Journal of Computer Science

Original Research Paper

Analysis of Student Mental Health Dataset Using Mining Techniques

[1]Yemima Monica Geasela, [1]Devi Yurisca Bernanda, [1]Johanes Fernandes Andry, [1]Christian Kurniadi Jusuf, [1]Samuel Winata, [2]Lydia and [3]Shierly Everlin

[1]Department of Information System, Universitas Bunda Mulia, Jakarta, Indonesia
[2]Department of Accounting, Universitas Bunda Mulia, Jakarta, Indonesia
[3]Department of Visual Communication Design, Universitas Bunda Mulia, Jakarta, Indonesia

Article history
Received: 23-08-2023
Revised: 06-11-2023
Accepted: 20-11-2023

Corresponding Author:
Yemima Monica Geasela
Department of Information System, Universitas Bunda Mulia, Jakarta, Indonesia
Email: yemima.geasela@gmail.com

Abstract: This study utilizes a decision tree model in RapidMiner to analyze a dataset from Kaggle, comprising 200 student records. Among these, 70 students reported mental health issues, while 130 did not. Strikingly, a significant majority of 58 out of the 70 students with mental health concerns do not seek assistance from professionals. This study underscores the pressing issue of underutilization of mental health services among students and offers practical solutions, such as enhancing awareness and education, improving access to mental health services, providing peer support, and addressing underlying issues. The research design includes data collection methods that maintained ethical standards and the decision tree model's application for analysis. This study's contribution lies in its identification of the prevalence of students with mental health issues who do not seek help and the proposed solutions to address this critical issue.

Keywords: Big Data, Mental Health, Educational, Institutions, Rapid Miner, Decision Tree

Introduction

The importance of mental health has increased recently among college students. Over 60% of college students reported having overwhelming anxiety, while approximately 40% had depressive symptoms, according to the American College Health Association. Furthermore, among college students, suicide happens to be the second-leading reason for death (Lederer and Hoban, 2022). These stark numbers underscore the pressing need for academic institutions to provide students with accessible mental health services and support. College students face a variety of stressors, including academic pressures, social isolation, and financial burdens, which contribute to poor mental health outcomes (Grasdalsmoen *et al*., 2020). Effective interventions and mental health support not only have the potential to improve students' academic achievements but also to enhance their overall well-being. Studies have indicated that high levels of stress and mental health issues are linked to poorer academic performance (Gao *et al*., 2020), emphasizing the interconnectedness of mental well-being and academic success among students. The COVID-19 pandemic has further exacerbated mental health challenges faced by college students, with reports of increased stress, anxiety, and depression levels, coupled with difficulties related to remote learning and social isolation (Zhang *et al*., 2021). Universities must prioritize mental health services and support to address these issues. This includes improving access to mental health services, reducing the stigma associated with seeking help, and implementing targeted interventions to assist students during times of stress or crisis (Grøtan *et al*., 2019). While online solutions like counseling services and mental health apps have been introduced, it is essential to recognize that not all students can benefit equally from these strategies, some prefer in-person interactions or lack access to the necessary technology or internet connectivity (Lattie *et al*., 2020). To ensure all students have the care and support they need, universities must continually assess and enhance their mental health services.

The current level of mental health challenges among college students underscores the importance of offering robust mental health services and support. Promising approaches include peer support groups, known to be effective in promoting mental health and reducing stigma (Conley *et al*., 2020). Additionally, recent studies have indicated the value of teletherapy and telepsychiatry services in reaching students who may face barriers to accessing traditional, in-person therapies due to stigma or

distance (Ahuvia *et al*., 2022). Turn our attention to data mining techniques, specifically decision tree analysis, which offers the potential to identify factors influencing college students' mental health challenges and predict their future mental health status (Bo and Gao, 2022).

Decision tree analysis is a valuable data mining tool used to investigate and understand the factors influencing student mental health. Previous research has demonstrated its effectiveness in predicting mental health issues in college students, such as depression and anxiety, by considering behavioral and demographic factors (Ebert *et al*., 2019). By analyzing large datasets, like the student mental health dataset, using the decision tree as a mining technique, authors can gain valuable insights into the intricate relationships between variables affecting mental health outcomes. These insights serve as a foundation for developing effective interventions and support services for college students.

