Proceedings of the Second International Conference on Artificial Intelligence and Smart Energy (ICAIS-2022)
IEEE Xplore Part Number: CFP22OAB-ART; ISBN: 978-1-6654-0052-7

Prediction of Public Mental Health by using Machine Learning Algorithms

CH.M.H.Saibaba
Department of CSE
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Guntur
saibaba.ch77@gmail.com

Alekhya K V K
Department of CSE
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Guntur
alekhyakolluru27@gmail.com

Yeshwanth K
Department of CSE
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Guntur
yeshwanthyesh6026@gmail.com

Praveen Tumuluru
Department of CSE
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Guntur
praveenluru@gmail.com

2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS) | 978-1-6654-0052-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICAIS53314.2022.9742837

***Abstract*- Interpretation public emotional wellness concerns utilizing data science and observing arrangements dependent on the discoveries from the data science tasks can be complicated and requires progressed methods, contrasted with regular data investigation projects. Have an extensive task the executive's interaction to guarantee that undertaking partners are capable and have sufficient information to carry out the data science procedure. Thus, this paper describes another structure that psychological wellness experts be able to use to address difficulties they realize utilizing data science. Albeit countless exploration papers have been distributed on open emotional well-being, few have tended to the utilization of data science in open psychological wellness. As of late, data science has altered the manner in which we oversee, investigate and influence data in medical care industry. High prevalence of mental health and the need for effective mental fitness care, blended with current advances in AI, has led to a growth in explorations of ways the sphere of system getting to know Machine Learning can assist inside the detection, prognosis and treatment of mental health issues.**

***Index Terms*- Psychological Health, Public Health, Machine Learning, Visual Data Exploration, Data Science.**

I. INTRODUCTION

Mental illness denotes all diagnosable mental disorders which are characterized by irregularities in reasoning, sentiments, or practices.[1] Psychological maladjustment is extremely normal and causes generous social and financial weight around the world, yet no standardized biological diagnostic tests are available, and the analysis is as yet reliant upon clinical abilities and well-qualified assessment. Mental health is vital for overall well-being of human. Mental prosperity is significant for personal satisfaction and the capacity to adapt to information to day life. Mental illness can have a huge impact on and correlation with physical illness.[2] Along with medical conditions, data management factors are examined in the experimental segment to highlight the proposed work performance.[15]

II. LITERATURE SURVEY

D. V. Dimitrov, "Medical internet of things and big data in healthcare," Healthcare Informatics Research.

A range of technology can reduce standard prices for the revention or organization of continual infections. These consist of gadgets that regularly reveal health signs, devices that car-administer treatment options, or gadgets that song real time health facts whilst a patient self-administers a remedy.

Because they've accelerated get right of entry to high-pace Internet and smartphones, many sufferers have started out to use cellular programs (apps) to control diverse health desires.[3] These devices and cell apps are actually increasingly more used and included with telemedicine and telehealth via the clinical Internet of Things (mIoT). When many data sets are used, these tests enable a proper channel of Comparison.[13]

mIoT is a essential piece of the digital revolution of healthcare, as it permits new business prototypes to develop and allows changes in work methods, throughput upgrades, cost containment and greater purchaser practices.[4]

A new classification of "personalized preventative fitness coaches" (Digital Health Advisors) will develop. These employees will acquire the abilities and the potential to clarify and understand health and nicely-being information. They will help their customers keep away from continual and weight loss program-related contamination, enhance intellectual function, achieve stepped forward mental fitness achieve stepped forward life overall. As the global population a long time, such roles turns into an increasing number of essential.[5]

III. THEORETICAL ANALYSIS

A. K-Nearest Neighbor

K-Nearest Neighbor is one of the top Machine Learning calculations fundamentally dependent on Supervised Learning strategy. K-NN computation expects the comparability between the fresh out of the box new case/measurements and to be had cases and put the new case into the classification this is most similar as the to be had classes.[6]

K-NN calculation stores every one of the accessible information and characterizes another information point dependent on the comparability. This implies when new information shows up then it tends to be effectively grouped into a well suite class by utilizing K-NN calculation.

