Journal of Information Systems Engineering and Management
2025, 10(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Predictive Analytics in Mental Health: A Machine Learning Approach to Assessing Depression Severity

Dr. Satya Kiranmai Tadepalli[1], Dr. Shobarani Salvadi[2], Madhuri Timmapuram[3], Srilakshmi Anireddy[4]

[1,2]Department of AI&DS, [3,4]Department of IT

[1,2,3,4] Chaitanya Bharathi Institute of Technology(A), Hyderabad, India

[1,2,3,4] Email – tadepallikiranmai84@gmail.com

ARTICLE INFO

Received: 02 Oct 2024

Revised: 30 Nov 2024

Accepted: 16 Dec 2024

ABSTRACT

Contemporary healthcare reforms are being significantly shaped by the ongoing advancements in technology. One area where these advancements are proving crucial is in the understanding and treatment of depression, which is increasingly emerging as a substantial public health concern. To address this issue, there is a growing interest in leveraging novel research methods and therapeutic approaches to identify the contributing factors to depression. This study adopts an innovative approach, utilizing machine learning techniques to carry out an exhaustive examination of diverse data sources. The primary aim is to gain a profound comprehension of the complex interplay between various facets of individuals' quality of life and the presence of depression. To undertake this investigation, the researchers have harnessed the National Health and Nutrition Examination Survey (NHANES) Dataset provided by the Centers for Disease Control and Prevention, a rich source of valuable health-related information.

In this study, the focus is on exploring the behavioral and social dimensions of numerous subjects and their intricate connections to depression. To achieve this, a diverse array of machine learning classifiers has been deployed, including the Decision Tree Classifier, Random Forest Classifier, Gaussian Naive Bayes, KNN Classifier, Logistic Regression, Support Vector Machine Classifier, and Multilayer Perceptrons. By applying these classifiers, the researchers aim to assess their performance across various metrics, providing valuable insights into which models are best suited to discern the connection between depression and aspects of quality of life.

Keywords: Depression, Mental health, Machine Learning, Artificial Neural Network.

1. INTRODUCTION

"Depression" is one of the most hazardous and complex psychological diseases in the world and is thought to be the primary cause of the sickness burden of all disorders. As a result, many researchers in the fields of science and medicine have focused on the study of depression. By 2030, depression would be among the top illnesses and causes of mortality, according to WHO. Depression's effects last and continue to have a detrimental impact on a person's performance and quality of life (QoL) even after they get effective therapy. The term "quality of life" refers to several dimensions of an individual's existence, such as their mental, emotional, and physical health. Numerous academics and medical professionals are researching these characteristics because they provide insight into people's perspectives on life. Like this, it centres on two aspects: (1) being knowledgeable about the relationships between food consumption, depression, and patient quality of life; and (2) finding out how to improve psychosocial results and patient quality of life. Some people keep their problems with money, relationships, and sexuality private until someone asks them about them. Therefore, the evaluation of QoL traits reveals these people's underlying issues more accurately and results in greater care. Over the last 30 years, a great deal of research has been done in several medical specialties to investigate quality of life. Especially in the fields of psychology and psychiatry, it is important to understand how QoL relates to mental health issues. The association involving depression and QoL characteristics may be found using machine learning's larger search capability for QoL components.

Healthcare remains among the most significant global concerns, regardless of the state of the nation or its level of development. To improve quality of life, intelligent, efficient, and secure medical systems are being developed as a global priority. Researchers from a variety of fields have been drawn to the fields of neuroscience and psychological science by the initial investigations of human behaviour. This also holds true for the quickly expanding domains of machine learning and computer science research. For medical professionals and institutions, determining whether an individual has psychological issues is a recurring challenge, particularly with the patients who are younger. Recent research has demonstrated that machine learning and deep learning are capable of recognising mental health issues in individuals and comprehending how such issues affect lifestyle choices. The most significant alteration in human growth that can be observed everywhere is a shift in mental health. Because of this, it is believed that the two primary illnesses associated with ageing are mental health conditions including depression and anxiety. Both negatively impact the quality of life (QoL) of patients and impede their capacity for decision-making. This causes a great deal of misery, which frequently culminates in self-harm or suicide.

