Article │ Received 26 July 2023; Accepted 2 August 2023; Published 11 December 2023
https://doi.org/10.55092/pcs2023020018

Classification of mental illness risk among employees in workplace using machine learning

Anis Nadiah Rahmat*, Nor Erne Nazira Bazin, Zuriahati Mohd Yunos, Razana Alwee, Chan Weng Howe and Roliana Ibrahim

Faculty of Computing, Universiti Teknologi Malaysia, 81310 Skudai, Johor Bahru, Johor, Malaysia

* Correspondence author; E-mail: anisnadiah@graduate.utm.my.

Abstract: This paper addresses the growing concern of workplace mental illness, impacting productivity and organizational well-being. Employing machine learning, a classification method is developed to assess the likelihood of mental illness among employees. The study encompasses comprehensive mental health analysis and risk factor identification. Data retrieval and preprocessing yield crucial mental health insights. Machine learning methods including Logistic Regression, Decision Tree, and Random Forest are utilised to train a classification model. Recursive Feature Elimination (RFE) enhances model performance by selecting impactful features. Among models, Random Forest achieved 84.10% accuracy. RFE is applied, comparing Feature Sets 1, 2, and 3, with Feature Set 3 exhibited the highest accuracy at 84.19%. This highlights the potency of Feature Set 3 in enhancing the accuracy of Random Forest by 0.09%. The results highlight the effectiveness of the classification method in assessing the risk of mental illness. Implementing this approach can enable organizations to address mental health concerns, elevating productivity and well-being proactively. The study concludes by proposing real-world applications and emphasizing the method's potential for workplace mental health improvement.

Keywords: mental illness classification; machine learning method; recursive feature elimination

1. Introduction

This paper addresses the crucial issue of mental health and its impact on employees worldwide. Mental health plays a vital role in emotional and psychological well-being, affecting actions, behaviors, and feelings. Unfortunately, the neglect of mental health can lead to serious conditions, such as mental illness, which affects ability of individuals to perform effectively at work and may result in severe consequences. As we know, every employee is entitled to a safe and healthy work environment. A quality job, a means of support, and a sense of assurance, purpose, and accomplishment all contribute to good mental health [1]. The main objective of this project is to classify the risk of mental illness among employees using machine learning models, specifically Logistic Regression, Random Forest, and Decision Tree. To enhance the classification performance, the Recursive Feature Elimination (RFE) approach is implemented as a feature selection method. The study also involves the development of a dashboard to visualize the mental health risk among employees in the IT/Technology field. The significance of this project lies in its potential to aid mental health departments in identifying patients with a higher risk of severe mental illness more efficiently and helping employers recognize mental health risk of their employees in the workplace. By providing early intervention and support, this classification can contribute to better mental health outcomes and create a safer and healthier work environment for employees.

2. Mental health risk factors

Mental illness risk factors in this paper explores the various aspects that contribute to the development of mental health issues among employees in the workplace. Mental health plays a vital role in self-awareness, daily problem-solving, and overall well-being, impacting ability of someone to function effectively in a community. However, if not addressed properly, mental health concerns can progress to mental illness. Several specific risk factors increase the likelihood of employees developing mental disorders over time, and these factors are crucial for understanding and addressing mental health issues in the workplace.

Among the risk factors, pre-existing mental conditions are significant contributors to mental illness risk. Family history and background play a role, as individuals with a family history of mental illness have a higher risk of experiencing similar conditions [2]. Diagnosed mental conditions also increase vulnerability, as employees may face burnout, decreased job satisfaction, and social isolation due to mental health challenges [3]. Demographic factors such as age, gender, and country of residence also play a role [4]. It also shown that there are substantial variations between female and male patients with mental illness who are at higher risk of developing a mental illness [5]. Young adults are statistically more likely to experience mental illness, and gender differences exist in the types of mental illnesses individuals are prone to [6]. Additionally, countries with high rates of mental illness may have a lower emphasis on physical activity and overall mental health awareness [7].

