

Machine Learning Predictions for South Korean Higher Education Students Work-Life Balance

Ryan Hatcher

Assistant Professor, Department of General Education, Professor Hannam University, Korea.

*Corresponding Author

Ryan Hatcher

Assistant Professor, Department of General Education, Professor Hannam University, Korea.

Article History

Received: 16.11.2023

Accepted: 28.11.2023

Published: 25.12.2023

Abstract: In today's quickly evolving job and educational scenario, striking a healthy work-life balance is becoming more and more crucial. The purpose of this project is to investigate the variables affecting students' work-life balance and to determine the best machine learning methods for doing so. At Hannam University in South Korea, 231 students participated in a survey that produced a dataset with 59 characteristics. The dataset was cleaned up, and its size was reduced, using data preparation techniques. The top 15 relevant characteristics were determined using the SelectKBest feature selection method. Using performance criteria including precision, recall, F1 score, and accuracy, three classification algorithms—Logistic Regression, Balanced Random Forest Classifier, and Neural Network—were trained and assessed. The findings showed that work-life balance was highly impacted by variables such place of residence, number of majors, time spent in college, and weekend class attendance. The Neural Network performed best among the methods, demonstrating how well it can forecast work-life balance based on the discovered features. In order to create interventions and policies that support a better balance between academic and personal duties, stakeholders in the education and employment sectors would benefit greatly from understanding the implications of these findings.

Keywords: Machine learning, Work-Life Balance, Education, University Students, Data analysis.

Introduction

The idea of work-life balance has been much more popular recently, especially because the global pandemic forced the quick adoption of remote work and e-learning. Students have benefited most from these new paradigms since they have made it possible for people to achieve a better balance between their personal and professional lives. Increased productivity, improved knowledge retention, and a more upbeat view on life are just a few benefits of being able to maintain a healthy work-life and study balance. The work-life balance has a lot of room for improvement as the world embraces the digital revolution, and remote work culture has emerged as a key component in this endeavor [1].

It is impossible to overestimate the significance of work-life balance for both general well-being and mental wellness. Numerous advantages of maintaining a healthy work-life balance have been demonstrated, including lowered levels of stress and anxiety, enhanced mental and physical health, and increased levels of job satisfaction and engagement [2]. Increased productivity, improved knowledge retention, and a more upbeat view of life are all advantages of a balanced study and work-life culture for students. Industries face a variety of problems that call for creative solutions as the speed of technological development quickens. Universities have reacted by embracing digitization and providing online education, which enables students to access resources and submit projects remotely. Because it reduces the stress and anxiety brought on by uneven work-life frameworks, this flexible approach

to education has a direct impact on students' productivity and quality of life [3].

Given the significance of work-life balance, it is imperative to create tools and approaches that help forecast and enhance student work-life balance. The development of intelligent computer programs with the ability to learn on their own is made possible by the possibility for automation and predictive analysis offered by machine learning, a young technology. In this study, a machine learning model is put forth to forecast the work-life balance of students while they are enrolled in classes. This model can help colleges optimize their fee schedules and provide better services to their students. The proposed model would examine a variety of work-life balance-related variables, including age, gender, education, marital status, children, adult dependents, general seniority, years of experience, organization size, and occupation category [4].

This study's organization is as follows: The first step is to present a thorough overview of the study work-life balance, its effects, and the possible advantages of anticipating and enhancing it using machine learning approaches. The need of work-life balance for students is emphasized in this part along with the opportunities and challenges that the digital revolution and the quick speed of technology improvements provide.

After then, a review of previous studies that have looked at pertinent topics is given. In addition to discussing the limits and

gaps in the existing literature, this section also discusses the methods currently used to forecast work-life balance. The most promising directions for future study are found by synthesizing the results from prior investigations, and the development of the machine learning model is guided by this information.

The study experiment is thoroughly discussed in the methodology section, along with the data sources, preprocessing methods, feature selection, and machine learning algorithms used for classification and prediction. The evaluation measures and the justification for their choice are also described in this part along with the performance of the various models. The study's conclusions are presented in the results and analysis section along with the accuracy, precision, recall, and F1 scores of the various classification methods. The performance of the suggested model is also contrasted with that of existing models described in the literature, and the implications of these findings for the forecasting of students' study-work-life balance are examined. The discussion of the research findings and their consequences for students, academic institutions, and politicians comes to a close the report. The creation of more sophisticated machine learning models and the inclusion of more variables that can affect work-life balance are examples of potential directions for future research. This research intends to contribute to a deeper knowledge of the elements that drive work-life balance and to provide useful insights for the creation of tailored treatments and support systems. It does this by researching novel techniques to forecasting study work-life balance.