The application of AI technology in recent years and machine learning has proven to be a great tool for understanding and analyzing human behavioral patterns and decision-making capabilities. Machine learning is one of AI's most active subfields. It is a strategy that learns from patterns and situations and suggests the most probable decision/solution. Its learning method is used in a number of intelligent operations, including self-driving automobiles and voice recognition software, and it also makes suggestions (based on search history, Google proposes what users wish to look for). Additionally, machine learning approaches are improving many other fields, such as multi-omics illness discovery, human activity detection, and the prediction of suicidal behaviour. Current examples have demonstrated how machine learning is assisting in the creation of algorithms that can compete with human doctors in terms of performance. In the healthcare sector, machine learning is doing a remarkable job of identifying patterns in data that help medical professionals and organisations identify a number of serious diseases.

We used machine learning techniques in this investigation to analyze and examine heterogeneous data to further comprehend the relationship between aspects of quality of life (QoL) along with depression. Providing an overview of the current state of understanding, we highlight key concepts, theories, methodologies, and findings.

2. BACKGROUND STUDY

[1] Researchers used machine learning to study the link between depression and quality of life, analyzing NHANES data. They introduced the PPMCSVM model, achieving high precision, recall, and accuracy in their findings. However, SVM predictions can vary with sensitive data.

[2] the application of persuasive technologies (PT) in the mental health sector to advance equitable utilisation of mental health services. [7] Given the close link between depressed symptoms and quality of life. [8] implied the potential to support choices related to the identification, prognosis, and therapy of patients who experience mental illness utilizing

clinical and biological data. The usage of ML techniques and other techniques like data tweaking is discussed using SVM. [12] showed the correlation of depressed symptoms acted as a

mediating factor in the relationship between smoking and HRQoL. [14] using the Beck Depression Inventory, the International Physical

Activity Survey, and the Leary Interpersonal Orientation Paper Test, respectively. ANOVA,

Pearson's correlation analysis, reliability analysis, and The data and the relationships between the quantity of physical physical activity and depressive symptoms and social contacts were all examined using descriptive statistics. [15] using the social network data obtained a deep integrating support vector machine (DISVM) technique, which enables depression to be recognized. The proposed obtains 86% of precision on test set when compared to other algorithms RBF-NN, SVM, KNN, with precisions 82%, 80%, 79% respectively. [20] Using information gathered from Reddit on models like SVM with co-training, NB, and RF, the utilisation of social media platforms and big data analysis for detecting patients with mental diseases produced an F-measure of 0.84, 0.83, and 0.83. [21Covariance analysis in many variables The viability and effectiveness of using Tess, a unified psychological AI, to reduce college students' self-reported feelings of hopelessness and anxiety were assessed using a MANOVA on two groups. [24] The study employed logistic regression analysis to investigate the potential associations between depression and reported body mass index (BMI) and body picture, including self-expressed BMI, BMI difference, perceived mass, and planned weight. [25] The following classification techniques are used to a dataset: K-nearest neighbours, decision trees, support vector machines, logistic regression, and naïve bayes. Furthermore, models were constructed via the ensemble bagging technique and the random forest tree ensemble approach. [28] show that Support Vector Machine

(SVM) achieves a high F1-score of 89.7%, outperforming both lexical classifiers as well as other machine learning (ML) classifiers.

3. DATASET

The National Health and Nutrition Examination Survey (NHANES) dataset was utilized in this study. This is a research project created to assess the nutritional and overall health of elders and kids in the United States. This survey stands out since it incorporates both interviews and practical exams. The main initiative of the National Centre for Health Statistics (NCHS) is NHANES. The task of gathering crucial and national health statistics falls to NCHS, a branch of the Centres for Disease Control and Prevention (CDC).

The NHANES interview includes questions on socioeconomic status, diet, health, and demography. As part of the examination procedure, highly qualified medical specialists conduct physiological, dental, and medical tests in the lab.