Big Data

Big data is a collection of datasets that are very large and rich in variety and they grow exponentially over time. Data is so extensive and complex that traditional data management systems cannot effectively store and process it. Big data analytics interactively analyze and visualize data, connecting numerous variables related to research topics. It enables researchers to access vast amounts of data from cyber databases, emphasizing the relationship between the digital and human realms to make the research process interactive. Big data is characterized by volume, velocity, and variety, often referred to as the 3 Vs (Madyatmadja *et al*., 2021). Large data volumes have significant capacities, ranging from sizes between MB, GB, TB, PB and ZB. A substantial amount of big data makes it possible to include a vast array of existing data. Velocity in big data demonstrates a very high transfer speed, allowing users to access data in real time within minutes or even seconds. Variety in big data consists of various data types, including structured (e.g., tables), semi-structured (XML documents), and unstructured data (documents, emails, text messages, audio, video, images, graphics, etc.).

Besides 3V, big data has other characteristics, namely veracity and value. The term "veracity" describes the caliber and correctness of data. The more data you have, the more difficult it becomes to manage it properly. If there are any errors or mistakes along the way, the consequences are enormous. From a deeper perspective, big data is not just data that is generated. However, it should correctly identify the data and provide benefits to the user. Whereas value refers to the added value and usefulness of big data. Because of the focus on making information valuable and the idea that big data is codified and thus easily communicated and shared, big data can be seen as explicit information from a knowledge-based perspective (Franke and Hiebl, 2022).

The goal of companies that use big data in their systems is to improve operations, provide better customer service, create more personalized campaigns, and take other actions that ultimately result in profits (Andry *et al*., 2023a). In the highly complex healthcare sector, where integration with technology is critical, big data also offers ways and opportunities to ensure the efficient use of resources and quality patient care. In addition, it is widely recognized that BI can provide significant benefits to healthcare organizations, such as Increased clinical performance, efficient use of personnel, process improvement, and cost efficiency. Although the amount of information generated increases, improvement in the quality of clinical performance can also reduce costs that are issued. This growing amount of data (known as big data) holds promise not only for supporting various clinical functions but also for supporting decision-making (Andry *et al*., 2023b).

Student Mental Health

Mental health is a very important aspect for everyone, especially for the nation's next generation who are still in the world of education, either primary education or tertiary education. Mental health can be influenced by many factors including genetics, family, friendships, lifestyle, society, and many other factors. These factors can have positive or negative effects on college students. However, many students are not aware of the positive and negative impacts arising from these factors, causing them to lose their mental health because they are too focused on assignments, organizations, class schedules, and the demands of the people around them. However, it turns out that these tensions have pushed some students to feel anxious and depressed, affecting their mental health. Students' lack of knowledge about mental health makes them unable to control and control themselves, even though many students who experience mental health do not consult doctors or experts such as psychiatrists in a fashionable way (Essangri *et al*., 2021).

Transitioning from a pandemic to an epidemic and returning to normal activities also affect our ability to adapt to the environment. Students often experience this concerning their mental state. The term "mental health" refers to a person's cognitive, behavioral, and emotional well-being, which influences how they feel, think, and act. It also has an impact on how they go about their daily lives. and physical health student educational activities. This online survey was conducted with a total of 207 university students in Pakistan to collect information on socio-demographics, concerns, and fears related to COVID-19 and mental health issues. Validated tools; Perceived Stress Scale (PSS), Generalized Anxiety Disorder Scale (GAD-7), and Patient Health Questionnaire (PHQ-9) depression were used to assess stress, anxiety, and depression. According to the survey, about

14% of students suffer from severe stress and anxiety and 8.2% suffer from major depression (Ali *et al*., 2023).

Beyond seeing a psychiatrist, people with depression and mild mental illness may be cured by paying more attention to the patient than to their loved ones., give this patient some peace. Providing reassurance and comfort also helps patients slowly recover from panic. People with mild mental illness usually need someone to support and support them. Especially if the people around you have positive personalities, it will help them a lot (Niemiec, 2023).

