Journal of Intelligent Decision Support System (IDSS) 7 (3) (2024), pp. 210-217

Published by: Institute of Computer Science (IOCS)

Journal of Intelligent Decision Support System (IDSS)

Journal homepage: www.idss.iocspublisher.org

Mental disorder classification with exploratory data analysis (EDA)

Juanto Simangunsong[1], Mutiara S Simanjuntak[2], Nurmala Dewi Simanjuntak[3]

[1,2] *Universal Computer and Informatics Management Academy*
[3] *HKBP Nomensen University*

Article Info

Article history:

Received Jun 27, 2024
Revised Jul 05, 2024
Accepted Jul 15, 2024

Keywords:

Classification;
Diagnose;
Exploratory Data Analysis;
Mental Disorder.

ABSTRACT

Classification of mental disorders is the process of grouping mental disorders into categories based on their symptoms, causes and consequences. EDA is a data analysis strategy that emphasizes open-mindedness, creativity and diverse perspectives. EDA aims to explore as much data as possible, without imposing previous assumptions or models, until a coherent, coherent story emerges. EDA can help generate new hypotheses, identify patterns and outliers, and uncover underlying structures and relationships in data. This paper shows how EDA can be used to analyze and understand mental disorders data from a variety of sources and perspectives. We used EDA methods to explore the characteristics, prevalence, and distribution of mental disorders, as well as the relationships and interactions between mental disorders and other variables. We also compared EDA results with mental disorder classification systems such as the Diagnostic and Statistical Manual of Mental Disorders (DSM). We show that EDA can provide a more comprehensive and nuanced understanding of mental disorder data, as well as highlight the challenges and limitations of mental disorder classification. We hope this paper will illustrate the potential and benefits of EDA for mental disorders research and practice.

Corresponding Author:

Juanto Simangunsong,
Universal Computer and Informatics Management Academy ,
D3 Informatics Engineering Department
Medan – North Sumatera – Indonesia – 20132
Email : juantosimangunsong1980@gmail.com

Introduction

Many mental disorders exhibit overlapping symptoms, making it challenging to differentiate between them, which can lead to misdiagnosis or over-diagnosis. Unlike many physical illnesses, mental disorders often lack clear biological markers. Diagnoses are mainly based on observed behaviors and self-reported symptoms, which can be subjective. The criteria and definitions in diagnostic manuals are often influenced by cultural norms, which may not be applicable or accurate for individuals from diverse cultural backgrounds. Labeling individuals with specific disorders can result in stigmatization and discrimination, impacting their social and professional lives. There is ongoing debate about whether mental disorders should be classified categorically (as distinct disorders) or dimensionally (as a spectrum of symptoms). The current systems primarily use a categorical approach, which may not fully capture the complexity of mental health conditions.

EDA offers an objective way to measure physiological responses linked to emotional and psychological states, helping to reduce the subjectivity of self-reported symptoms. EDA has shown

potential as a biomarker for mental health conditions like depression and anxiety. It can distinguish between different phases of mood disorders and identify individuals at risk of suicidal behavior. EDA enables real-time monitoring of emotional and stress responses, providing continuous data that can track changes over time and in response to treatment. When used alongside traditional diagnostic tools, EDA provides complementary data that enhances the accuracy and depth of mental health assessments. Additionally, EDA is non-invasive, making it suitable for regular monitoring without causing discomfort to patients.

The DSM provides detailed and specific descriptions of mental disorders, aiding clinicians in the United States in making precise diagnoses. In contrast, the ICD, which is used globally, is broader and more flexible, allowing for greater clinical discretion. The ICD is designed for international use and incorporates a wider range of cultural perspectives, leading to more culturally sensitive diagnoses compared to the DSM, which is more focused on Western norms. The DSM requires the presence of functional impairments for a diagnosis, whereas the ICD does not always mandate this, affecting the threshold for diagnosing certain disorders. The ICD-11 has adopted a more dimensional approach for some disorders, such as personality disorders, considering severity and specific traits. Meanwhile, the DSM-5 retains a more categorical approach, listing specific disorders. These differences in classification can lead to variations in treatment approaches. For instance, the DSM's detailed criteria can guide specific therapeutic interventions, while the ICD's broader categories might influence more generalized treatment plans.