This research aims to pinpoint the major determinants of a balanced study and work-life culture by looking at the relationship between work-life balance and numerous aspects. The research findings can be used to guide the creation of laws and plans that support kids' academic achievement, personal development, and mental health. Additionally, the outcomes can help universities create curricula and support services that better meet the needs of their students, fostering a more favorable environment for learning and individual growth.

Literature review

It is essential to have a solid understanding of the role that educational institutions play in the process of cultivating an atmosphere of support as the world is fast adjusting to a new normal that places a greater emphasis on a better work-life balance for both employees and students. The writers of this study work [7] examine the ways in which the future generation of faculty members can be prepared to give quality education. This, in turn, encourages the development of well-rounded individuals who can tackle a variety of obstacles in their professional lives. They contend that engaging in conversation with one's contemporaries is an essential element to accomplishing this objective.

A comparative study on work-life balance among college students enrolled in ASEM classes was carried out by [8], which found that maintaining a healthy work-life balance is becoming an increasingly vital component of their everyday existence. Achieving a healthy equilibrium between one's professional and personal lives depends on several factors, including one's capacity for emotional, imaginative, spiritual, intellectual, and physical equilibrium. These are the aspects that need to be taken into consideration by employers since they have the potential to affect employees' levels of productivity and creativity [9]. Since the

beginning of the COVID-19 pandemic in the year 2020, working from home has become increasingly common. Many employees like this arrangement since it enables them to spend more time with their families [10], and this is one of the primary reasons why this trend has grown so widespread.

The field of artificial intelligence has recently seen the emergence of a powerful new tool known as machine learning, which is being progressively put to use to automate the resolution of a wide variety of problems [11]. Data scientists and academics who are active in the topic of work-life balance have been working on the development of machine learning models for the purpose of analyzing employees' work-life balance and making improvements to it. Using techniques from machine learning, the researchers in one such study [12] investigated the impact that emotional intelligence has on productivity as well as job satisfaction. The study evaluates the strengths and weaknesses of employees by assessing their levels of abilities, self-awareness, sensitivity, influence, and motivation. These results are then analyzed using machine learning algorithms. In order to provide a concise summary of the findings, the WEKA tool is utilized.

The epidemic has had a significant impact on the work-life balance of IT personnel, which has prompted many of them to transition working remotely. Even though the information technology business was one of the first to start working again, other fields have also been impacted [1]. During this time, the information technology industry made a smooth transition to a model that allows employees to work from home, which resulted in uninterrupted business continuity and sustained production [10]. Productivity among IT workers can be significantly affected by a variety of factors, including the working environment, the presence of distractions, the level of stress, and the quality of communication [3]. It is interesting to note that working from home has, in some situations, resulted in an improvement in productivity since employees have been able to better balance their work and family lives.

In a similar vein, students who are pursuing higher education have reported increased levels of satisfaction and performance after embracing the connection between maintaining a healthy work-life balance and making use of e-learning capabilities. 1032 students enrolled in higher education programs in a variety of countries were polled as part of an investigation into how they feel about e-learning, which was carried out by [13]. Using machine learning regression models, an evaluation of work-life balance was conducted, and the results showed that e-learning has a favorable influence on living standards and culture [5]. According to the findings of the research, students may be able to reduce costs associated with travelling to and from school, maintaining their health, purchasing food, and finding suitable housing [6]. In addition, students who participate in e-learning are more likely to demonstrate innovative thinking, creative output, and productive behavior, providing their training is adequate and they keep open lines of communication with their teachers.

A rising number of businesses are becoming concerned about employee turnover, which can frequently be traced back to an imbalance between their personal and professional lives [4]. Researchers and professionals working in human resources have been putting a lot of effort into developing machine learning models that can accurately forecast employee turnover in order to

cut employee turnover rates. In a research study titled [14], the authors recommend using machine learning models to predict employee turnover. Some examples of these models include decision tree classifiers, artificial neural networks, and support vector machines. The optimized support vector machine (SVM) showed the best performance, achieving an accuracy of 88.87% [15].

