, the methods can be additionally classified as classification- or regression-based methods. Classification-based methods try to classify the data by discrete and categorical labels, while regression-based methods fit the data to a continuous function and thus work with continuous labels for the data [5]. For unsupervised learning, the methods are primarily categorized as clustering method, which group the data into clusters based on underlying similarities [5].

IV. BACKGROUND STUDY

In the current works,a nature motivated and unique FS algorithm grounded on standard Ant Colony Optimization(ACO), known as improved ACO(IACO), was utilized to lessen the number of attributes by eliminating unrelated and out of work information [6]. The nominated attributes were then providing for into support vector machine (SVM), in order to categorize MDD and BD subjects [6]. Quantitative electroencephalography (QEEG) coherence data of 46 BD and 55 MDD subjects were fed into IACO first [6]. Selected more informative feature subset from 16 electrodes of alpha, delta and theta frequency bands was then used as input in SVM [6]. Enactment of IACO–SVM methodology identified that it is promising to distinguish 46 BD 55 MDD themes using 22 of 48 attributes with 80.19% total taxonomy accuracy (Turker Tekin Erguzel et al., 2015).

Current works also determine that guilt-selective modifications in functional connectivity of the ATL are satisfactory to differentiate the remitted MD set from the control set with in elevation accuracy [7]. The outcomes demonstrate that guilt-selective functional disconnection of the ATL has the impending to be moreestablished into a clinically beneficial fMRI biomarker of MD liability [7]. Expending a lately recognized neural signature of guilt-selective functional disconnection, the machine learning classification procedure was capable to differentiate remitted MD from control members with 78.3% accuracy [7]. This determines the high possibility of our fMRI signature as a biomarker of MD vulnerability (João R. Sato et al., 2015).

Cognitive models predict that vulnerability to major depressive disorder (MDD) is due to a bias to blame oneself for failure in a global way resulting in excessive self-blaming emotions, decreased self-worth, hopelessness and depressed mood [8]. 132 patients with remitted MDD and no relevant lifetime comorbid axis-I disorders were assessed using a phenomenon

logical psychopathology based interview (AMDP) including novel items to assess moral emotions (n=94 patients) and the structured clinical interview-I for DSM-IV-TR [8]. Cluster analysis was engaged to recognize indicator rationality for the most severe incidence. As anticipated, moods of insufficiency and uselessness were part of the central depressive disorder, closely co-occurring with depressed frame of mind. Self-blaming feelings were very recurrent and concerning but not limited to guilt [8]. This calls for a sophisticated valuation of self-blaming sentiments to progress the identification and stratification of MDD (Roland Zahn et al., 2015).

Current works also utilized a three-step procedure fusing multiple imputation, a machine learning boosted regression procedure and logistic regression, to find key biomarkers related with depression in the NHANES (2009–2010) [1]. After the formation of 20 imputation data collections from multiple chained regression sequences, machine learning boosted regression at first identified 21 biomarkers related with depression [1]. Using traditional logistic regression methods, counting monitoring for probable confounders and moderators, an ultimate set of three biomarkers were nominated [1]. The methodical utilization of an amalgam procedure for variable choice, merging data mining procedures consuming a machine learning procedure with traditional statistical modeling, accounted for lost information and composite survey selection approach and was validated to be a suitable tool for spotting three biomarkers related with depression for future theory generation: red cell distribution width, serum glucose and total bilirubin (Joanna F. Dipnall et al., 2016).

Also, the WMH ML prototypes were practiced to this baseline information to produce expected result scores that were matched to detected marks measured 10–12 years after baseline [9]. ML prototypical expectation accuracy was also equaled to that of conventional logistic regression models [9]. Area under the receiver operating characteristic curve (AUC) centered on ML (.63 for high chronicity and .71–.76 for the other probable outcomes) was dependably greater than for the logistic models (.62–.70) even though the later models with more predictors [9]. 34.6–38.1% of individuals with consequent high persistence-chronicity and 40.8–55.8% with the severity signs were in the top 20% of the baseline ML predicted risk distribution, while only 0.9% of respondents with consequent hospitalizations and 1.5% with suicide tries were in the lowermost 20% of the ML predicted risk distribution [9]. Outcomes approve that clinically beneficial MDD risk stratification representations can be produced from baseline patient self-reports and that ML approaches advance on conventional procedures in emerging such models (Ronald C. Kessler et al., 2016).