Mental disorders are complex and heterogeneous phenomena that affect the psychological well-being and functioning of individuals (Jain et al., 2024) . Over recent years there has been growing interest in anduse of self-report mental health and wellbeing measures out-side traditional clinical contexts. (Davies et al., 2024). They are associated with various biological, psychological, and social factors, and can have significant impacts on the quality of life and health outcomes of the affected individuals and their families (Xu et al., 2021). Therefore, it is important to classify mental disorders into meaningful and useful categories that can facilitate diagnosis, treatment, prevention, and research (DIAGNOSTIC AND STATISTICAL MANUAL OF DSM-5 TM, n.d.)

A mental disorder is characterized by a clinically significant disturbance in an individual's cognition, emotional regulation, or behaviour (Laganaro et al., 2023). It is usually associated with distress or impairment in important areas of functioning. There are many different types of mental disorders. Mental disorders may also be referred to as mental health conditions (Laganaro et al., 2023) The latter is a broader term covering mental disorders, psychosocial disabilities and (other) mental states associated with significant distress, impairment in functioning, or risk of self-harm. This fact sheet focuses on mental disorders as described by the International Classification of Diseases 11th Revision (ICD-11) (WHO, 2022).

However, mental disorder classification is not a simple or straightforward task (Fransvea et al., 2022) There are different systems and criteria for defining and categorizing mental disorders, and they may not always agree or correspond with each other (Xiong et al., 2020). Moreover, mental disorders are dynamic and evolving, and they may vary across cultures, contexts, and time periods (Bredström, 2019). Hence, there is a need for flexible and comprehensive approaches to analyze and understand mental disorder data (Bhavani et al., 2023).

Exploratory Data Analysis (EDA) is one such approach that can complement the existing systems of mental disorder classification (Wongsuphasawat et al., 2019). EDA is a data analysis strategy that emphasizes open-mindedness, creativity, and multiple perspectives (Da Poian et al., 2023). EDA aims to explore the data in as many ways as possible, without imposing any prior assumptions or models, until a plausible and coherent story emerges. EDA can help generate new hypotheses, identify patterns and outliers, and reveal the underlying structure and relationships of the data (Willa Dhany & Izhari, 2023)

EDA can be applied to various types of data, such as numerical, categorical, textual, or image data (Munappy et al., 2022). EDA can also use different methods and techniques, such as descriptive statistics, data visualization, data transformation, data reduction, clustering, or classification.(Wongsuphasawat et al., 2019). EDA can be performed using various software tools, such as R, Python, or MATLAB. EDA can also be integrated with other data analysis approaches, such as confirmatory data analysis (CDA), Bayesian statistics, resampling, or data mining (Bouabdallaoui et al., 2022).

Exploratory Data Analysis (EDA) aids in revealing hidden patterns and trends within complex datasets, offering new insights into the causes and progression of mental disorders. By examining data without preconceived notions, researchers can formulate new hypotheses about mental health conditions for future testing. EDA focuses on using visual tools to comprehend data, making it easier to spot anomalies, correlations, and other significant features that traditional statistical methods might overlook. Through EDA, researchers can identify and correct errors, outliers, and missing values, ensuring higher quality and more reliable results. EDA enhances the development of predictive models by providing a deeper understanding of the data, leading to more accurate and robust predictions of mental health outcomes. Additionally, by identifying individual differences and patterns in the data, EDA can help create personalized treatment plans tailored to the specific needs of patients.

In this paper, we will demonstrate how EDA can be used to analyze and understand mental disorder data from different sources and perspectives. We will use EDA methods to explore the characteristics, prevalence, and distribution of mental disorders, as well as the associations and interactions between mental disorders and other variables, such as demographics, symptoms, treatments, or outcomes. We will also compare and contrast the results of EDA with the existing systems of mental disorder classification, such as the DSM and the ICD. We hope that this paper will illustrate the potential and benefits of EDA for mental disorder research and practice.