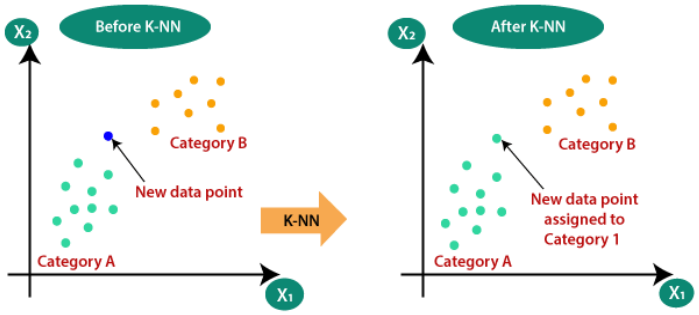

Fig. 1. Graphs represents before and after applying k-nn, new data point assigned to category 1ine in the 2nd graph

B. Random Forest

A random forest is an AI method that is utilized to take care of relapse and arrangement issues. It fabricates decision trees on various examples and takes their larger part vote in favor of arrangement and normal in the event of relapse. A random forest calculation comprises of numerous decision trees.[7]

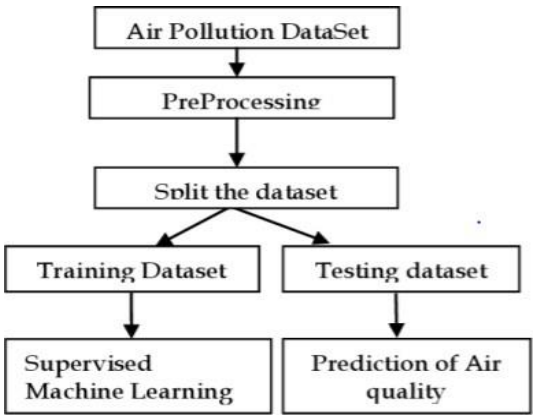

Fig 2. Flowchart representation of random forest with an example, like splitting, training and testing the dataset

C. Logistic Regression

Logistic regression utilizes a condition as the portrayal, particularly like linear regression. Input esteems (x) are joined straightly utilizing loads or coefficient esteems to foresee a output esteem (y).

Logistic regression is fundamentally a regulated grouping calculation. In a characterization issue, the objective variable (or result), y, can take just discrete qualities for a given arrangement of features (or data sources), X. [8]

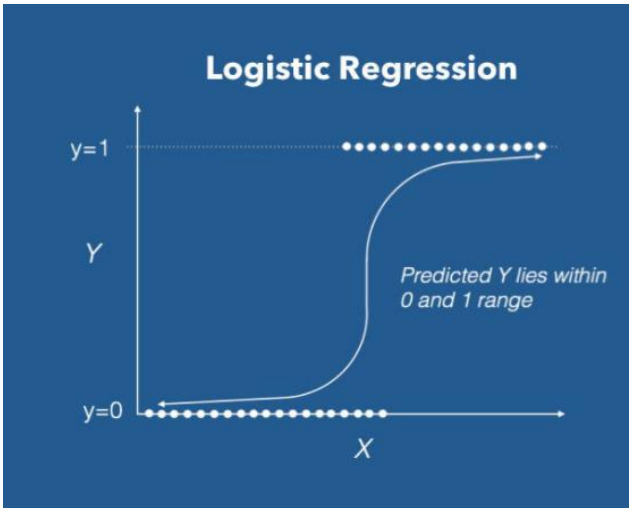

Fig. 3. Graphical representation of logistic regression, where the probability of y lies between 0 and 1

D. Decision Tree

Decision Trees are a kind of Supervised Machine Learning. The objective of calculation is to make a prototype that calculates the worth of an actual variable where interior nodes address the elements of a dataset, branches address the decision rules, and each leaf hub addresses the result.[9]

In a Decision tree, the tree can be clarified by two substances, in particular decision hubs and leaves. The leaves are the decisions or the ultimate results. Furthermore, the decision hubs are the place where the information is parted. The latest developments in the signal processing have lengthened the conventional methods of the wavelet transform to the domains that are with irregularities, that is the graph.[12]

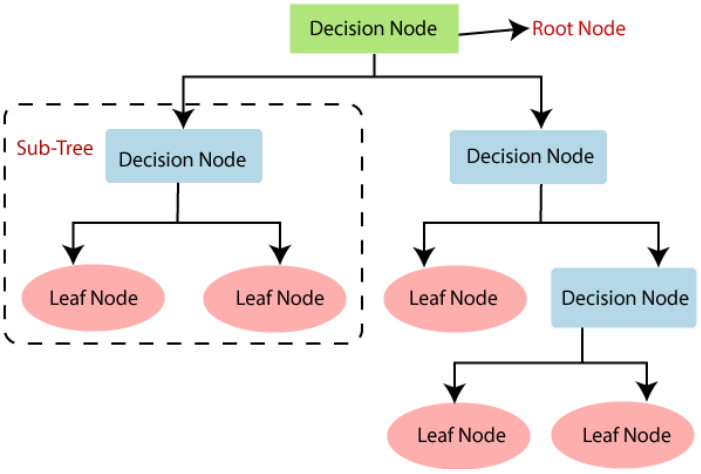

Fig. 4. Decision tree represents node with ultimate results.