Demographic 10175 records 47 columns, questionnaire 10175 records 953 columns, diet 9813 records 168 columns, medication 20187 records 13 columns, laboratory 9813 records 424 columns, examination 9813 records 224 columns.

4. PROPOSED WORK

4.1 SYSTEM ARCHITECTURE

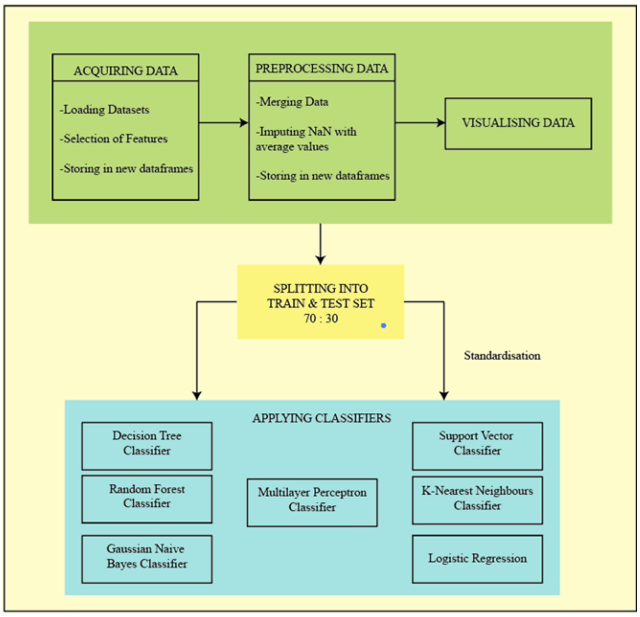

Fig. 1 System Architecture

In the architecture represented above, the dataset is acquired and utilized by loading it.

Carefully, some of features are selected and these features are stored into a new separate dataframes. After performing certain preprocessing steps, these dataframes are merged into a single dataframe. After merging, it is observed that there are some NaN values which needed to be handled and these are imputed by the average values obtained for the variables. The target variable Depression is handled with a unique value and is converted into categorical type for classification. The modified data is then stored into a new dataframe. This data is then visualized for insights.

The obtained data is segregated into training and testing data in 70 : 30 and applied on classifiers.

Next, the same procedure is followed but after splitting of data for training and testing standardization is performed upon the data. This standardized data is applied on the classifiers.

The results were obtained individually.

4.2 WORKFLOW

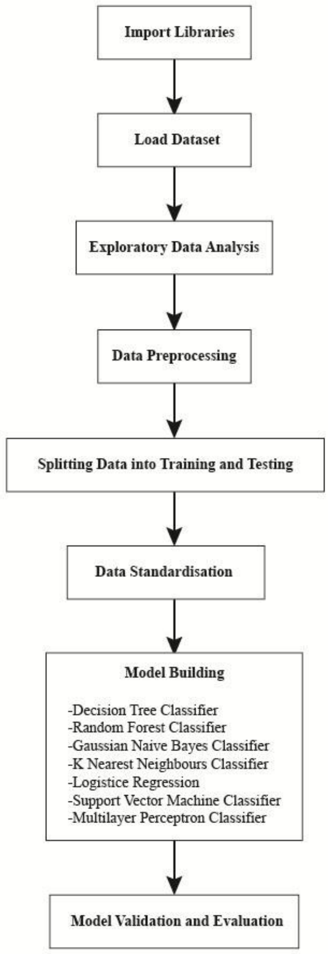

Fig. 2 Workflow

4.2.1 IMPORTING LIBRARIES

The first step is importing the necessary libraries to carry out the work. Libraries provide the comfort of performing tasks in an easy manner.

4.2.2 LOADING THE DATA

The National Health and Nutrition Examination Survey (NHANES) is a comprehensive programme designed to evaluate the health and nutritional status of adults and children in the United States. The dataset used in this investigation is derived from this programme.

Numerous questions on demographics, socioeconomic status, eating habits, and general health are included in the NHANES interviews. Additionally, NHANES involves rigorous medical, dental, and physiological assessments carried out by highly trained medical professionals. The dataset in question is accessible online and is provided in CSV format.