Employment-related risk factors include high working hours, particularly for employees under employers, leading to burnout and dissatisfaction [8]. The technology and IT industry, known for high-pressure work environments, shows a high prevalence of mental health issues among employees. Furthermore, the lack of a robust mental health support system in the workplace, coupled with inadequate communication and low social support, can contribute to the development and exacerbation of mental health issues among employees [9]. In order to create a healthier work environment, it is crucial for employers to understand and address these risk factors. This can be achieved by implementing appropriate mental health support systems, promoting open communication, and offering social support to employees [10]. Early detection and prevention measures can help identify employees at higher risk and provide timely interventions to support their mental well-being. By addressing the various risk factors, workplaces can foster a positive and supportive environment that prioritizes mental health, leading to improved well-being and productivity among employees.

3. Machine learning on mental health risk

The application of machine learning techniques are explored to identify and classify the risk factors of mental illness among employees. The three machine learning algorithms used for this classification task are Decision Trees, Logistic Regression, and Random Forest. Supervised machine learning is employed for classifying the workplace mental health risk among employees, as it utilizes labeled training data to uncover patterns and connections. The classification models, such as Logistic Regression, Decision Tree, and Random Forest, make predictions and assign labels to new instances based on the patterns learned from the labeled data. Logistic Regression is primarily used to predict the likelihood of a binary event occurring, such as "Yes" or "No" [11]. It can also be applied to datasets with more than two sets of variables. This algorithm is effective for classifying mental illness risk among employees and offers a straightforward probabilistic viewpoint on class predictions. The formula for the Logistic Regression hypothesis is used to calculate the probability that the target variable y would be 1, given the input features x. Random Forest is an ensemble machine learning framework that combines multiple decision trees to produce precise predictions. It handles datasets with both continuous and categorical variables, making it suitable for classification tasks [12]. The Random Forest model is well-suited for identifying complex interactions between features and can spot correlations and non-linear patterns that simpler models might miss. Decision Trees are supervised learning algorithms used in regression and classification problems. They represent a tree structure, with nodes representing attribute-based questions, edges for the responses, and leaves for the final class label [13]. Decision Trees are effective for understanding and displaying nonlinear data patterns, and they are efficient for analyzing exploratory data. They offer the advantage of ignoring or excluding unimportant parts of data and allow the selection of branches for more precise classification. In conclusion, the application of machine learning algorithms such as Decision Trees, Logistic Regression, and Random Forest holds great potential in identifying and classifying the risk factors of mental illness among employees. These techniques can uncover patterns and relationships within the dataset, leading to effective prediction and classification of mental health risks. By understanding the features and equations of each algorithm, researchers can develop accurate models to assess mental illness risk and support the well-being of employees in the workplace.

4. Research methodology

This paper is described in this study to accomplish its goals and objectives. The research technique and process flow are presented, aiming to select the most appropriate machine learning algorithms for classifying the risk of mental illness. The chosen methodologies depend on the type of study conducted and the classification tasks required. The research framework, illustrated in Figure 1, serves as a roadmap for the successful completion of the entire process and achieving purpose and objectives of the study. It begins with problem formulation, defining the problem description, project aims, objectives, and scopes. The subsequent stage involves conducting a literature review to gain insights into the risk factors affecting mental illness among employees. The literature review utilizes keywords from various sources, such as postgraduate theses, IEEE Xplore, Google Scholar, and others, to understand and investigate the problem comprehensively. Data preparation follows, which includes retrieving data from open-source dataset websites like Kaggle and performing data cleansing, reduction, and transformation. Data visualization is then carried out through exploratory data analysis and dashboard design and development. This phase focuses on graphical representations to present patterns, trends, and insights for easier comprehension and interpretation by the audience. The model training phase utilizes three machine learning algorithms: Decision Trees, Random Forest, and Logistic Regression. The data is split into target and feature variables, and encoding is applied to prepare the data for training. The three classifiers are then trained on the cleaned dataset.