IV. PROCESS OF EXECUTION

1. Library and data loading
2. Data cleaning
3. Encoding data

4. Covariance matrix. Variability comparison between categories of variables
5. Some charts to see data relationship
6. Scaling and fitting
7. Tuning
8. Evaluating models
    a) Logistic eegression
    b) Kneighbors classifier
    c) Decision tree classifier
    d) Random forests
    e) Bagging
    f) Boosting
    g) Stacking
9. Predicting with neural network
10. Success method plot
11. Creating predictions on test set
12. Submission
13. Conclusion

V. RESULTS AND DISCUSSION

Think of portrayed our methodology as well as examination, we now represent our discoveries. For expanded clearness, the uncovered associations will be outlined as Association rules. Our investigational conclusions recommend that, under 2 affiliation rules are normal in both female and male and shows high-level certainty among different standards. We detected the correlation between the states of user's stress and their social interaction behavior in social networks by utilizing real world social media data.[16]

As per the beneath affiliation rules, we can contend that places with low exorbitant drinking rate, significant degree of weight, high smoking rate and individuals have less proactive tasks has the most elevated continuous of mental pain. Likewise, these principles have shown high certainty, Support and Lift limits among females rather than guys.[10] We can use the selection of input weights and biases which is random in an improved version of ELM.[11] Based on the collected data, it includes relevant columns, such as volume, close, open, high costs, and market capitalization. "NaN" values are substituted with the meaning of a specific attribute. Next, all the datasets are merged into one, resulting in one overall dataset.[14]

```
logisticRegression()
########### Logistic Regression ###############
Accuracy: 0.7962962962962963
Null accuracy:
 0    191
1    187
Name: treatment, dtype: int64
Percentage of ones: 0.4947089947089947
Percentage of zeros: 0.5052910052910053
True: [0 0 0 0 0 0 0 0 1 1 0 1 1 0 1 1 0 1 0 0 0 1 1 0 0]
Pred: [1 0 0 0 1 1 0 1 0 1 0 1 1 0 1 1 1 1 0 0 0 0 1 0 0]
```

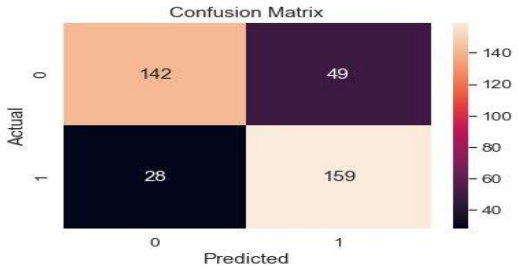

Fig. 5. Results for logistic regression algorithm with confusion matrix and accuracy

```
Knn()
Rand. Best Score:  0.8209714285714286
Rand. Best Params:  {'weights': 'uniform', 'n_neighbors': 2
[0.815, 0.818, 0.811, 0.811, 0.812, 0.811, 0.811, 0.821, 0.
0.823, 0.815]
########### KNeighborsClassifier ###############
Accuracy: 0.8042328042328042
Null accuracy:
 0    191
1    187
Name: treatment, dtype: int64
Percentage of ones: 0.4947089947089947
Percentage of zeros: 0.5052910052910053
True: [0 0 0 0 0 0 0 0 1 1 0 1 1 0 1 1 0 1 0 0 0 1 1 0 0]
Pred: [1 0 0 0 1 1 0 1 1 1 0 1 1 0 1 1 1 1 0 0 0 0 1 0 0]
```

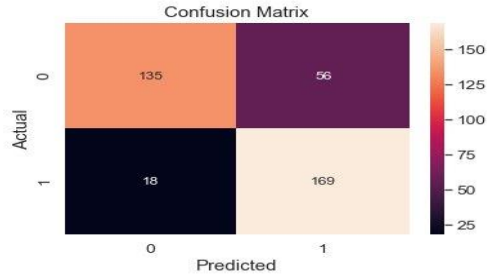

Fig. 6. Results for k-nearest neighbor with confusion matrix and accuracy

```
treeClassifier()
Rand. Best Score:  0.8305206349206349
Rand. Best Params:  {'criterion': 'entropy', 'max_depth': 3,
[0.828, 0.831, 0.81, 0.83, 0.831, 0.831, 0.831, 0.821, 0.819,
5, 0.83]
########### Tree classifier ###############
Accuracy: 0.8068783068783069
Null accuracy:
 0    191
1    187
Name: treatment, dtype: int64
Percentage of ones: 0.4947089947089947
Percentage of zeros: 0.5052910052910053
True: [0 0 0 0 0 0 0 0 1 1 0 1 1 0 1 1 0 1 0 0 0 1 1 0 0]
Pred: [1 0 0 0 1 1 0 1 1 1 0 1 1 0 1 1 1 1 0 0 0 0 1 0 0]
```