Classification

Classification is the process of classifying or subdividing unlabeled data or objects into categories or labels determined based on certain attributes. In classification, a machine learning algorithm or model learns patterns from identified training data and applies those patterns to new data to predict appropriate classes or labels. Classification can be done manually without the aid of algorithms or intelligent computing techniques using multiple algorithms such as logistic regression, Naive Bayes, decision trees, random forests, k-nearest neighbors, and artificial neural networks. Decision trees are commonly used in classification studies. A decision tree builds multiple observations from a sample to make a classification decision. Training requires obedience as well as information sharing. This selects the order in which to check the observations (Trivedi and Panigrahi, 2018).

This decision tree is typically constructed from one or more original nodes. The note then branches to display a list of available options. This decision tree can be used to simplify complex decision-making. For those who often have trouble making big decisions, this method can be used as an effective solution to overcome their problems. A decision tree starts with the problem you are facing. Then create a new branch containing all the options or choices in question. Each previously created option as a result. The following branch will help you write what the result will be.

The existence of decision trees is expected to generate survival-friendly decision trees and provide information and decisions about the mental health of students being studied by researchers. A classification procedure using a decision tree algorithm facilitates decision-making that is transferred from research results to everyday life. The classification procedure provided by the decision tree algorithm also has strengths such as naturally generated rules, processed continuous and categorical variables, and the ability to indicate the relative importance of variables for classification (Marzukhi *et al*., 2021).

Materials and Methods

The methodology is a crucial aspect used by authors to analyze, work on, and solve problems at hand. The theoretical or scientific framework serves as a guide to the scientific methods applied in conducting research. In this study, the research process involves selecting data, processing data, analyzing data and implementation can be seen in Fig. 1. Some of the steps for research carried out in this scientific are:

1. Data selection: In the data selection phase, the dataset was carefully chosen by the authors from www.kaggle.com. The dataset consists of 200 data points, with 130 not having any mental health issues and 70 students facing mental health challenges. This step involved considering various criteria to ensure the dataset's relevance to our research
2. Data preprocessing: Following data selection, the authors imported the chosen dataset into RapidMiner studio, a powerful data science platform. Data preprocessing is a critical step to ensure data quality. During this phase, the dataset was cleaned, missing values were handled and feature engineering was performed to prepare the data for analysis. These preprocessing steps are essential to obtain reliable and accurate results
3. Data analyst: In the data analysis stage, the authors utilized RapidMiner studio to conduct a comprehensive analysis of the dataset. The primary analysis technique employed was the development of a decision tree model. Decision trees are powerful tools for visualizing and understanding complex data relationships. Decision tree algorithms were applied to the dataset, resulting in a structured tree representation that revealed critical insights into the factors influencing students' mental health challenges and their likelihood to seek professional help
4. Implementation: The implementation phase involved the application of data mining techniques using the RapidMiner platform. The authors used this tool to construct, fine-tune, and evaluate the decision tree model. It enabled the identification of patterns, correlations, and predictive features within the data, ultimately leading to the generation of valuable insights

Method of Collecting Data

This study employed a method of data collection that involved utilizing public data sources from the Kaggle platform. Kaggle is a widely recognized platform for data science, machine learning, and data mining practitioners, providing access to a diverse range of datasets that are essential for research and analysis in these fields.

The specific dataset used in this study was selected from one of the competitions hosted on Kaggle. The selection process adhered to predetermined criteria to ensure the dataset's relevance to the research objectives. The dataset chosen for this study comprises a total of 200 data points, making it substantial enough to yield meaningful insights.