Method

The data that will be used in the research is taken from the UCI Machine Learning Repository, consisting of 4980 data, 11 attributes and 2 classes. Therefore, it is necessary to clean the data to remove or correct missing, duplicate or inconsistent data to ensure data quality. Data is often normalized so that all features are on the same scale, which is important for certain machine learning algorithms. Datasets are usually divided into training set and test set; The training set is used to train the model, while the testing set is used to evaluate the model performance. This dataset is published in a repository and can be accessed by researchers and practitioners around the world for various machine learning applications.

The aim of this study is to apply EDA methods to mental disorder data from different sources and perspectives, and to compare and contrast the results with the existing systems of mental disorder classification, such as the DSM and the ICD. The research questions are what are the characteristics, prevalence, and distribution of mental disorders in different populations and regions, what are the associations and interactions between mental disorders and other variables, such as demographics, symptoms, treatments, or outcomes? How do the EDA results agree or disagree with the DSM and the ICD classifications of mental disorders?.

The EDA methods (Alcantara & Calvo, 2022) for this study are Descriptive statistics, such as mean, median, mode, standard deviation, range, frequency, and percentage, to summarize the basic features of the data, such as the number of cases, the distribution of variables, (Ma*, 2023) and the descriptive labels of categories. Data visualization, such as histograms, boxplots, scatterplots, bar charts, pie charts, and heatmaps, to display the data graphically and reveal the patterns, trends, outliers, and relationships of the data (Special Issue: Responsible Writing in Science, n.d.) Data transformation, such as log, square root, or inverse transformations, to improve the normality, linearity, or homoscedasticity of the data, and to reduce the effects of skewness, kurtosis, or outliers (*DescriptiveStatistics-Organizingsummarizingdescribingandpresentingdata*, n.d.) Data reduction, such as principal component analysis (PCA), factor analysis (FA), or cluster analysis (CA), to reduce the dimensionality of the data and identify the underlying structure and groups of the data. (Komorowski et al., 2016)

The EDA methods will be performed using R, Python, or MATLAB, (Rahmany et al., 2020) depending on the availability and suitability of the software tools. The EDA results will be interpreted and discussed in relation to the research questions and the existing systems of mental disorder classification. The EDA results will also be compared and contrasted with the results of CDA, Bayesian statistics, resampling, or data mining, if applicable. The EDA results will be presented in tables, figures, or text, with appropriate captions, labels, and references. The EDA results will be evaluated for their validity, reliability, and generalizability, as well as their implications and limitations for mental disorder research and practice (Olatunde & Falola, 2021).

Results and Discussions

The dataset comprises data from 120 individuals undergoing psychological evaluation, focusing on 17 critical symptoms that are instrumental in diagnosing Bipolar Disorder (both Manic and Depressive types), Major Depressive Disorder, as well as distinguishing individuals without these conditions. These symptoms include levels of Sadness, Exhaustion, Euphoria, Sleep Disturbances, Mood Fluctuations, Suicidal Ideation, Anorexia, Anxiety, Explanatory Attempts, Emotional Breakdowns, Disregard and Progression, Acknowledgment of Errors, Rumination, Hostile Reactions, Hopefulness, Sexual Drive, and Attention Focus. The 'Normal' classification is assigned to those who seek therapy for personal growth, skill development, and counseling, and while they may exhibit some psychological issues, their conditions are not as severe as those diagnosed with Major Depressive or Bipolar Disorders.

Table 1. Dataset Mental Disorder

Patient Number	Sadness	Euphoric	Exhausted	Sleep dissorder	...	Expert Diagnose	Notes
0	Patiant-01	Usually	Seldom	Sometimes	...	4 From 10	Bipolar Type-2
1	Patiant-02	Usually	Seldom	Usually	...	5 From 10	Depression
2	Patiant-03	Sometimes	Most-Often	Sometimes	...	7 From 10	Bipolar Type-1
3	Patiant-04	Usually	Seldom	Usually	...	2 From 10	Bipolar Type-2
4	Patiant-05	Usually	Usually	Sometimes	...	6 From 10	Normal
...	...	...	...	...	...	...	...
120	Patiant-120	Sometimes	Usually	Seldom		6 From 10	Normal

In this section, we present and discuss the results of applying EDA methods to mental disorder data from different sources and perspectives. We also compare and contrast the results with the existing systems of mental disorder classification.

