

Machine Learning in the Fight Against Burnout:

A Study of Young Adults

De Rizzo, Matteo (2749303)

m.de.rizzo@student.vu.nl

Vrij Universiteit Amsterdam

Teacher: Iryna Frankova

Abstract

Generally, burnout is detected with the use of the Maslach Burnout Inventory. However, different approaches to its detection are available, especially with the improvements in the field of machine learning (ML). Three papers are considered, where ML techniques are used to predict levels of burnout in young adults (18-25). The results show impressive performance in the prediction of burnout by part of Random Tree, SVM, and Neural Networks, suggesting the validity and effectiveness of machine learning to predict burnout in its early stages, enabling a better chance of its prevention.

Introduction

Burnout is a serious problem in the modern workplace, affecting the well-being of employees and the productivity of organizations. It is characterized by chronic emotional exhaustion, depersonalization, and reduced personal accomplishment, which can lead to a range of negative outcomes such as absenteeism, reduced job performance, and even mental health issues (Maslach et al., 2001). The prevention of burnout is therefore a critical issue for employers. Similarly, burnout is highly present in students during their university years, due to the learning conditions that demand excessively high levels of effort often without providing support mechanisms that would facilitate an effective coping (Bellido-Medina et al., 2022).

In recent years, machine learning (ML) has emerged as a powerful tool for addressing a wide range of problems in various domains, including healthcare, finance, transportation and bioinformatics (Chauhan et al., 2022). Machine learning is a branch of artificial intelligence that involves the use of algorithms and statistical models to analyze and make predictions based on large amounts of data. The ability of machine learning to identify patterns and make predictions based on complex data sets has led to the development of innovative approaches to the detection of burnout.

The use of machine learning in the detection of burnout raises several important research questions. Among the most fundamental ones is whether machine learning is an effective method for detecting professional burnout at all, which is the focus of the current paper. Research Question:

"Is machine learning an effective method to detect burnout in young adults?"

The research question aims to investigate the capability of machine learning techniques to detect burnout in the population of young adults (between the ages of 18-25), facing high stress due to university and first years of work. It will explore three recent papers on this topic and investigate the potential of machine learning to improve the burnout prevention process.

Currently, the most popular method to detect burnout in an individual is the use of the MBI (Maslach Burnout Inventory). The MBI consists of 22 items that assess three dimensions of burnout: emotional exhaustion, depersonalization, and reduced personal accomplishment. Each item is rated on a 7-point Likert scale, with higher scores indicating greater levels of burnout (Koeske et al., 1989). The MBI is widely used in research and clinical settings and has been translated into multiple languages for use in international research. It is considered as a reliable and valid measure of burnout (Galanakis et al., 2009; Lim et al., 2019).

Alternatively, in ML there are a multitude of models able to learn from grand amounts of data and make independent predictions based on it. Most models differ greatly, each having advantages and disadvantages, depending on the type of data, sample size, and task at hand.

The papers that will be taken into account used many different models to investigate the effectiveness of ML in predicting Burnout, however, for the purpose of this study only three of these models will be considered: Random Tree, SVM and Neural Networks (NN).

A random tree is a decision tree that is built by randomly selecting a subset of the data to use as training data. It is a type of ensemble learning, which means it combines multiple decision trees to create a more accurate model. Random tree models are often used for classification and regression tasks and are particularly useful for tasks where the relationships between the input variables and the output variable are complex (Gall, 2005).

SVM, on the other hand, is a type of supervised learning algorithm used for classification and regression tasks. It works by finding the best boundary, called a hyperplane, that separates different classes of data in a high-dimensional space. SVM models are particularly useful for tasks where the data is not linearly separable, and they are also useful in tasks where the number of features is high and the number of samples is low (Boswell, 2002).

Neural networks (NN), are a type of machine learning model that is inspired by the structure and function of the human brain. They consist of layers of interconnected nodes, called neurons, which work together to process information. Neural networks are particularly useful for more complex tasks (Kaur, 2016).

Paper search process

Gathering enough evidence to answer this research question came with a fair number of difficulties. The main issue that was encountered was the lack of research done in regard to burnout and machine learning combined. While the originally planned research question was more specific, only looking at individuals in the workplace, the inability to find enough literature to back it up caused the RQ to include students in the population, therefore extending it from workers to young adults. This allowed for slightly more papers to be taken as potential evidence to support the research question, finally deciding on the following three.

Burnout in company employees

The first paper (Zhernova et al., 2020) focused on the development of approaches for the early identification of individuals with burnout symptoms. The study was conducted using data collected from 105 employees from a company, gathered according to the methodology provided by the MBI, but with a reduced set of questions. However, while creating the reduced version of the MBI,

the researchers considered the ratio of the original total possible number of points. The data collected by the reduced MBIs were used to construct two models using Random Forest and another algorithm. The study found that the best results were achieved using the Random Forest algorithm, where the precision and recall metrics were 0.8. Precision and recall are two metrics used to evaluate the performance of a binary classification model. Precision is a measure of the accuracy of the classifier when it predicts positive and recall is the ability of the classifier to find all the positive instances. Generally, a good classifier should have high precision and recall (ranging from 0 to 1). This suggests that the proposed models have a high level of accuracy in identifying burnout early. Overall, the conclusion is that the proposed models can be used for early detection of professional burnout, which can help prevent negative consequences and improve the well-being of employees.