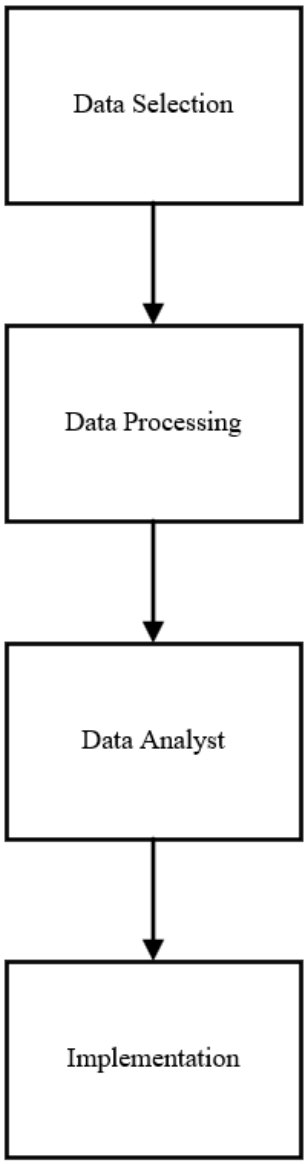

Fig. 1: Research stage

The dataset consists of several key variables that are directly related to the research topic, specifically focusing on the mental health of college students. These variables include information related to academic performance, socio-demographic factors, and, most importantly, whether the students sought professional help for their mental health challenges.

Furthermore, during the data collection process, the author made sure to review the dataset for any anomalies or inconsistencies to ensure data quality and reliability. Data cleaning steps were implemented to address missing values and maintain the dataset's integrity. The comprehensive approach to data collection and preparation is essential to guarantee the accuracy and validity of the research findings.

Data Analysis Technique

RapidMiner studio served as our primary tool for data analysis, offering a versatile environment for predictive data analysis, descriptive data analysis, and statistical analysis (AL-Ma'aitah, 2020). These techniques enabled us to derive valuable insights and knowledge that informed our decision-making throughout the research process. Our data analysis process in RapidMiner involved a combination of descriptive and inferential statistics. Descriptive statistics were employed for a comprehensive examination of the dataset, enabling us to create data summaries and visual representations without drawing general conclusions. In contrast, predictive analytics focused on making data-driven predictions based on the available information, while inferential statistics allowed us to conduct significance tests for drawing conclusions from a sample (Uska *et al*., 2020). In our study, harnessed the power of predictive models, specifically decision trees, for classification purposes. Decision trees proved to be an effective model at different stages of our research. They were initially employed as the foundational model in the modeling process and later considered as the final model choice after evaluating various other techniques. Emphasized the importance of precise wording when articulating research objectives and selecting statistical analysis techniques. By carefully choosing phrases such as describing, identifying, evaluating, checking, designing, reviewing, showing, and measuring, among others, and tailoring our statistical procedures to align with our research questions and study design.

Results

In this section, provide a detailed discussion of the data attributes and their significance. The "student mental health dataset" sourced from Kaggle, serves as the foundation for our analysis. This dataset, collected from students worldwide, comprises 200 instances, each enriched with a combination of numeric and alphanumeric data. These instances are defined by various attributes, each playing a crucial role in our analysis. Table 1 provides a comprehensive list of these attributes along with their explanations.

Table 2 offers a comprehensive understanding of the dataset's attributes, enabling a clear perspective on the variables to be analyzed and their role in predicting student mental health outcomes.

Implementation of the Dataset in RapidMiner

The successful implementation of the dataset in RapidMiner was a critical step in our analysis. This phase involved importing both the training and testing data into RapidMiner, which included all 200 instances from the dataset.

Table 1: Attribute explanation

No.	Attribute	Description
1.	Choose your gender	Indicate the gender of the related student
2.	Age	Indicate the age of the related student
3.	What is your course	Indicate the program of study taken by the related student
4.	Your current year of study	Indicate what year of education the related student is currently in
5.	What is your CGPA	Indicate the CGPA data of the related student
6.	Marital status	Indicate the marital status of the related student
7.	Do you have depression	Indicate whether the related student is experiencing depression
8.	Do you have anxiety	Indicate whether the related student is having anxiety problems
9.	Do you have panic attacks	Indicate whether the related student has panic problems
10.	Did you seek any specialist for treatment	Indicate whether the related student has ever undergone medical treatment regarding their mental health

Table 2: Data type

Instance	Data type
Choose your gender, marital status, do you have depression, do you have anxiety, do you have panic attacks Did you seek any specialist for a treatment	Binominal
What is your course, your current year of study	Polynomial
What is your CGPA	Real
Age	Integer