Initially, we applied descriptive statistical methods and graphical representations to examine the traits, frequency, and geographical spread of mental health conditions across various demographics. We utilized data on patient behavior, specifically evaluating levels of Sadness, Fatigue, Elation, Insomnia, Emotional Variability, Suicidal Ideation, Loss of Appetite, Worry, Rationalization, Psychological Collapse, Dismissal and Advancement, Confession of Faults, Excessive Contemplation, Belligerent Reactions, Positivity, Libido, and Mental Focus (Schinka & Velicer, n.d.).

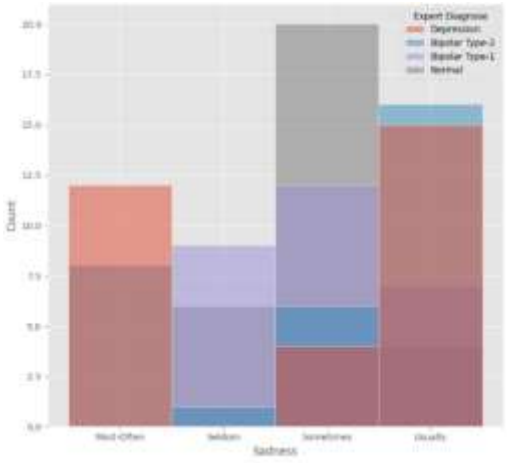

Figure 1. The Sadness Distribution

Figure 1 shows the histograms of the sadness distribution of the respondents in most-often, seldom, sometimes, usually datasets. We can see that the Psychology Patients dataset has a wider range of ages, from 18 to 99. sadness can cause depression in someone which if left unchecked can cause mental disorders in someone.

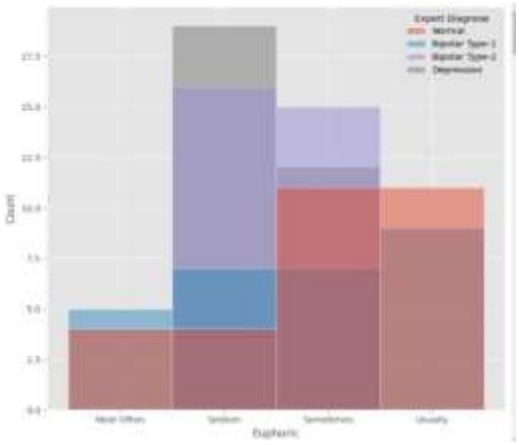

Figure 2. The Euphoric Distribution

Figure 2 shows the histograms of the Euphoric distribution of the respondents in most-often, seldom, sometimes, usually datasets.

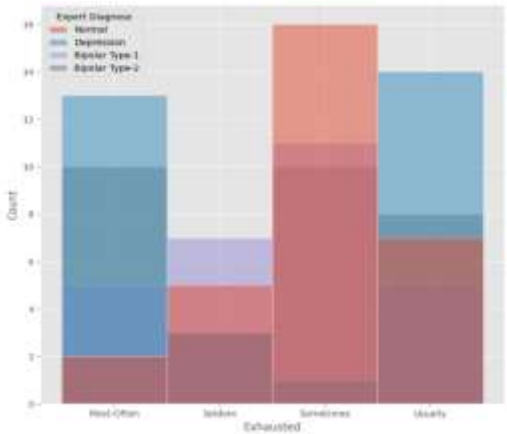

Figure 3. The Exhaustness distribution

Figure 3 shows the histograms of the Exhaustness distribution of the respondents in most-often, seldom, sometimes, usually datasets.