Information on a person's physical fitness, levels of activity, smoking and alcohol use, housing and work status, financial information, general health issues, sexual activity, drug use, and food intake are just a few of the various aspects that are included in the dataset. These comprehensive data sources form the foundation for the proposed research work, offering a rich repository for exploring the multifaceted relationships between various factors and depression.

4.2.3 EXPLORATORY DATA ANALYSIS

The loaded data is explored and analyzed to gain an understanding about the data. It is an important step to understand the data, recognize the quality issues embedded in the data, and get an insight into how to carry out the preprocessing steps of the data.

The datasets used are demographic, questionnaire, and diet. The demographics dataset as whole consists of 10175 records with 47 columns, The questionnaire dataset consists of 10175 records with 953 columns. The demographics dataset consists of 9813 records with 168 columns. We selected two features from demographic and diet datasets, and 12 from the questionnaire dataset.

4.2.4 DATA PREPROCESSING

This is a significantly important phase that necessitates meticulous execution. It involves a series of steps, starting with Data Preparation, where data is collected and organized to ensure easy access for analysis. The subsequent phase is Data Integration, where disparate datasets are consolidated into a unified dataframe. Following this is Data Cleaning, a pivotal step in which data validity is scrutinized, and the overall data quality is assessed. This process includes the identification and handling of any extraneous noise present in the dataset, which can commonly manifest as missing values or outliers.

Notably, the acronym "NaN" signifies "not a number," and it arises when certain input parameters in floating-point operations yield an undefined result. Within this research, the management of NaN values involves the imputation of either the average value of a specific feature or the introduction of a unique value that does not reoccur in the dataframe. Additionally, data is visually represented in various forms, such as heatmaps and histograms, aiding in the comprehensive understanding of the dataset and its characteristics.

Fig. 3 Heat Map

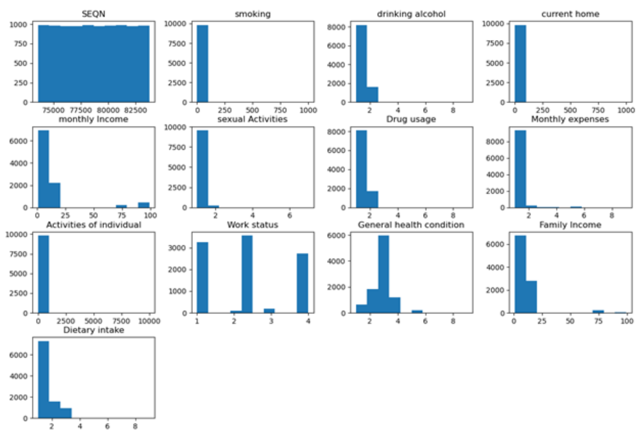

Fig. 4 Histogram

4.2.5 SPLITTING THE DATA INTO TRAINING AND TESTING SET

Two subsets of the pre-processed data are separated out: a training set that makes up 70% of the data and a testing set that makes up 30%.

4.2.6 DATA STANDARDIZATION

It is important for data to be in a standard form. The features are standardized in a machine-learning pipeline. The Standard Scaler helps create a standardized distribution with a mean of zero and a standard deviation of one (unit variance). To standardize the data points, the following mathematical formula can be used,

$$z = (x - u) / s$$

where,

z = standardized data points

u = Mean value

s = Standard Deviation

4.2.7 MODEL BUILDING

Many machine learning approaches exist, and some of these are used in this study: Gaussian Naïve Bayes Classifier, Random Forest Classifier, Knn Classifier, Decision Tree Classifier, Logistic Regression, and Support Vector Classifier., And A Customized Multilayer Perceptron Classifier are used.

The classifiers are implemented with and without standardization of the data and these observation were discussed later on. A multilayer perceptron classifier with customized layers, optimization, learning rate and other parameters is adopted.

4.1.8 MODEL VALIDATION AND EVALUATION

At last, we validate the models and evaluate them by getting the predictions made by them and use different metrics to get the reports. Metrics used are accuracy, precision, recall, F1- score and support.