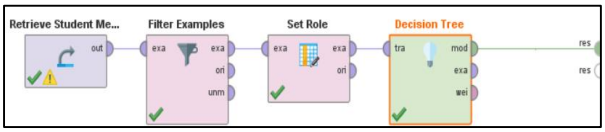

Fig. 2: Decision tree process model using RapidMiner

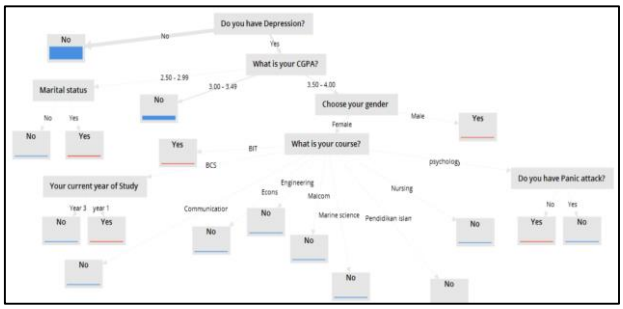

Fig. 3: The formed decision tree model

Tree

```
Do you have Depression? = No: No {No=130, Yes=0}
Do you have Depression? = Yes
|   What is your CGPA? = 2.50 - 2.99
|   |   Marital status = No: No {No=2, Yes=0}
|   |   Marital status = Yes: Yes {No=0, Yes=4}
|   What is your CGPA? = 3.00 - 3.49: No {No=38, Yes=0}
|   What is your CGPA? = 3.50 - 4.00
|   |   Choose your gender = Female
|   |   |   What is your course? = BCS
|   |   |   |   Your current year of Study = Year 3: No {No=2, Yes=0}
|   |   |   |   Your current year of Study = year 1: Yes {No=0, Yes=2}
|   |   |   What is your course? = BIT: Yes {No=0, Yes=2}
|   |   |   What is your course? = Communication: No {No=2, Yes=0}
|   |   |   What is your course? = Econs: No {No=2, Yes=0}
|   |   |   What is your course? = Engineering: No {No=2, Yes=0}
|   |   |   What is your course? = Malcom: No {No=2, Yes=0}
|   |   |   What is your course? = Marine science: No {No=2, Yes=0}
|   |   |   What is your course? = Nursing: No {No=2, Yes=0}
|   |   |   What is your course? = Pendidikan islam: No {No=2, Yes=0}
|   |   |   What is your course? = psychology
|   |   |   |   Do you have Panic attack? = No: Yes {No=0, Yes=2}
|   |   |   |   Do you have Panic attack? = Yes: No {No=2, Yes=0}
|   |   Choose your gender = Male: Yes {No=0, Yes=2}
```

Fig. 4: Decision tree description

Attribute-type assignment was a meticulous process, as it directly influenced the accuracy and reliability of the final results. In this context, several attribute types were utilized, each carefully chosen to match the nature of the data:

a. Binominal: For attributes with two possible values, such as YES/NO or 0/1
b. Polynominal: For attributes with more than two possible values
c. Real: For attributes with decimal point numerical values
d. Integer: For attributes with whole number values, without decimals

This attribute type assignment was crucial to ensure the dataset was accurately represented within the RapidMiner environment. Any misassignment could have led to inaccurate results.

The dataset underwent further preprocessing to optimize it for the data mining process. This included transforming the data and reducing it by removing any unnecessary attributes, ultimately streamlining the dataset for modeling.

Modeling marked a pivotal phase where data mining techniques were directly applied. Selected the decision tree algorithm as our primary tool for analysis. The decision tree model's outcome can be observed in Figs. 3-4.

Process Implementation Using RapidMiner

To begin, launch RapidMiner and create a new repository in the repository table. Import the student performance dataset into the repository and then use the retrieve operator to import the student mental health dataset into the rapid miner design process field.