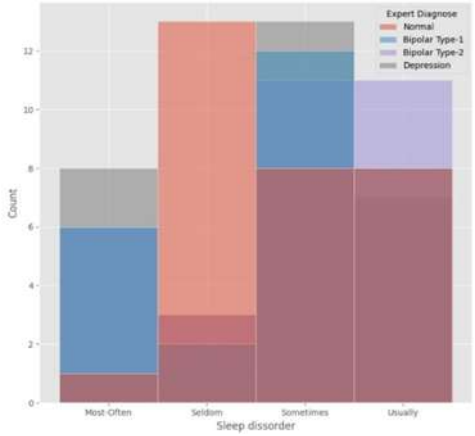

Figure 4. The Sleep Disorder Distribution

Figure 4 shows the histograms of the sleep disorder distribution of the respondents in most-often, seldom, sometimes, usually datasets.

People who are never sad and sometimes euphoric are diagnosed with Bipolar type 2 and people who are euphoric most-often are also Bipolar Type 1. So, it is better to be sometimes sad and happy usually because those people are most normal. People who are usually or most-often exhausted are depressed and those who are never exhausted are Bipolar Type 1 and sometimes exhausted people are normal. People who are sleepless most often are depressed whereas normal people are usually never sleepless. People with no mood swings, are anorexic, no authority respect, try-explanatory, highly aggressive responses, who don't ignore and move-on, have constant nervous breakdowns, and do not admit their mistakes are diagnosed with Bipolar Type 1. People who overthink a lot, don't admit their mistakes, have no aggressive responses, don't try explaining, no respect for authority and have higher levels of suicidal thoughts and mood swings are usually depressed. People who get less action are usually Bipolar Type 2 whereas those who get more action are usually Bipolar Type 1. People with lower concentration levels are Bipolar Type 2 and surprisingly with higher concentration are depressed. People with higher levels of optimism are Bipolar Type 1 and lower levels are depressed.

Conclusions

One of the main challenges is the high variability of the data, where the datasets used show significant variations in the symptoms reported by patients, making it difficult to identify consistent patterns. This study used a relatively small data sample, which may limit the generalization of the results, because the dataset only includes 4980 patient data with 17 main symptoms, which may not be representative enough for the wider population. Therefore, mental disorders need to be classified categorically or dimensionally. The categorical approach views mental disorders as distinct entities, while the dimensional approach views them as a spectrum. These differences can affect the results of analysis and interpretation of data.

In this paper, we demonstrated how EDA can be used to analyze and understand mental disorder data from different sources and perspectives. We used EDA methods to explore the characteristics, prevalence, and distribution of mental disorders, as well as the associations and interactions between mental disorders and other variables. We also compared and contrasted the results of EDA with the existing systems of mental disorder classification, such as the DSM and the ICD. We showed that EDA can provide a more comprehensive and nuanced understanding of mental disorder data, as well as generate new hypotheses, identify patterns and outliers, and reveal the underlying structure and relationships of the data. We also discussed the challenges and limitations of EDA, such as the variability and complexity of mental disorder data, the lack of biological markers for most mental disorders, and the potential for stigma and discrimination associated with mental disorder labels. We hope that this paper will illustrate the potential and benefits of EDA for mental disorder research and practice.

Exploratory Data Analysis (EDA) and the DSM/ICD classifications offer different yet complementary approaches to understanding mental disorders. Here's a detailed explanation of how they qualitatively differ and align objective vs. Subjective Data, EDA focuses on uncovering patterns and trends in data through statistical and visual methods. It relies on objective data analysis to identify correlations and anomalies. DSM/ICD is primarily based on clinical observations, self-reported symptoms, and expert consensus. These classifications can be more subjective, influenced by cultural and societal norms, Data-Driven Insights vs. Established Criteria, EDA is Generates new hypotheses and insights by exploring data without preconceived notions. It can reveal unexpected relationships and trends that may not be captured by existing diagnostic criteria. DSM/ICD is Provides established criteria for diagnosing mental disorders, ensuring consistency and reliability in clinical practice. These criteria are based on extensive research and expert consensus, flexibility vs. Standardization, EDA is Offers flexibility in analyzing data, allowing researchers to adapt their methods based on the data's characteristics. This can lead to more personalized and nuanced understandings of mental health conditions. DSM/ICD is Emphasizes standardization, providing a consistent framework for diagnosing and treating mental disorders across different settings and populations.