Precision: the proportion of correct predictions that were positive among all positive predictions produced by the model. To calculate it, divide the total number of true positives by the sum of all true positives and false positives.

$$Precision = \frac{TP}{TP + FP}$$

Recall: the proportion of all genuine positive data points to accurate positive forecasts. It is calculated as the ratio of genuine positives to the sum of true positives + false negatives.

$$Recall = \frac{TP}{TP + FN}$$

Accuracy: The ratio of the model's correct predictions to all other forecasts. The number of accurate forecasts divided by the total number for predictions is used to compute it.

$$Acuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

F1 score: The F1 score is the harmonic mean of recall and accuracy.

$$F1Score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} = \frac{2 * (Precision * Recall)}{(Precision + Recall)}$$

5. RESULTS

In this research, we have employed several machine learning methodologies and customised an artificial neural network in order to identify the factors that contribute to depression and their correlation with life quality. The National Health and Nutrition Examination Survey (NHANES) dataset was utilised in this study. Neural networks and machine learning techniques were used in this experimental investigation. Techniques like Decision Tree, Random Forest, Gaussian Naive Bayes, KNN, Logistic Regression, Support Vector Machine, and Multilayer

Perceptron classifiers are used in this experimental investigation and will be compared with existing research work on different models.

The accuracies obtained for Decision Tree Classifier is 83.00%, Random Forest Classifier is 88.00%, Naive Bayes is 89.13%, KNN Classifier is 88.80%, Support Vector Machine Classifier is 91.16%.

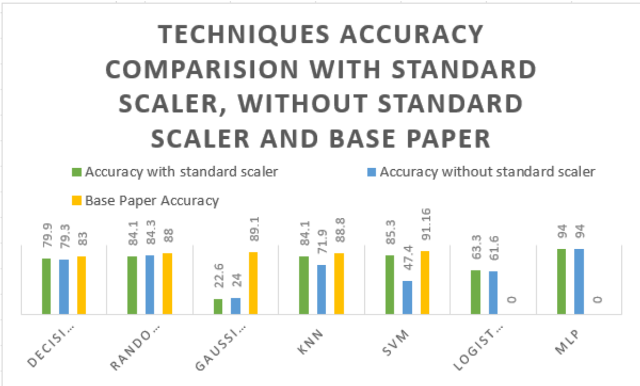

Fig. 5 Comparison of the accuracies of the models with existing work, proposed work with and without standardization of data.

The results obtained without standardisation of data for Decision Tree Classifier is 79.3%, Random Forest Classifier is 84.3%, Gaussian Naive Bayes is 24%, KNN Classifier is 71.9%, Logistic Regression is 61.6%, Support Vector Machine Classifier is 47.4%.

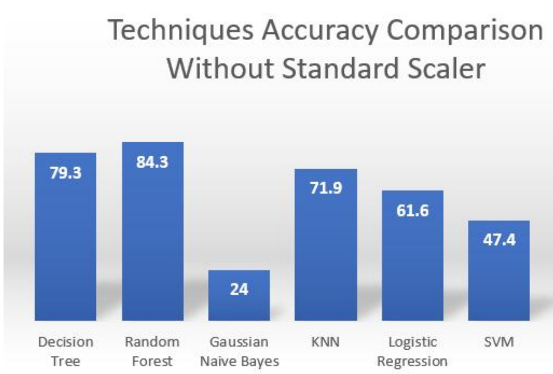

Fig. 6 Accuracies of the models without standardization of data.

The results obtained with standardisation of data for Decision Tree Classifier is 79.9%, Random Forest Classifier is 84.1%, Gaussian Naive Bayes is 22.6%, KNN Classifier is 84.1%, Logistic Regression is 63.3%, Support Vector Machine Classifier is 85.3%, Multilayer perceptron classifiers is 94%.