V. METHOD

A. Study participants and measurements

In this research study two different datasets were utilized. One dataset is collected from the National Health and Nutrition Examination Survey (NHANES) (2015–2016) were utilized. Relevant NHANES data files were downloaded from the website. NHANES is a cross-sectional, population-based study of civilians aged 18 to 80 years, conducted in two-year blocks. The sample size of this dataset is 5735[10].

Second dataset is collected from LDRP-ITR college of KSV University, Gandhinagar(INDIA). The respondents aged 16 to 25 years in the study. The sample size of this dataset is 342.

The Patient Health Questionnaire-9 (PHQ-9) was cast-off to evaluate depressive indicators(depression). The PHQ-9 is a well-validated, self-report instrument for discovering and observing depression, with worthy concordance with a medical analysis of major depressive disorder (MDD). Things evaluate the existence of nine Diagnostic and Statistical Manual of Mental Disorder Fourth Edition (DSM-IV) depression indicators over the past two weeks, and are counted on a four-point scale representing the grade of severity from 0 (not at all) to 3 (nearly every day). Things were then totaled to form a complete severity mark going from 0 to 27 where those with a over-all score of 10 or more were measured depressed (i.e. moderately to severely depressed).

B. Methodology

The two datasets collected for the research study contains missing values, which needs to be processed after which the ML classifiers are implemented to predict the severity of MDD.

Data Preprocessing

In several ‘big data’ circumstances, absent information are not a subject due to the large size of observations and variables or features existing. In difference, absent data in readings with lesser model amounts can influence the results greatly. There have been a several missing data techniques recommended in the works over the last years: listwise deletion; pairwise deletion; mean substitution; regression imputation; Maximum Likelihood (ML) estimation. Most of these techniques can only be cast-off when there is no arrangement for the missing data. The selection of technique cast-off for handling with missing data is frequently not as much of significant when the ratio of missing data is less than 5%. Besides, it is not rare for the ratio of missing data in big epidemiological readings to surpass this fraction.

Mean substitution is a common strategy for addressing the missing value problem. In the research, attribute mean value is used to replace all the missing values in both the datasets.

ML classifiers

Machine learning classifiers are implemented to predict the severity of MDD present among respondents. MDD is assessed using PHQ-9. PHQ-9 measures the severity based on the 0-27 score as none (0-4 score), mild (5-9 score), moderate (10-14 score), moderately severe (15-19) and severe (20-27 score). Machine learning multiclass classifiers can be implemented to predict the type of severity of MDD present among respondents. Several multiclass classifiers such as Support Vector Machine (SVM) classifier, Decision Tree, k-Nearest Neighbors, Naïve Bayes classifier can be implemented to classify the severity of MDD. These classifiers can be cast-off for multiclass classification issue by extending the binary classification methods.

- **Support Vector Machine**

Support Vector Machines are amid the best robust and effective classification procedures [11]. They are centered upon the notion of exploiting the margin i.e. maximizing the minimum space from the separating hyperplane to the closest sample [11]. The elementary SVM provisions only binary classification, but additions have been recommended to handle the multiclass classification instance as well [11]. In these additions, extra factors and restraints are added to the optimization problem to handle the separation of the unlike classes [11].

- **Decision Tree**

Decision trees are a powerful classification technique. Two broadly recognized procedures for constructing decision trees are Classification and Regression Trees and ID3/C4.5 [11]. The tree attempts to deduce a split of the training information created on the values of the accessible attributes to create a good simplification [11]. The split at every nodule is centered on the attribute that provides the maximum information gain [11]. Every leaf nodule relates to a class label [11]. A different instance is classified by following a path from the root nodule to a leaf nodule, where at every nodule a test is implemented on certain attribute of that instance [11]. The leaf node reached is considered the class label for that example. The algorithm can naturally handle binary or multiclass classification problems. The leaf nodes can refer to either of the K classes concerned [11].