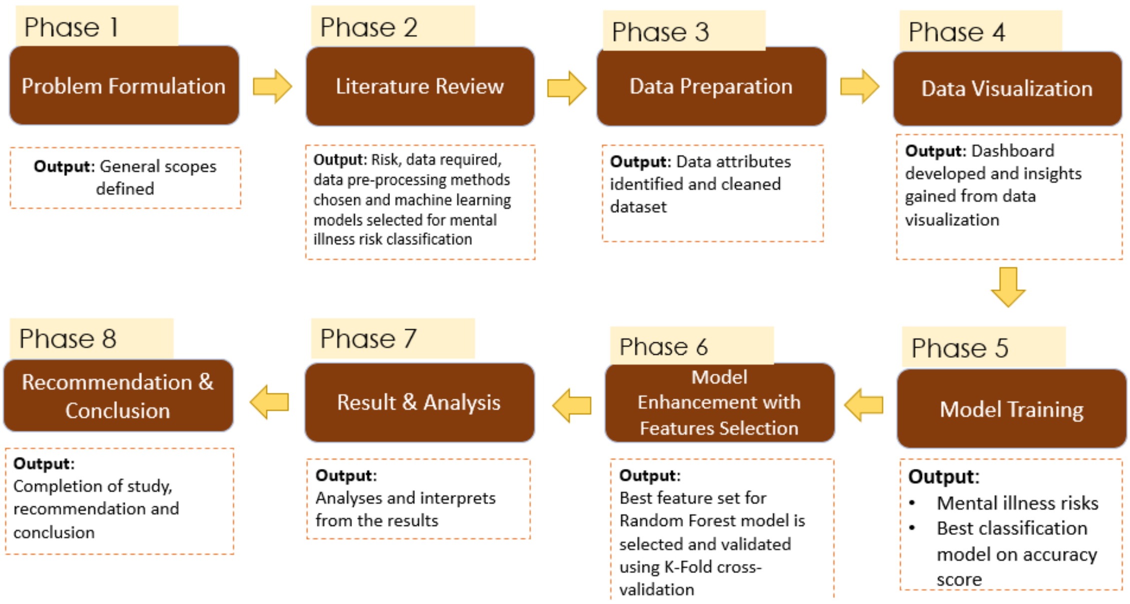

Figure 1. Framework diagram of research methodology.

The model evaluation is performed by displaying the classification report, including accuracy, f1-score, precision, and recall scores. The classifier with the highest accuracy is selected for further research. In the model enhancement with Feature Selection phase, the model performance is improved using Recursive Feature Elimination. This step aims to identify the best combination of features that yield the highest accuracy and the smallest error value, validated by K-Fold cross-validation. The comparisons between methodologies, set features, and overall project results are analyzed during the result and analysis phase. The results of the project, including the performance of the applied machine learning models and the enhancement achieved through Feature Selection, are summarized. Finally, the recommendation and conclusion phase conclude the research, providing recommendations for future works in the field of mental illness among employees. The Gantt chart provided offers a visual representation of the project planning timeline. Overall, this research methodology allows for a comprehensive investigation of the mental health risk among employees using machine learning algorithms and data analysis techniques.

5. Model development

Model development focuses on presenting the initial findings of the project, including exploratory data analysis and data pre-processing. It discusses the results obtained from these initial findings and their relevance in evaluating the data. This phase also highlights the importance of descriptive analysis and how it aids in understanding the data.

The data preparation process involved two main tasks: data retrieval and data pre-processing. The dataset was retrieved from Kaggle and underwent data cleaning, reduction, and transformation using Python scripts. This resulted in an optimized dataset ready for analysis. Data cleaning ensured data correctness and integrity by handling missing values, outliers, and irregularities. Data reduction techniques, such as feature selection, were used to simplify the dataset for analysis, highlighting the most relevant characteristics. Data transformation involved subjecting the data to mathematical or statistical modifications, ensuring all features were on the same scale and ready for analysis. These steps improved the accuracy and reliability of our conclusions. The dimensionality of the dataset was reduced using feature selection methods, which improved computational efficiency and helped avoid overfitting.