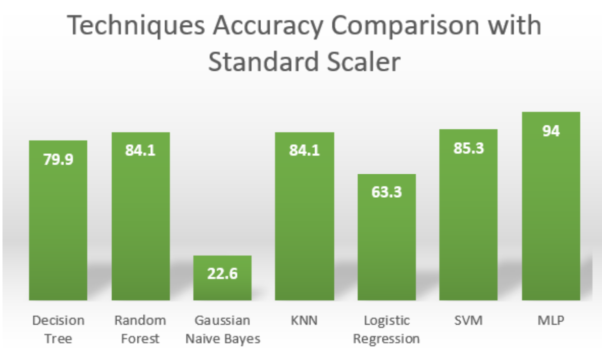

Fig. 7 Accuracies of the models with standardization of data.

We have observed that the classifiers show variation in results and the adopted MLP seems to perform better than the techniques applied.

6. CONCLUSION AND FUTURE SCOPE

In the course of our research, we have made a notable observation regarding the application of the Multilayer Perceptron (MLP) classifier. This implementation has yielded substantial enhancements in the classification process, resulting in an impressive accuracy score of 94%. Our findings have unveiled a significant improvement in comparison to all the machine learning classifiers previously employed in similar studies. Remarkably, our MLP classifier has outperformed the highest-scoring classifier in those prior works, which happened to be the Support Vector Machine (SVM), by a notable margin of 2.86%.

Furthermore, we have identified an additional factor contributing to the enhanced performance of our classifiers. The incorporation of the Standard Scaler into our methodology has proven to be pivotal in achieving more accurate and reliable results. This preprocessing technique has played a crucial role in enhancing the classifiers' ability to handle the data effectively, ultimately contributing to the robustness of our analysis. These combined advancements signify a substantial leap forward in the field of depression classification research, offering promising prospects for more accurate and informed insights into this intricate and critical health issue.

The aforementioned research's potential for future growth is encouraging, since it may result in a number of significant developments and applications for the study and treatment of depression. The research highlights the importance of considering various dimensions of an individual's quality of life in relation to depression. This insight can be used to develop personalized treatment plans that take into account the specific factors contributing to a person's depression, thereby improving treatment outcomes. With more accurate classification models, future research can focus on identifying early warning signs of depression, allowing for timely intervention and prevention strategies. This could have a significant impact on reducing the overall burden of depression in communities. As technology continues to advance, integrating the research findings with telehealth solutions and wearable devices can provide real-time monitoring and support for individuals with depression. Machine learning models can be used to analyze data from wearable devices and offer timely interventions or suggestions. The study's insights into the behavioral and social factors associated with depression can inform public health policies aimed at improving mental health in communities. This can lead to targeted programs and initiatives that address these factors proactively.

REFERENCES

[1] M. Habib, Z. Wang, S. Qiu, H. Zhao, and A. S. Murthy, “Machine Learning Based Healthcare System for Investigating the Association Between Depression and Quality of Life”, IEEE J. Biomed. Health Informat., vol. 26, no. 5, pp. 2168-2194, May 2022, doi: 10.1109/JBHI.2022.3140433.

[2] T. Kolenik and M. Gams, “Persuasive technology for mental health: One step closer to (Mental health care) equality?,” IEEE Technol. Soc. Mag., vol. 40, no. 1, pp. 80–86, Mar. 2021, doi: 10.1109/MTS.2021.3056288.t

[3] S. Chen, Y. Conwell, C. Cerulli, J. Xue, and H. F. K. Chiu, “Primary care physicians’ perceived barriers on the management of depression in China primary care settings,” Asian J. Psychiatry, vol. 36, pp. 54–59, 2018, doi: 10.1016/j.ajp.2018.06.019.

[4] H. Dibeklioglu, Z. Hammal, and J. F. Cohn, “Dynamic multimodal measurement of depression severity using deep autoencoding,” IEEE J. Biomed. Health Informat., vol. 22, no. 2, pp. 525–536, Mar. 2018, doi: 10.1109/JBHI.2017.2676878.