Fig. 7. Results for tree classifier with confusion matrix and accuracy

```
randomForest()
Rand. Best Score:  0.8321206349206349
Rand. Best Params:  {'criterion': 'gini', 'max_depth': 3,
[0.831, 0.831, 0.831, 0.833, 0.831, 0.831, 0.831, 0.831,
0.831, 0.833]
########### Random Forests ###############
Accuracy: 0.8121693121693122
Null accuracy:
 0    191
1    187
Name: treatment, dtype: int64
Percentage of ones: 0.4947089947089947
Percentage of zeros: 0.5052910052910053
True: [0 0 0 0 0 0 0 0 1 1 0 1 1 0 1 1 0 1 0 0 0 1 1 0 0]
Pred: [1 0 0 0 1 1 0 1 1 1 0 1 1 0 1 1 1 1 0 0 0 0 1 0 0]
```

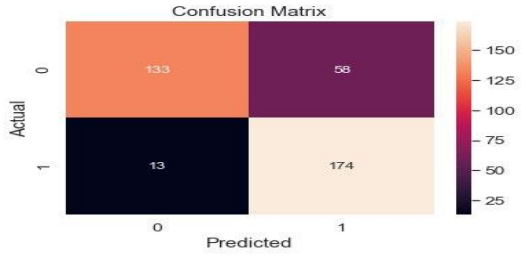

Fig. 8. Results for logistic regression algorithm with confusion matrix and accuracy

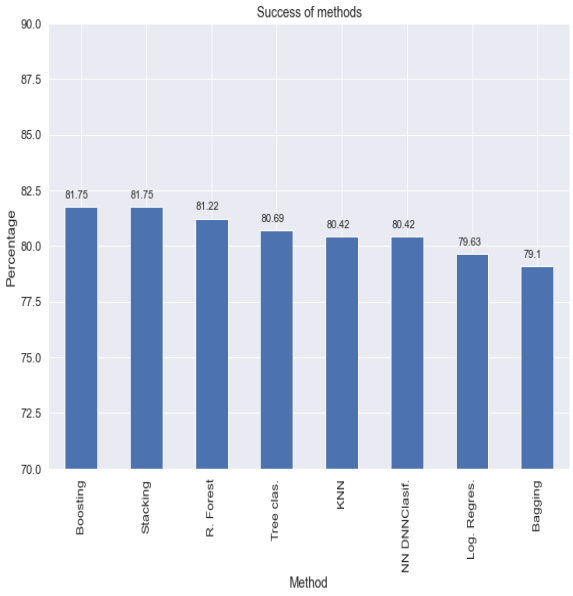

Fig. 9. Graph represents success rates for different methods

Index	Treatment				
5	1	161	1	302	1
494	0	811	1	790	0
52	0	1104	0	1253	1
984	0	45	1	443	0
186	0	1241	1	402	0
18	1	874	0	900	1
317	0	921	1	1186	1
511	1	1191	1	578	0
364	1	481	1	1043	1
571	1	308	0	1089	1
609	0	269	0	156	1
1147	1	731	0	184	1
922	1	1017	1	856	0
461	0	1193	0	14	0
740	1	875	1	142	0
955	1	1	1	709	1
814	1	796	1	1243	1
1160	1	1092	1	416	1
85	0	141	1	418	0
733	0	1231	0	725	0
1112	0	953	1	254	0
124	0	240	0	320	1
1040	1	775	1	299	0
492	0	55	1	979	1
1159	0	641	0	688	0
211	1	333	1	471	1
1020	0	723	0	420	1
892	0	1116	0	539	1
453	0	39	1	994	0
646	1	152	0	436	0

Fig. 10. Final output obtained is represented in a .csv file

VI. CONCLUSION

Ongoing investigations on mental health have shown a solid connection between the way of life factors and mental health. Additionally, late healthcare concentrated on discovering that way of life factors was emphatically connected with the results. The point of a review is to look at relationships in the middle of mental health and human lead factors as smoking, proactive tasks, drinking, and eating inclinations with occupants at the urban or state level utilizing accumulated information. More explicitly, the principal point of the review was to determine the connection between Frequent Mental Distress (FMD) and human being conduct designs among United States populace. This exploration needs utilized information from the most recent six years of AHR report; individual investigations have been done on female and male datasets, then, at that point, contrasted the conduct factors related and psychological health along with both female and male. An investigation discovered that places that get low extreme consumption rate high-level corpulence, and the high smoke rate has the most noteworthy successive of psychological trouble. Additionally, these guidelines have demonstrated great certainty limit as well as women rather than guys.