- **k- Nearest Neighbors**

kNN is considered among the oldest non-parametric classification algorithms [11]. To classify an unidentified instance, the expanse (using some distance measure e.g. Euclidean) from that instance to every further training instance is evaluated [11]. The k minimum distances are recognized, and the most characterized class in these k classes is deliberated the output class label [11]. The value of k is normally determined using a validation set or using cross-validation [11].

- **Naïve Bayes**

Naive Bayes is an effective classifier centered upon the norm of Maximum A Posteriori (MAP) [11]. Assumed a problem with K classes $\{M_1, \dots, M_K\}$ with so-called prior probabilities $P(M_1), \dots, P(M_K)$, we can allot the class label m to an unidentified instance with attributes $a = (a_1, \dots, a_N)$ such that $m = \text{argmax}_m P(M=m|a_1, \dots, a_N)$, that is select the class with the maximum a posterior probability assumed the experimental data [11]. This aposterior probability can be expressed, utilizing Bayes theorem, as follows: $P(M=m|a_1, \dots, a_N) = P(M=m)P(a_1, \dots, a_N|M=m)/P(a_1, \dots, a_N)$ [11]. As the denominator is the same for all classes, it can be let go from the evaluation. Now, we should calculate the so-called class conditional probabilities of the attributes assumed the available classes [11]. This can be quite challenging taking into explanation the dependencies among attributes [11]. The Naive

Bayes methodology is to assume class conditional independence i.e. x_1, \dots, x_N are independent specified the class [11]. This shortens the numerator to be $P(M = m)P(a_1|M = m) \dots P(a_N|M = m)$, and then selecting the class m that maximizes this value over all the classes $m = 1, \dots, K$. Evidently this method is certainly extensible to the instance of having more than two classes, and was presented to implement well inspite of the original simplifying supposition of conditional independence [11].

The performance of the above mentioned ML multiclassifiers to predict the severity of MDD among the individuals are assessed by utilizing Data mining tool WEKA. **Waikato Environment for Knowledge Analysis (Weka)** is a collection of machine learning software written in Java, established at the University of Waikato, New Zealand [12]. Weka is a gathering of machine learning procedures for data mining errands [13]. The procedures can either be practiced straight to a dataset or called from your own Java code [13]. Weka comprises tools for data pre-processing, classification, regression, clustering, association rules, and visualization [13]. It is also well-suited for evolving novel machine learning patterns. From the results of Weka, the suitable ML multiclass classifier can be implemented on the datasets to obtain better results of classification of severity of MDD among respondents.

VI. RESULTS

WEKA tool is utilized to implement ML multiclass classifiers on the two datasets to predict the severity of the MDD. Results are shown below.

NHANES dataset

Logistic, SVM, k-NN, Naïve Bayes and Decision Table multi class classifiers are applied to the dataset in WEKA. As shown in figure, Logistic classifier classified the most correct instances (records) followed by SVM, k-NN, Naïve Bayes and Decision Table. Also, Logistic results have high accuracy, precision, recall, F-measure, TP rate and Kappa Statistic followed by SVM, k-NN, Naïve Bayes and Decision Table. While Logistic Classifier have the lowest incorrectly classified instances and FP rate followed by SVM, k-NN, Naïve Bayes and Decision Table. The results are shown in below graph.

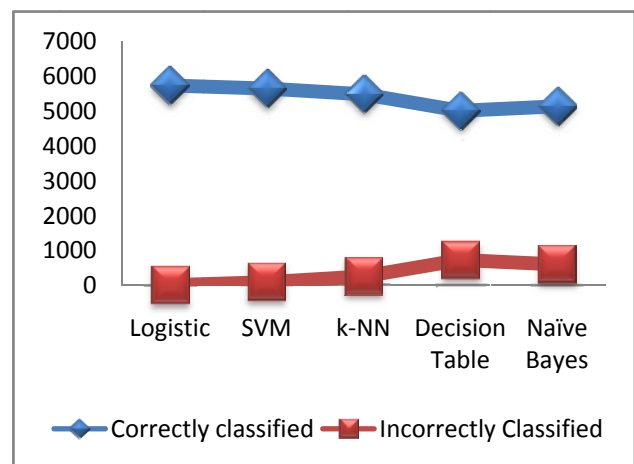


Figure 2: Graph of correctly and incorrectly classified instances by ML classifiers