Complementary Data, EDA can enhance DSM/ICD classifications by providing additional data that supports or refines existing diagnostic criteria. For example, EDA can identify physiological markers or behavioral patterns that align with DSM/ICD-defined disorders. DSM/ICD can benefit from EDA's

insights to update and improve diagnostic criteria, making them more accurate and reflective of real-world data. Improving Diagnostic Accuracy, EDA helps in identifying subtle differences and patterns that might be missed by traditional diagnostic methods. This can lead to more accurate and early diagnoses. DSM/ICD provides a structured approach to diagnosis, which, when combined with EDA's findings, can result in more comprehensive and precise assessments. Personalized Treatment Plans, EDA by identifying individual differences and patterns in the data, EDA can contribute to the development of personalized treatment plans tailored to the specific needs of patients. DSM/ICD is offers a framework for treatment based on established diagnostic categories, which can be further refined and personalized using EDA insights.

Exploratory Data Analysis (EDA) encounters several specific challenges and limitations. Researchers often face significant data variability due to differences in how symptoms are reported and recorded across various populations, making it difficult to identify consistent patterns or trends. The quality and representativeness of the data sample can also be a major limitation. For instance, if the sample lacks diversity, the findings may not be generalizable to the broader population. This is particularly relevant in mental health research, where cultural and demographic factors are crucial. EDA frequently uncovers errors, outliers, and missing values in the data. While these issues can be addressed to some extent, they can still impact the reliability of the findings. For example, missing data on key variables can limit the ability to draw robust conclusions.

Exploratory Data Analysis (EDA) in mental disorders research can have several practical implications for clinicians, policy makers, and other researchers by analyzing physiological data and self-reported symptoms, doctors can gain a more comprehensive understanding of the patient's condition., continuous monitoring of electrodermal activity can help track a patient's response to therapy in real-time, allowing for timely treatment adjustments, data-driven insights can help policymakers allocate resources more efficiently, targeting areas with the highest need for mental health services, policies can be designed to address specific mental health challenges faced by different cultural groups, improving access and effectiveness of mental health services, researchers can use EDA findings to design studies that explore the underlying mechanisms of mental disorders, resulting in new discoveries and advances in the field, by improving data quality, researchers can produce more robust and generalizable findings, thereby contributing to the overall progress of mental health research, when used alongside traditional diagnostic tools, EDA provides complementary data that increases the depth and accuracy of mental health assessments, EDA is a non-invasive method, making it suitable for routine monitoring without causing discomfort to the patient.

References

Alcantara, T., & Calvo, H. (2022). *Exploratory Data Analysis for the Automatic Detection of Question Paraphrasing in Collaborative Environments* (pp. 193–211). https://doi.org/10.1007/978-3-031-19496-2_15

Bhavani, V., Sirivarshitha, A. K., Sravani, K., & Priya, K. S. (2023). Music Classification and Mental Health Analysis using Exploratory Data Analysis. *2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)*, 555–561. https://doi.org/10.1109/ICIDCA56705.2023.10099605

Bouabdallaoui, I., Guerouate, F., Bouhaddour, S., Saadi, C., & Sbihi, M. (2022). *Advanced Exploratory Data Analysis for Moroccan Shopping Places in TripAdvisor* (pp. 257–271). https://doi.org/10.1007/978-3-031-20319-0_20

Bredström, A. (2019). Culture and Context in Mental Health Diagnosing: Scrutinizing the DSM-5 Revision. *Journal of Medical Humanities*, *40*(3), 347–363. https://doi.org/10.1007/s10912-017-9501-1

Da Poian, V., Theiling, B., Clough, L., McKinney, B., Major, J., Chen, J., & Hörst, S. (2023). Exploratory data analysis (EDA) machine learning approaches for ocean world analog mass spectrometry. *Frontiers in Astronomy and Space Sciences*, *10*. https://doi.org/10.3389/fspas.2023.1134141

Davies, J., McKenna, M., Denner, K., Bayley, J., & Morgan, M. (2024). The emoji current mood and experience scale: the development and initial validation of an ultra-brief, literacy independent measure of psychological health. *Journal of Mental Health*, *33*(2), 218–226. https://doi.org/10.1080/09638237.2022.2069694

DescriptiveStatistics-organizingsummarizingdescribingandpresentingdata. (n.d.).