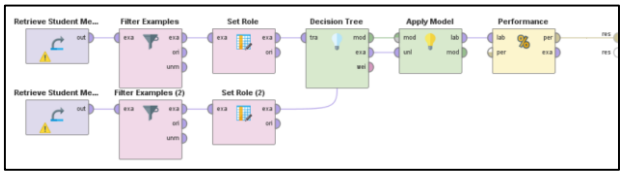

Fig. 5: Decision tree performance process model

Apply the filter example operator to filter the data and then utilize the set role operator to designate the attribute's role as a label. Introduce the decision tree operator into the field and connect it to the previous operators. Finally, add the apply model and performance operators to the design field, reintroduce the performance dataset as testing data, and connect it to the apply model operator. The result will resemble Figs. 2 and 5.

Figures 3-4, can see how students who suffer from depression and other mental health issues tend not to seek professional help. These findings emphasize the importance of enhancing mental health awareness and support on campus.

Discussion

In the context of our research on students' utilization of mental health services, it's essential to consider Shafiee's study, 'Prediction of mental health problems among higher education students using machine learning.' Shafiee's work, conducted in Malaysia, in 2020, addresses the prevalence of mental health issues among higher education students. It highlights the challenges in identifying contributing factors and the utilization of machine learning to predict mental health problems among this demographic. In our study, which analyzed data from 200 students, focused on students' decision-making regarding professional assistance for mental health issues. Our findings indicate that a significant majority of students with mental health concerns do not seek professional help, a concerning trend that aligns with the broader context described by Shafiee.

High accuracy in our modeling process is essential, as it enables us to detect whether a student is seeking professional help. Achieving 100% accuracy, precision and recall signifies the model's effectiveness in predicting students' mental health status and their likelihood to seek assistance. In comparison with traditional statistical methods like linear regression, which might assume linear relationships and have limited capacity to capture non-linear dependencies, decision tree analysis offers a distinct advantage. It excels in uncovering hidden patterns in the data and doesn't rely on the assumption of linear relationships. These results have profound implications for our research question and emphasize the need for increased mental health support and awareness on campus.

Conclusion

Based on the analysis of the student mental health dataset using a decision tree in RapidMiner, it can be concluded that the majority of students with mental health problems tend not to seek any help from a specialist. This finding is a cause for concern, as it suggests that many students may be suffering in silence and not getting the support they need.

Furthermore, the decision tree model developed in RapidMiner shows promising results, with an accuracy, precision, and recall rate of 100. This indicates that the model is highly effective in predicting whether a student is likely to seek help for their mental health issues or not. However, it is important to note that while the model is highly accurate, it should not be the sole basis for decision-making regarding a student's mental health. It has certain limitations that should be taken into account:

1. Data limitations: The accuracy of the model is contingent on the quality and representativeness of the data. Our model relies on the dataset collected from 200 students and any biases or inaccuracies in the data could affect the results. Additionally, the dataset's characteristics, such as the balance of classes (students seeking help and those who don't), might have contributed to the high accuracy. It's important to acknowledge that the small dataset size and class imbalances might have influenced the model's performance. Further research with a larger and more diverse dataset would provide a more realistic assessment of the model's generalization capabilities
2. Generalizability: This study focused on a specific group of students and may not be fully generalizable to all student populations. Variations in cultural, institutional, or regional factors could influence students' decisions regarding mental health services
3. Model scope: The decision tree model is designed to predict the likelihood of students seeking help based on specific variables in the dataset. It may not account for all potential factors influencing a student's mental health decisions, such as personal experiences and unique circumstances
4. Dynamic nature of mental health: Mental health is a complex and evolving aspect of an individual's well-being. The model's predictions are based on static data and may not capture changes in a student's mental health status over time

In conclusion, the findings of this study highlight the need for greater awareness and support for students with mental health problems, as well as the potential of decision tree models to aid in identifying and supporting

these students. To address the problem of students with mental health problems not seeking help from a specialist, solutions include improving awareness and education, increasing access to mental health services, providing peer support, and addressing underlying issues. These solutions can improve the mental health outcomes of students and ensure they receive the support they need. It is imperative to use the model's predictions as a supportive tool in a broader context that considers individual circumstances and experiences.