The data retrieval process involved collecting data from a survey on employee mental health in IT/Technology organizations from Kaggle. However, the raw data required cleaning and reduction due to missing data and unrelated information in some columns. The data pre-processing steps, performed using Python, included data cleaning, reduction, and transformation to ensure the data was suitable for machine learning algorithms.

Data reduction was done to focus on the most significant columns for the project, resulting in a dataset with 29 features. Additionally, a new ‘year’ column was created to distinguish data from different years, allowing the merging of the datasets. After that, the column ‘year’ is being dropped as it may lead to bias results because majority of the data are from the year of 2016. Next, in data transformation, as shown below in Figure 2, it shows the 29 features after undergoing data cleaning steps which then saved into a .csv file for further use on the next steps for this study.

Descriptive analysis provided a detailed insight into the dataset. Information about data types and missing values was examined, and charts and graphs were used for data visualization. Notably, age and gender distributions, countries with the highest number of respondents diagnosed with mental illness, and trends related to self-employment and mental illness were analyzed. Predictive analysis was conducted using various machine learning models to classify mental health conditions. The selected model was enhanced through feature selection using Recursive Feature Elimination (RFE). Cross-validation was employed to assess the performance of the model. Finally, a dashboard was designed using Microsoft Power B.I., featuring different pages for an overview, age, gender, countries, employment, mental condition, mental health support systems, and conclusions. The dashboard allowed audiences to explore and gain insights into workplace mental health based on the dataset. The exploratory data analysis, data pre-processing, predictive analysis, and visualizations offer valuable insights into employee mental health in IT/Technology organizations, aiding in decision-making and further research.

```
RangeIndex: 2952 entries, 0 to 2951
Data columns (total 29 columns):
 #   Column                            Non-Null Count  Dtype
---  ------                            --------------  -----
 0   self_employed                     2952 non-null   object
 1   total_employees                   2952 non-null   object
 2   tech_company                      2952 non-null   object
 3   mental_health_benefit             2952 non-null   object
 4   mental_health_available           2952 non-null   object
 5   mental_health_discussion          2952 non-null   object
 6   mental_health_resource            2952 non-null   object
 7   anonimity_protected               2952 non-null   object
 8   mental_medical_leave              2952 non-null   object
 9   comfortable_discuss_coworker      2952 non-null   object
 10  comfortable_discuss_supervisor    2952 non-null   object
 11  have_previous_employer            2952 non-null   object
 12  discuss_mental_coworker           2952 non-null   object
 13  discuss_mental_surpervisor        2952 non-null   object
 14  bring_physical_employer_interview 2952 non-null   object
 15  bring_mental_employer_interview   2952 non-null   object
 16  share_friends_family_mental       2952 non-null   object
 17  negative_individual_mental_workplace 2952 non-null   object
 18  history_mental                    2952 non-null   object
 19  diagnosed_mental                  2952 non-null   object
 20  find_treatment_health             2952 non-null   object
 21  mental_treatment_effective        2952 non-null   object
 22  mental_treatment_not_effective    2952 non-null   object
 23  age                               2952 non-null   int64
 24  gender                            2952 non-null   object
 25  country                           2952 non-null   object
 26  state                             2952 non-null   object
 27  work_country                      2952 non-null   object
 28  US_state                          2952 non-null   object
dtypes: int64(1), object(28)
memory usage: 668.9+ KB
```

Figure 2. Features after data cleaning.

In conclusion, this phase presented the research methodology, data preparation, model classification, and model enhancement results. Other than that, the recommended feature set for better prediction of mental illness risk among employees also highlighted in this study.