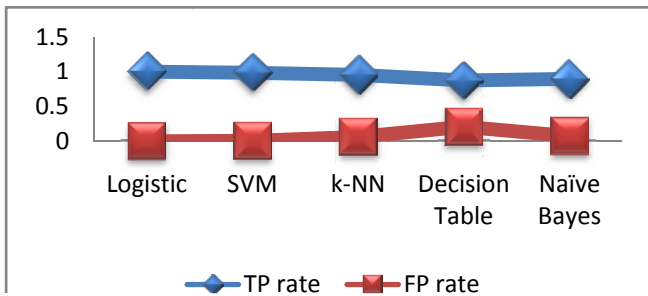


Figure 3: Graph of TP rate & FP rate ML classifiers

Average precision and F measure isn't calculated for SVM, Decision Table and Naïve Bayes as there are small no of instances. The results are shown in below graph.

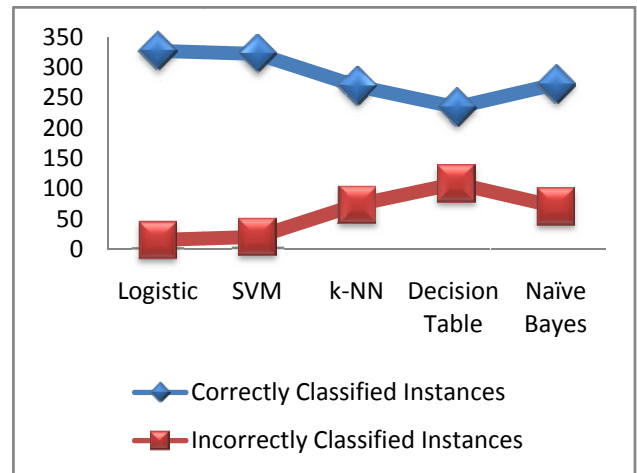


Figure 6: Graph of correctly and incorrectly classified instances by ML classifiers