REFERENCES

[1] Adrian B. R. Shatte, Delyse M. Hutchinson, and Samantha J. Teague. 2019. Machine learning in mental health: A scoping review of methods and applications. Psychological Medicine 49, 9 (2019), 1426–1448.

[2] Stephen M. Schueller, Adrian Aguilera, and David C. Mohr. 2017. Ecological momentary interventions for depression and anxiety. Depression and Anxiety 34, 6 (2017), 540–545.

[3] SAMHS (Substance Abuse and Mental Health Services Administration). 2018. Key substance use and mental health indicators in the United States: Results from the 2017 National Survey on Drug Use and Health (HHS Publication No. SMA 18-5068, NSUDH Series H-53). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration, Rockville, MD.

[4] Guyon, A. Elisseeff, An introduction to variable and feature selection,J. Mach. Learn. Res. 3 (2003) 1157–1182 .

[5] S. Monteith, T. Glenn, J. Geddes, and M. Bauer, “Big data are coming to psychiatry: a general introduction,” International Journal of Bipolar Disorders.

[6] P. Kellmeyer, “Big brain data: on the responsible use of brain data from clinical and consumer-directed neurotechnological devices,” Neuroethics.

[7] Praveen Tumuluru, Burra Lakshmi Ramani, CH.M.H.Saibaba, B.Venkateswarlu, N.Ravindar, “OpenCV Algorithms for facial recognition”, International Journal of Innovative Technology and Exploring Engineering, Vol.8, Issue.8, pp.927-933, June 2019.

[8] Viranjitha Lakshmi Maturi, Nagarjuna Reddy Boya, Jaiakanth Polisetti, Sripujitha Adavi, CH.M.H.Saibaba, “Twitter Sentimental Analysis using Machine Learning Techniques”, International Journal of Innovative Technology and Exploring Engineering, Vol.8, Issue.6, pp.1592-1594, April 2019.

[9] K.Prasuna, K.V.S.N.Rama Rao, CH.M.H.Saibaba, “Application of Machine Learning Techniques in Predicting Breast Cancer – A Survey”, International Journal of Innovative Technology and Exploring Engineering, Vol.8, Issue.8, pp.826-832, June 2019.

[10] Burra Lakshmi Ramani, Padmaja Poosapati, Praveen Tumuluru, CH.M.H.Saibaba, Mothukuri Radha, K.Prasuna, “Deep Learning and Fuzzy Rule-Based Hybrid Fusion Model for Data Classification”, International Journal of Recent Technology and Engineering, Vol.8, Issue.2,pp3205-3213, July 2019.

[11] Mugunthan, S. R., and T. Vijayakumar. "Design of Improved Version of Sigmoidal Function with Biases for Classification Task in ELM Domain." Journal of Soft Computing Paradigm (JSCP) 3, no. 02 (2021): 70-82.

[12] Manoharan, Samuel. "Study on Hermitian graph wavelets in feature detection." Journal of Soft Computing Paradigm (JSCP) 1, no. 01 (2019): 24-32

[13] Pandian, A. Pasumpon. "Performance Evaluation and Comparison using Deep Learning Techniques in Sentiment Analysis." Journal of Soft Computing Paradigm (JSCP) 3, no. 02 (2021): 123-134.

[14] Andi, Hari Krishnan. "An Accurate Bitcoin Price Prediction using logistic regression with LSTM Machine Learning model." Journal of Soft Computing Paradigm 3, no. 3 (2021): 205-217.

[15] Balasubramaniam, Vivekanadam. "IoT based Biotelemetry for Smart Health Care Monitoring System." Journal of Information Technology and Digital World 2, no. 3 (2020): 183-190.

[16] Sharma, Shraddha, Ila Sharma, and A. K. Sharma. "Automated System for Detecting Mental Stress of Users in Social Networks Using Data Mining Techniques." In International conference on Computer Networks, Big data and IoT, pp. 769-777. Springer, Cham, 2019.