6. Results

The results of this study are presented focuses on the classification of mental illness risk among employees using machine learning techniques. After model training, feature selection was performed to enhance model performance, interpretability, and explainability. By selecting a subset of the most relevant features, this study gained better insights into the behavior of the model and its relationship with the target variable.

Following the data preparation and feature selection, then proceeded with the model classification using three classifiers: Logistic Regression, Decision Tree, and Random Forest. Based on the accuracy results, the Random Forest classifier outperformed the other two, achieving an accuracy of 84.10% as shown in Table 1.

Consequently, Random Forest are selected for further enhancement with feature selection since it performs the highest accuracy. In the next step of model enhancement, this study then employed the Recursive Feature Elimination (RFE) approach on feature selection. Each classifier (Logistic Regression, Decision Tree, and Random Forest) selected its own set of features, resulting in 14 different combinations in total. After model training with the selected features, its performance metrics are being evaluated. Random Forest with Feature Set 3 showed the highest accuracy of 84.19%, followed by Logistic Regression and Decision Tree as shown in Table 2. Additionally, K-Fold cross-validation was used to calculate the smallest mean error and highest average accuracy scores. Feature set of Decision Tree achieved the smallest mean error of 14.91% as displayed in Table 3, while feature set of Random Forest obtained the highest average accuracy of 85.23% as shown in Table 4. Based on the overall results and performance evaluation, based on the results the Feature Set 3 showed in Table 5 as the recommended as it has the best set of features for enhancing model performance. This feature set included 14 variables that contributed significantly to predicting mental illness risk among employees.

Table 1. Machine learning model accuracy.

ML model	Accuracy (%)
Logistic regression	83.97
Decision tree	83.63
Random forest	84.10

Table 2. Machine learning set features accuracy.

ML model	Accuracy (%)
Feature set 1	83.52
Feature set 2	75.73
Feature set 3	84.19

Table 3. Smallest mean error score by feature set (K-Fold cross validation).

Feature set	SME (%)
Feature set 1	15.58
Feature set 2	14.91
Feature set 3	15.01

Table 4. Average accuracy score by feature set (K-Fold cross validation).

Feature set	Accuracy (%)
Feature set 1	84.18
Feature set 2	84.89
Feature set 3	85.23

Table 5. Feature set 3.

No.	Feature variables
1.	total_*employees*
2.	*mental_medical_leave*
3.	*discuss_mental_supervisor*
4.	*share_friends_family_mental*
5.	*negative_individual_mental_workplace*
6.	*history_mental*
7.	*find_treatment_health*
8.	*mental_treatment_effective*
9.	*mental_treatment_not_effective*
10.	*age*
11.	*country*
12.	*state*
13.	*work_country*
14.	*US_state*

Based on the results compared on the accuracy score of using model classifiers of Logistic Regression, Decision Tree and Random Forest with the accuracy score for Random Forest with feature selection, it is identified that the performance of Random Forest accuracy is increased by 0.09% by using feature selection, which is by using Recursive Feature Elimination (RFE) method.

Random Forest outperformed other machine learning models like Logistic Regression and Decision Tree in this study for classifying mental illness risk among employees. The ensemble nature of Random Forest, where multiple decision trees are combined, reduced the risk of overfitting and improved overall accuracy. Additionally, the ability of Random Forest to handle non-linear relationships between features and the target variable allowed it to capture complex patterns in the data effectively. The feature importance analysis of model helped identify the most relevant features for classification, enhancing its performance. Moreover, robustness of Random Forest to noisy data and ability to handle missing values and outliers saved time and effort. The ensemble approach and feature selection capabilities of Random Forest contributed to its superior performance, making it a suitable choice for this classification task. However, the selection of the best model should consider the specific characteristics and research objectives of the dataset.