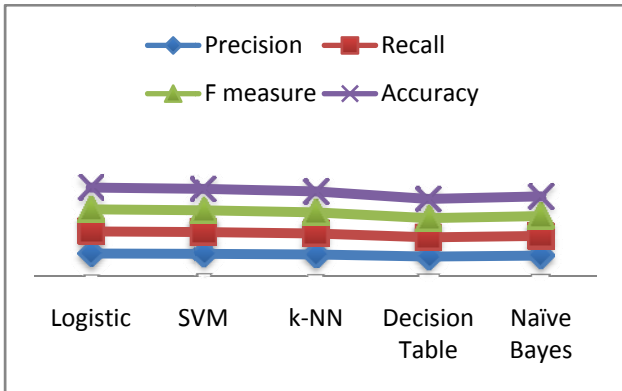


Figure 4: Graph of precision, recall, F measure & accuracy of ML classifiers

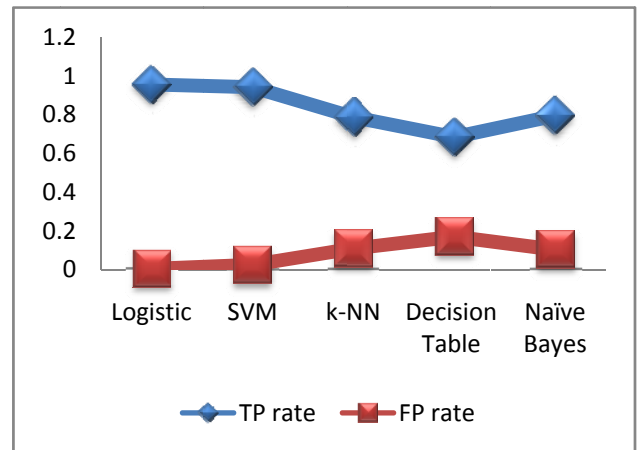


Figure 7: Graph of TP rate & FP rate ML classifiers

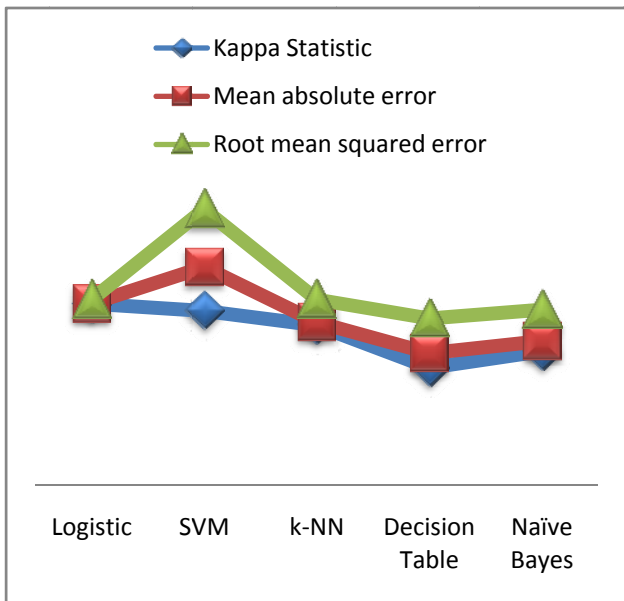


Figure 5: Graph of Kappa Statistic, Mean absolute error & Root mean squared error of ML classifiers

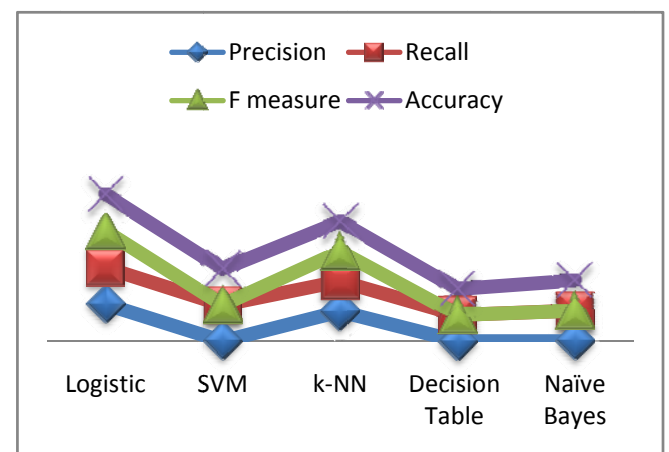


Figure 8: Graph of precision, recall, F measure & accuracy of ML classifiers

LDRP-ITR dataset

Logistic, SVM, Naïve Bayes, k-NN and Decision Table multi class classifiers are applied to the dataset in WEKA. As shown in figure, Logistic classifier classified the most correct instances (records) followed by SVM, Naïve Bayes, k-NN and Decision Table. Also, Logistic results have high accuracy, recall, TP rate and Kappa Statistic followed by SVM, Naïve Bayes, k-NN and Decision Table. While Logistic Classifier have the lowest incorrectly classified instances and FP rate followed by SVM, Naïve Bayes, k-NN and Decision Table.

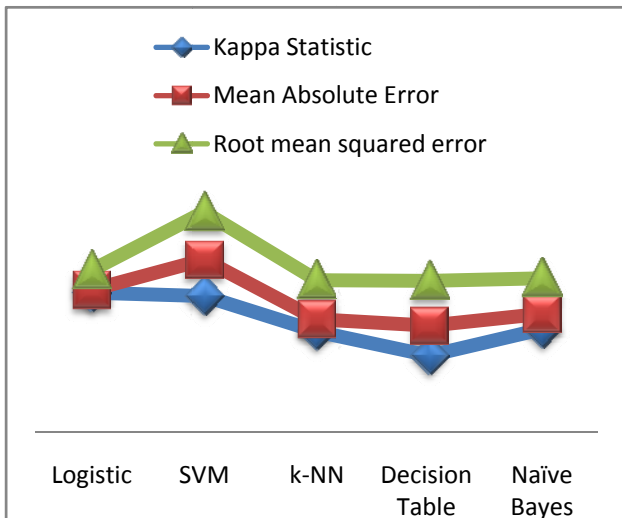


Figure 9: Graph of Kappa Statistic, Mean absolute error & Root mean squared error of ML classifiers