Acknowledgment

The manuscript's authors express appreciation and gratitude to our university especially to thank Center for Research and Community Service or Pusat Penelitian dan Pengabdian Masyarakat (P3M) University of Bunda Mulia, Jakarta, Indonesia for support for supporting this research.

Funding Information

The researchers would like to thank to University of Bunda Mulia for funding this research fiscal year 2023/2024, especially to the Center for Research and Community Service or Pusat Penelitian dan Pengabdian Masyarakat (P3M) with grant number: 162/ST.Dir.P3M&QA.01/X/2023.

Author's Contributions

Yemima Monica Geasela: Conceptualized the research problem, conducted a literature reviewed, interpreting the results, and written the manuscript for submission to the Journal of Counseling and Student Services (JCS).

Devi Yurisca Bernanda: Implementing the research design, collecting and analyzed the data, and examed the correlation between student mental health and academic performance.

Johanes Fernandes Andry: Defined the research questions, obtaining the data on student mental health from various sources such as Kaggle, and applying the decision tree model for data analysis.

Christian Kurniadi Jusuf: Performed the decision tree analysis, which allows the identification of factors that affect college students' mental health issues and the prediction of their future mental health outcomes.

Samuel Winata: Preparing the data for analysis, training and tested the decision tree model using the RapidMiner application, and evaluating the model's accuracy and validity based on static data.

Lydia: Explaining the attributes and data types of the source data on student mental health, and tested the data using the RapidMiner application.

Shierly Everlin: Processing the data using the RapidMiner application, constructing, fine-tuning, and evaluating the decision tree model, and discovering patterns, correlations, and predictive features from the data.

Ethics

This article is original and unpublished. In conducting this research, significant emphasis was placed on ethical considerations, particularly given the sensitivity of the topic, which revolves around mental health issues. The original dataset from Kaggle was anonymized and no personally identifiable information was included.

References

Ahuvia, I. L., Sung, J. Y., Dobias, M. L., Nelson, B. D., Richmond, L. L., London, B., & Schleider, J. L. (2022). College student interest in teletherapy and self-guided mental health supports during the COVID-19 pandemic. *Journal of American College Health*, 1-7. https://doi.org/10.1080/07448481.2022.2062245

Ali, N. A., Feroz, A., Akber, N., & Khoja, A. (2023). Students' psychological state and its associated factors during an outbreak of COVID-19 in Pakistan: a web-based survey. *Arab Gulf Journal of Scientific Research*. https://doi.org/10.1108/AGJSR-12-2022-0309

AL-Ma'aitah, M. A. (2020). Utilizing of big data and predictive analytics capability in crisis management. *J. Comput. Sci*, *16*(3), 295-304. https://doi.org/10.3844/jcssp.2020.295.304

Andry, J. F., Chakir, A., Silalahi, R. M. P., Liliana, L., & Clara, M. (2023a). The Impact of Business IT Strategic Alignment on Corporate Performance. *Journal of Computer Science*, *19*(3), 389-398. https://doi.org/10.3844/jcssp.2023.389.398

Andry, J. F., Sibarani, R., & Yefta, V. N. (2023b). Analysis of Big Data Football Club Market Value Using K-Means and Linear Regression Mining Methods. *Journal of Computer Science*, *19*(2), 286-294. https://doi.org/10.3844/jcssp.2023.286.294

Bo, Q., & Gao, X. (2022, January). Mental State Prediction of College Students Based on Decision Tree. In *LISS 2021: Proceedings of the 11th International Conference on Logistics, Informatics and Service Sciences* (pp. 345-357). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-16-8656-6_32

Conley, C. S., Hundert, C. G., Charles, J. L., Huguenel, B. M., Al-khouja, M., Qin, S., ... & Corrigan, P. W. (2020). Honest, open, proud-college: Effectiveness of a peer-led small-group intervention for reducing the stigma of mental illness. *Stigma and Health*, *5*(2), 168. https://doi.org/10.1037/sah0000185

Ebert, D. D., Buntrock, C., Mortier, P., Auerbach, R., Weisel, K. K., Kessler, R. C., ... & Bruffaerts, R. (2019). Prediction of major depressive disorder onset in college students. *Depression and Anxiety*, *36*(4), 294-304. https://doi.org/10.1002/da.22867