[5] J. L. Z. Montenegro, C. A. da Costa, and R. da Rosa Righi, “Survey of conversational agents in health,” Expert Syst. Appl., vol. 129, pp. 56–67, Sep. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417419302283t

[6] S. C. Daniel, A. Azuero, O. M. Gutierrez, and K. Heaton, “Examining the relationship between nutrition, quality of life, and depression in hemodialysis patients,” Qual. Life Res., vol. 30, no. 3, pp. 759–768, 2021, doi: 10.1007/s11136-020-02684-2.

[7] C. González-Blanch, F. Hernández-de-Hita, R. Muñoz-Navarro, P. Ruíz Rodríguez, L. A. Medrano, and A. Cano-Vindel, “The association between different domains of quality of life and symptoms in primary care patients with emotional disorders,” Sci. Rep., vol. 8, no. 1, 2018, Art. no. 11180, doi: 10.1038/s41598-018-28995-6.

[8] D. B. Dwyer, P. Falkai, and N. Koutsouleris, “Machine learning approaches for clinical psychology and psychiatry,” Annu. Rev. Clin. Psychol., vol. 14, no. 1, pp. 91–118, 2018, doi: 10.1146/annurev–clinpsy-032816-045037.

[9] S. Ledesma, M. A. Ibarra-Manzano, E. Cabal-Yepez, D. L. Almanza Ojeda, and J. G. Avina-Cervantes, “Analysis of data sets with learning conflicts for machine learning,” IEEE Access, vol. 6, no. 1, pp. 45062–45070, 2018, doi: 10.1109/ACCESS.2018.2865135.

[10] Y. Dong, E. C. Dragut, and W. Meng, “Normalization of duplicate records from multiple sources,” IEEE Trans. Knowl. Data Eng., vol. 31, no. 4, pp. 769–782, Apr. 2019, doi: 10.1109/TKDE.2018.2844176.

[11] S. Gu, Y. Cai, J. Shan, and C. Hou, “Active learning with error correcting output codes,” Neurocomputing, vol. 364, pp. 182–191, 2019, doi: 10.1016/j.neucom.2019.06.064.

[12] M. Milic et al., “Tobacco smoking and health-related quality of life among university students: Mediating effect of depression,” PLoS One, vol. 15, no. 1, pp. 1–18, 2020, doi: 10.1371/journal.pone.0227042.

[13] D. Riemann, L. B. Krone, K. Wulff, and C. Nissen, “Sleep, insomnia, and depression,” Neuropsychopharmacology, vol. 45, no. 1, pp. 74–89, 2020, doi: 10.1038/s41386-019-0411-y.

[14] Chang-Hwan Kim, Young-Eun Song, Yoo-Jeong Jeon "The Effect of College Students’ Physical Activity Level on Depression and Personal Relationships", Healthcare (Basel). 2021 May; 9(5): 526. Published online 2021 Apr 29. doi: 10.3390/healthcare9050526

[15] Yan Ding et al., "A Depression Recognition Method for College Students Using Deep Integrated Support Vector Algorithm", IEEE Access, vol. 8, pp. 75616 - 75629, 2018, doi: 10.1109/ACCESS.2020.2987523.

[16] Marcel Trotzek et al., "Utilizing Neural Networks and Linguistic Metadata for Early Detection of Depression Indications in Text Sequences", IEEE Transactions on Knowledge and Data Engineering, Vol. 32, No. 3, 2020, pp. 588 - 601, doi: 10.1109/TKDE.2018.2885515.

[17] Huifeng Zhang et al., "Divergent and Convergent Imaging Markers Between Bipolar and Unipolar Depression Based on Machine Learning", IEEE Journal of Biomedical and Health Informatics, Vol. 26, no. 8, 2022, pp. 4100 - 4110, doi: 10.1109/JBHI.2022.3166826.

[18] Chaoqun Yue et al., "Fusing Location Data for Depression Prediction", IEEE Transactions on Big Data, 2021, pp. 355 - 370, doi: 10.1109/TBDATA.2018.2872569.

[19] Yaakov Hacohen-Kerner et al., "Detection of Anorexic Girls-In Blog Posts Written in Hebrew Using a Combined Heuristic AI and NLP Method", IEEE Access, Vol. 10, pp. 34800 - 34814, 2022 , doi: 10.1109/ACCESS.2022.3162685.