VII. CONCLUSION

Thus MDD severity can be assessed on the epidemiological dataset based on the baseline reports using ML multiclass classifiers. Here data is preprocessed using mean value substitution, after which evaluation of ML classifiers on the preprocessed data to predict the severity of MDD can be done using Weka. Also in this research study, the comorbidity is not considered which can be taken into consideration for future work.

VIII. REFERENCES

- [1] Joanna F. Dipnall, Julie A. Pasco, Michael Berk, Lana J. Williams, Seetal Dodd, Felice N. Jacka, Denny Meyer, "Fusing Data Mining, Machine Learning and Traditional Statistics to Detect Biomarkers Associated with Depression", February 5, 2016, PLOS ONE, DOI:10.1371/journal.pone.0148195.
- [2] Stacie L. Nelson, "Major Depressive Disorder: Overview, Treatment and Recurrence Prevention", August 2012, unpublished.
- [3] Kurt Kroenke, Robert L. Spitzer, "The PHQ-9: A New Depression Diagnostic and Severity Measure", *Psychiatric Annals* 32:9, 2002.
- [4] <https://www.digitalocean.com/community/tutorials/an-introduction-to-machine-learning>.
- [5] Meenal J. Patel, Alexander Khalaf, Howard J. Aizenstein, "Studying depression using imaging and machine learning methods", *NeuroImage: Clinical* 10(2016), pp. 115-123.
- [6] Turker Tekin Erguzel, Cumhur Tas, Merve Cebi, "A wrapper based approach for feature selection and classification of major depressive disorder – bipolar disorders", *Computers in Biology and Medicine* 64(2015), pp.127-137.
- [7] Joao R. Sato, Jorge Moll, Sophie Green, John F.W. Deakin, Carlos E. Thomaz, Roland Zahn, "Machine learning algorithm accurately detects fMRI signature of vulnerability to major depression", *Psychiatry Research: Neuroimaging* 233 (2015), pp.289-291.
- [8] Roland Zahn, Karen E. Lythe, Jennifer A. Gethin, Sophie Green, John F. William Deakin, Allan H. Young, Jorge Moll, "The role of self blame and worthlessness in the psychopathology of major depressive disorder", *Journal of Affective Disorders* 186 (2015), pp. 337-341.
- [9] Ronald C. Kessler, Hanna M. van Loo, Klaas J. Wardenaar, Robert M. Bossarte, Lisa A. Brenner, Tianxi Cai, David Daniel Ebert, Irving Hwang, Junlong Li, Peter de Jonge, Andrew A. Nierenberg, Maria V. Petukhova, Anthony J. Rosellini, Nancy A. Sampson, Robert A. Schoevers, Marsha A. Wilcox, EdD, Alan M. Zaslavsky, "Testing a machine learning algorithm to predict the persistence and severity of major depressive disorder from baseline self reports", *Mol Psychiatry* 2016 October 21(10): 1366-1371.
- [10] <https://wwwn.cdc.gov/nchs/nhanes/search/datapage.aspx?Component=Questionnaire&CycleBeginYear=2015>
- [11] Mohamed Aly, "Survey on Multiclass Classification Methods", November 2005, https://pdfs.semanticscholar.org/a546/f2c88c588a2a46c054f67b39a3ebefdae694.pdf?_ga=2.146174396.963036132.1519347649-952649322.1519347649.
- [12] <https://www.cs.waikato.ac.nz/ml/weka/>
- [13] [https://en.wikipedia.org/wiki/Weka_\(machine_learning\)](https://en.wikipedia.org/wiki/Weka_(machine_learning))