Essangri, H., Sabir, M., Benkabbou, A., Majbar, M. A., Amrani, L., Ghannam, A., ... & Souadka, A. (2021). Predictive factors for impaired mental health among medical students during the early stage of the COVID-19 pandemic in Morocco. *The American Journal of Tropical Medicine and Hygiene*, *104*(1), 95. https://doi.org/10.4269/ajtmh.20-1302

Franke, F., & Hiebl, M. R. (2022). Big data and decision quality: the role of management accountants' data analytics skills. *International Journal of Accounting and Information Management*, *31*(1), 93-127. https://doi.org/10.1108/IJAIM-12-2021-0246

Gao, L., Xie, Y., Jia, C., & Wang, W. (2020). Prevalence of depression among Chinese university students: A systematic review and meta-analysis. *Scientific Reports*, *10*(1), 15897. https://doi.org/10.1038/s41598-020-72998-1

Grasdalsmoen, M., Eriksen, H. R., Lønning, K. J., & Sivertsen, B. (2020). Physical exercise, mental health problems and suicide attempts in university students. *BMC Psychiatry*, *20*(1), 1-11. https://doi.org/10.1186/s12888-020-02583-3

Grøtan, K., Sund, E. R., & Bjerkeset, O. (2019). Mental health, academic self-efficacy and study progress among college students-the SHoT study, Norway. *Frontiers in Psychology*, *10*, 45. https://doi.org/10.3389/fpsyg.2019.00045

Lattie, E. G., Nicholas, J., Knapp, A. A., Skerl, J. J., Kaiser, S. M., & Mohr, D. C. (2020). Opportunities for and tensions surrounding the use of technology-enabled mental health services in community mental health care. *Administration and Policy in Mental Health and Mental Health Services Research*, *47*, 138-149. https://doi.org/10.1007/s10488-019-00979-2

Lederer, A. M., & Hoban, M. T. (2022). The development of the American College Health Association-National College Health Assessment III: An improved tool to assess and enhance the health and well-being of college students. *Journal of American College Health*, *70*(6), 1606-1610. https://doi.org/10.1080/07448481.2020.1834401

Madyatmadja, E. D., Sembiring, D. J. M., Angin, S. M. B. P., Ferdy, D., & Andry, J. F. (2021). Big data in educational institutions using RapidMiner to predict learning effectiveness. *Journal of Computer Science*, *17*(4), 403-413. https://doi.org/10.3844/jcssp.2021.403.413

Marzukhi, S., Awang, N., Alsagoff, S. N., & Mohamed, H. (2021, August). RapidMiner and machine learning techniques for classifying aircraft data. In *Journal of Physics: Conference Series* (Vol. *1997*, No. 1, pp. 012012). IOP Publishing. https://doi.org/10.1088/1742-6596/1997/1/012012

Niemiec, R. M. (2023). Mental health and character strengths: The dual role of boosting well-being and reducing suffering. *Mental Health and Social Inclusion*. https://doi.org/10.1108/mhsi-01-2023-0012

Trivedi, S. K., & Panigrahi, P. K. (2018). Spam classification: A comparative analysis of different boosted decision tree approaches. *Journal of Systems and Information Technology*, *20*(3), 298-105. https://doi.org/10.1108/JSIT-11-2017-0105

Uska, M. Z., Wirasasmita, R. H., Usuluddin, U., & Arianti, B. D. D. (2020). Evaluation of rapidminer-aplication in data mining learning using persiva model. *Edumatic: Jurnal Pendidikan Informatika*, *4*(2), 164-171. https://doi.org/10.29408/edumatic.v4i2.2688

Zhang, Y., Wu, X., Tao, S., Li, S., Ma, L., Yu, Y., ... & Tao, F. (2021). Associations between screen time, physical activity and depressive symptoms during the 2019 coronavirus disease (COVID-19) outbreak among Chinese college students. *Environmental Health and Preventive Medicine*, *26*, 1-12. https://doi.org/10.1186/s12199-021-01025-0