[20] Subhan Tariq et al., "A Novel Co-Training-Based Approach for the Classification of Mental Illnesses Using Social Media Posts", IEEE Access, 2019, Vol. 7, pp. 166165 - 166172, doi: 10.1109/ACCESS.2019.2953087.

[21] R. Fulmer et al., “Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: Randomized controlled trial,” JMIR Mental Health, vol. 5, no. 4, p. e64, Dec. 2018. [Online]. Available: http://mental.jmir.org/2018/4/e64/

[22] J. Jetter, “The good, the bad, and the aesthetically challenged: The good, the bad, and the aesthetically challenged [Opinion],” IEEE Technol. Soc. Mag., vol. 38, no. 4, pp. 27–31, Dec. 2019.

[23] B. Cliffe et al., “Clinicians’ use of and attitudes towards technology to provide and support interventions in child and adolescent mental health services,” Child Adolescent Mental Health, vol. 25, no. 2, pp. 95–101, May 2020. [Online]. Available: https://acamh.onlinelibrary.wiley.com/doi/abs/10.1111/camh.12362

[24] Tobia Darimont et al., “Body weight and self-perception are associated with depression: Results from the National Health and Nutrition Examination Survey (NHANES) 2005-2016”, J Affect Disord., vol. 274, pp. 929-934, 2020 Sep, doi: 10.1016/j.jad.2020.05.130.

[25] M. Srividya, S. Mohanavalli & N. Bhalaji , “Behavioral Modeling for Mental Health using Machine Learning Algorithms”, April 2018, Journal of Medical Systems 42(5):88, doi:10.1007/s10916-018-0934-5

[26] ML Tlachac, Elke Rundensteiner, “Screening For Depression With Retrospectively Harvested Private Versus Public Text”, IEEE Journal of Biomedical and Health Informatics, Volume: 24, Issue: 11, pp. 3326 – 3332, November 2020, doi: 10.1109/JBHI.2020.2983035.

[27] Ellen W. McGinnis et al., “Giving Voice to Vulnerable Children: Machine Learning Analysis of Speech Detects Anxiety and Depression in Early Childhood”, IEEE Journal of Biomedical and Health Informatics, Volume: 23, Issue: 6, November 2019, pp. 2294 – 2301, doi: 10.1109/JBHI.2019.2913590.

[28] Adenrele Oduntan, “I Let Depression and Anxiety Drown Me …”: Identifying Factors Associated With Resilience Based on Journaling Using Machine Learning and Thematic Analysis”, Vol. 26, No. 7, pp. 2168-2194, July 2022, doi: 10.1109/JBHI.2022.3149862.

[29] I. C. Passos, P. Ballester, J. V. Pinto, B. Mwangi, and F. Kapczinski, “Big data and machine learning meet the health sciences,” in Personalized Psychiatry, vol. 81, Cham, Switzerland: Springer, 2019, pp. 1–13.

[30] A. B. R. Shatte, D. M. Hutchinson, and S. J. Teague, "Machine learning in mental health: A scoping review of methods and applications," Psychol. Med., vol. 49, no. 9, pp. 1426–1448, 2019, doi: 10.1017/S0033291719000151.
[31] Centers for Disease Control and Prevention, "National health and nutrition examination survey: Analytic guidelines, 2011–2016," pp. 1–40, 2018, [Online]. Available: https://wwwn.cdc.gov/nchs/nhanes/analyticguidelines.aspx
[32] J. Pitt, "From trust and loyalty to lock-in and digital dependence [Editorial]," IEEE Technol. Soc. Mag., vol. 39, no. 1, pp. 5–8, Mar. 2020.
[33] "Depressive disorder (depression)", March 2023, https://www.who.int/news-room/fact-sheets/detail/depression [online]
[34] "Mental health", https://www.who.int/health-topics/mental-health#tab=tab_1, [online]
[35] Machine Learning, Tom Mitchell, McGraw Hill, 1997.