Burnout in high risk individuals

The second paper (Almayyan, 2021) studied the psychological stress and job burnout among a very high risk population: front-line personnel, specifically medical staff and police officers, during the COVID-19 pandemic in Kuwait. A dataset from 599 people was collected using a version of the MBI questionnaire, then used to train a wide variety of ML models, after different feature selection techniques. The most valuable results were provided by SVM and a particular type of feed-forward neural network called “deep multilayer perceptron” (DMLP). The SVM Classifier accuracy results range from 0.925 to 0.963 among feature selection techniques while the accuracy results of DMLP range from 0.927 to 0.957. The results also show that DMLP provides better results than other models when the dataset is small. Overall, this study shares the same conclusion as the previous paper, highlighting the effectiveness of machine learning approaches in the field of prevention for burnout.

Burnout in students

Lastly, the third paper (Gonzales et al., 2021) investigated the use of machine learning techniques among which NNs, SVM, and Random Forest to detect burnout risk in university students. It used a sample of 791 students across 4 different universities. This study uses data acquired from psychological evaluation and virtual resources, and it aims to generate preventive actions to address the complications of burnout, such as school dropout and low academic performance, through the analysis of information and the creation of risk projection algorithms. The algorithm with the highest predictive result is Random Forest with a 95% projection percentage. This study could allow for the projection of student burnout at the university level and generate strategies and actions for prevention and intervention, such as tutoring programs that can contribute to reducing dropouts and lag in higher education.

Connection and added value

All three studies are connected by their focus on burnout and their use of machine learning techniques to identify and prevent it. Each study used similar models to learn from the data and provide evidence for the effectiveness of machine-learning approaches in identifying and preventing burnout. Furthermore, the three papers complement each other's value by demonstrating the effectiveness of ML in this field in different populations, ranging from students to high risk front-line personnel.

Discussion

Three papers have been investigated to determine whether Machine Learning methods can be used to effectively predict and detect the appearing of burnout in young adults facing stressful tasks. All studies used a variety of approaches to train their models, and overall Random Forest, SVM

and Neural networks demonstrated to be able to predict cases of burnout with very high accuracy in identifying burnout early. The studies also suggest that the combination of experience of professionals in psychology, education and ML engineering can be beneficial in targeting and treating people on risk of burnout, before it evolves into full-blown burnout, more difficult to treat.

In the field of prevention, machine learning is a relatively new approach, but it has been shown to be effective in the early detection of burnout in different populations. The use of machine learning techniques can provide a more efficient and accurate way of identifying individuals at risk of burnout compared to conventional methods, which can aid in the development of preventive actions to address burnout complications and improve well-being. Therefore, it is important that this approach is further studied and expanded in the future. As machine learning is a rapidly evolving field and technology is advancing quickly, it is important to continue researching and developing new techniques and models to improve the early detection of burnout risk. Additionally, it is also important to conduct more studies in different populations, on other mental health problems that could be predicted, to validate the generalization of the results and to ensure the applicability of the models in different settings.

Limitations

As stated in a previous paragraph, the availability of literature in this topic heavily hindered the development of this paper by only providing a handful of research to support the research question. Furthermore, the population size in the first paper is quite small. In order to avoid these obstacles in the future, it is crucial to do more research, to provide a comprehensive set of studies to support further research in the field of prevention of mental disorders with the use of machine learning approaches.

Conclusion

In conclusion, burnout is a growing concern that affects individuals across different fields and industries. This research shows that machine learning techniques can be an effective tool for early detection of burnout risk in young adults that can allow for the development of preventive actions to address burnout complications and improve well-being.

References

- Almayyan, W. (2021). Developing a Machine Learning Model for Detecting Job Burnout During the COVID-19 Pandemic Among Front-line Workers in Kuwait.
- Bellido-Medina, R.S., Evangelista-Aliaga, J.L., Gamarra-Castellanos, M.E., García-Apaza, J.Y., Gonzales-Manrique, E.J., & Calizaya-López, J. (2022). Academic burnout syndrome in university students during the Covid-19 pandemic. *International journal of health sciences*.
- Boswell, D. (2002). Introduction to Support Vector Machines.
- Chauhan, H. (2022). Overview of Machine Learning. *International Journal of Advanced Research in Science, Communication and Technology*.
- Galanakis, M., Moraitou, M., Garivaldis, F.J., & Stalikas, A. (2009). Factorial Structure and Psychometric Properties of the Maslach Burnout Inventory (MBI) in Greek Midwives. *Europe's Journal of Psychology*, 5, 52-70.
- Gall, J.F. (2005). Random trees and applications. *Probability Surveys*, 2, 245-311.
- Kaur, A. (2016). An introduction to neural network. *International journal of applied research*, 2, 164-167.
- Koeske, G.F., & Koeske, R. (1989). Construct Validity of the Maslach Burnout Inventory: A Critical Review and Reconceptualization. *The Journal of Applied Behavioral Science*, 25, 131 - 144.
- Lara-González, L.R., Delgado-Luna, M.A., De León-Galván, B.E., & Venegas-Guerrero, J.C. (2021). Comparison of Machine Learning algorithms for the Burnout projection. *ECORFAN Journal-Democratic Republic of Congo*.

Lim, W.Y., Ong, J., Ong, S., Hao, Y., Abdullah, H.R., Koh, D.L., & Mok, U.S. (2019). The Abbreviated Maslach Burnout Inventory Can Overestimate Burnout: A Study of Anesthesiology Residents. *Journal of Clinical Medicine*, 9.

Maslach, C., Schaufeli, W.B., & Leiter, M.P. (2001). Job burnout. *Annual review of psychology*, 52, 397-422.

Zhernova, P., Bodyanskiy, Y.V., Yatsenko, B., & Zavgorodnii, I. (2020). Detection and Prevention of Professional Burnout Using Machine Learning Methods. 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), 218-221.