DIAGNOSTIC AND STATISTICAL MANUAL OF DSM-5 ™. (n.d.).

Fransvea, G., Vendrame, E., Patron, E., Mura, F., Cipriani, C., Gentili, C., & Mannini, A. (2022). Inertial measurement units-based assessment of gait to identify depression. *Gait & Posture*, *97*, 5–6. https://doi.org/10.1016/j.gaitpost.2022.09.016

Jain, D., Arora, S., Jha, C. K., & Malik, G. (2024). Text classification models for personality disorders identification. *Social Network Analysis and Mining*, *14*(1), 64. https://doi.org/10.1007/s13278-024-01219-8

Komorowski, M., Marshall, D. C., Salciccioli, J. D., & Crutain, Y. (2016). Exploratory data analysis. In *Secondary Analysis of Electronic Health Records* (pp. 185–203). Springer International Publishing. https://doi.org/10.1007/978-3-319-43742-2_15

Laganaro, F., Mazza, M., Marano, G., Piuzzi, E., & Pallotti, A. (2023). Classification-Based Screening of Depressive Disorder Patients Through Graph, Handwriting and Voice Signals. *2023 International Workshop on Biomedical Applications, Technologies and Sensors (BATS)*, 6–10. https://doi.org/10.1109/BATS59463.2023.10303104

Ma*, X. (2023). Data science for geoscience: Recent progress and future trends from the perspective of a data life cycle. In *Recent Advancement in Geoinformatics and Data Science* (pp. 57–69). Geological Society of America. https://doi.org/10.1130/2022.2558(05)

Munappy, A. R., Bosch, J., Olsson, H. H., Arpteg, A., & Brinne, B. (2022). Data management for production quality deep learning models: Challenges and solutions. *Journal of Systems and Software*, *191*, 111359. https://doi.org/10.1016/j.jss.2022.111359

Olatunde, O., & Falola, B. (2021). *Classification of Mental Health Disorders*. https://doi.org/10.13140/RG.2.2.34918.60483

Rahmany, M., Zin, A. M., & Sundararajan, E. A. (2020). COMPARING TOOLS PROVIDED BY PYTHON AND R FOR EXPLORATORY DATA ANALYSIS. *IJISCS (International Journal of Information System and Computer Science)*, *4*(3), 131. https://doi.org/10.56327/ijiscs.v4i3.933

Schinka, J. A., & Velicer, W. F. (n.d.). *HANDBOOK of PSYCHOLOGY VOLUME 2 RESEARCH METHODS IN PSYCHOLOGY*.

Special issue: Responsible writing in science. (n.d.).

WHO. (2022, June 8). *Mental disorders*. Https://Www.Who.Int/News-Room/Fact-Sheets/Detail/Mental-Disorders/?Gad_source=1&gclid=CjwKCAjwkJm0BhBxEiwAwT1AXOo8VxQ1at-WjEgzZRwwvtS7jn2BzDjwuXB4aGJbW-Dp3fs5EagGXRoCUeUQAvD_BwE.

Willa Dhany, H., & Izhari, F. (2023). Journal of Intelligent Decision Support System (IDSS) Exploratory Data Analysis (EDA) methods for healthcare classification. In *Journal of Intelligent Decision Support System (IDSS)* (Vol. 6, Issue 4).

Wongsuphasawat, K., Liu, Y., & Heer, J. (2019). *Goals, Process, and Challenges of Exploratory Data Analysis: An Interview Study*.

Xiong, Z., Zhang, X., Zhang, M., & Cao, B. (2020). Predicting Features of Human Mental Disorders through Methylation Profile and Machine Learning Models. *2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, 67–75. https://doi.org/10.1109/MLBDBI51377.2020.00019

Xu, L., Zheng, Y., Xu, D., & Xu, L. (2021). Predicting the Preference for Sad Music: The Role of Gender, Personality, and Audio Features. *IEEE Access*, *9*, 92952–92963. https://doi.org/10.1109/ACCESS.2021.3090940