7. Conclusion

In conclusion, this study successfully classified mental illness risk among employees using various machine learning techniques. The implemented models, including Logistic Regression, Random Forest, and Decision Tree, demonstrated promising results, offering an efficient and cost-effective tool for future company use. The feature selection approach, RFE, further enhanced the classification performance, with Random Forest and its selected features showing the best accuracy. However, data limitations and the presence of unidentified values may have influenced the performance of the model. For future work, attentive feature engineering and incorporating data from multiple sources could further increase classification capabilities of the study. The contributions of this research lie in its implementation of machine learning in mental health illness risk classification, encouraging employers to prioritize employee well-being. Overall, this study plays a significant role in advancing mental health support and understanding within work environments.

Acknowledgments

The authors would like to acknowledge the financial support from the Ministry of Education Malaysia under Fundamental Research Grant Scheme (FRGS) (FRGS/1/2021/ICT06/UTM/02/3) and Universiti Teknologi Malaysia for the funding under UTM Fundamental Research (UTMFR) (R.J130000.7851.5F479).

Conflicts of interests

The authors declare no conflict of interests.

References

[1] Stoewen DL. Wellness at work: Building healthy workplaces. *Can. Vet. J.* 2016, 57(11):1188–1190.
[2] Lee AA, Laurent SM, Wykes TL, Kitchen Andren KA, Bourassa KA, *et al*. Genetic attributions and mental illness diagnosis: effects on perceptions of danger, social distance, and real helping decisions. *Soc. Psychiatry Psychiatr. Epidemiol*. 2014, 49(5):781–789.
[3] Merikangas KR, He JP, Burstein M, Swanson SA, Avenevoli S, *et al.* Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *J. Am. Acad. Child Adolesc. Psychiatry* 2010, 49(10):980–989.
[4] Suanrueang P, Peltzer K, Suen MW, Lin HF, Er TK. Trends and Gender Differences in Mental Disorders in Hospitalized Patients in Thailand. *Inquiry* 2022, 59:469580221092827.
[5] Waghorn G, Chant D. Overworking among people with psychiatric disorders: results from a large community survey. *J. Occup. Rehabil.* 2012, 22(2):252–261.
[6] Oudejans SCC, Spits ME, van Weeghel J. A cross-sectional survey of stigma towards people with a mental illness in the general public. The role of employment, domestic noise disturbance and age. *Soc/ Psychiatry Psychiatr/ Epidemiol.* 2021, 56(9:1547–1554.
[7] Moll S, Zanhour M, Patten SB, Stuart H, MacDermid J. Evaluating Mental Health Literacy in the Workplace: Development and Psychometric Properties of a Vignette-Based Tool. *J. Occup. Rehabil.* 2017, 27(4):601–611.
[8] Rose DM, Seidler A, Nübling M, Latza U, Brähler E, *et al.* Associations of fatigue to work-related stress, mental and physical health in an employed community sample. *BMC Psychiatry* 2017, 17(1):167.
[9] Hassan MF, Hassan NM, Kassim ES, Hamzah MI. Issues and Challenges of Mental Health in Malaysia. *Int. J. Acad. Res. Bus. Soc. Sci.* 2018, 8(12):1685–1696.
[10] Smith RA, Applegate A. Mental Health Stigma and Communication and Their Intersections with Education. *Commun. Educ.* 2018, 67(3):382–393.

[11] Zhang S, Tjortjis C, Zeng X, Qiao H, Buchan I, *et al*. Comparing data mining methods with logistic regression in childhood obesity prediction. *Inf. Sys. Front.* 2009, 11(4):449–460.
[12] Sujal BH, Neelima K, Deepanjali C, Bhuvanashree P, Duraipandian K, *et al.* Mental Health Analysis of Employees using Machine Learning Techniques. *14th International Conference on COMmunication Systems & NETworkS (COMSNETS)*. 2022, pp. 1–6.
[13] Battista K, Patte KA, Diao L, Dubin JA, Leatherdale ST. Using Decision Trees to Examine Environmental and Behavioural Factors Associated with Youth Anxiety, Depression, and Flourishing. *Int. J. Environ. Res. Public Health* 2022, 19(17):10873.