The trend of working from home, particularly among those employed in information technology, is gaining popularity. employment-life balance and overall well-being are topics that are investigated in [16]'s study paper, which focuses on the setting of remote employment. According to the findings of the survey, there are advantages to working remotely for both businesses and employees. As a result, an increasing number of workers, particularly women, are selecting this method of employment. In a separate piece of research [2], researchers analyzed the work-life balance of 12,7756 employees using data acquired from a dataset containing those individuals. Eighty percent of the dataset was analyzed using several machine learning techniques, such as the random forest classifier, SVM, and Naive Bayes. The accuracy of the models was determined to be 71.5%. To achieving a better work-life balance, the data was visualized with the use of bar plots and correlation plots. Factors such as sleep hours and job satisfaction were taken into consideration.

Work engagement is another factor that might affect work-life balance, and researchers have used machine learning methods to analyse data related to the prediction of job involvement in their studies. The authors of the research presented in [11] describe a methodology that is based on machine learning techniques and makes use of the IBM dataset. The purpose of the study is to examine the prediction performance of GLM, which incorporates linear regression and binomial classification, as well as to gain a better understanding of the roles that variables play in job involvement. This model is the first of its kind to predict job involvement using machine learning techniques and employee demographics [17]. It achieves an accuracy of 68% and is the first of its kind.

Learning machines can make improvements to a variety of facets of daily living, which can help employees enhance their quality of life overall and find it easier to strike a healthy balance between their personal and professional lives. The authors of a research paper [18] make use of machine learning technology to determine the aspects that influence a subject's perception of whether or not they have a healthy work-life balance. The survey included participation from a significant number of employees 8,800 in total. The SelectKBest algorithm was put into action, functioning in conjunction with ANN and featuring two hidden layers with a total of 50 and 25 neurons each. The accuracy of the model, which made use of Relu and ADAM, came in at 81%. According to the findings of the study, those who have a satisfactory balance between their job and personal lives have a higher level of motivation [19].

Motivation

The idea of striking a healthy balance between one's professional and personal life has received a lot of attention in recent years, particularly since individuals and organizations have been striving to adapt to the rapidly shifting employment and educational landscape. This change, which includes the increased prevalence of

remote work and e-learning, has been further hastened by the COVID-19 pandemic. Because of this, there is an increasing demand to comprehend and perfect the art of striking a healthy balance between one's professional or academic responsibilities and one's personal life to ensure the health and success of individuals in this new environment. This research is driven by the goal of finding a solution to this important problem. More specifically, the researchers are interested in the factors that influence the work-life balance of students as well as the machine learning algorithms that are most effective at predicting this balance. In the end, the purpose of this research is to contribute to the expanding body of knowledge on the topic of work-life balance by providing useful insights and suggestions for stakeholders in the education and employment sectors to help them cultivate an environment that promotes the health and success of both students and professionals.

Methodology

3.1 Data Collection

The data set was compiled by the administration of a survey to the student body at Hannam University, which is in South Korea. The purpose of the survey was to inquire about a variety of topics pertaining to the academic and personal life of students. The dataset contains a total of 231 cases and 59 attributes. Of these, 175 students reported having a healthy balance between their personal and professional lives, while 56 students did not.

3.2 Data Preprocessing

The initial dataset contained several errors, such as data that were missing, data types that did not match up, and irrelevant characteristics. In order to enhance the performance of the classifier, various data pre-treatment strategies were utilized. The following procedures were carried out during the preprocessing stage:

Handling missing values, eliminating duplicate entries, and altering column names to have shorter lengths are all part of the data-cleaning process.

Feature Reduction is the process of removing columns from a dataset that are deemed to be superfluous.

Data Splitting is the process of separating the dataset into test and training subsets.

3.3 Feature Selection

The process of selecting the characteristics within a dataset that are most pertinent to the problem at hand is known as feature selection. The purpose of this study is to determine which elements are the most significant in determining whether students can maintain a healthy work-life balance. In this investigation, the SelectKBest technique was utilised in order to zero in on the top 15 characteristics of the dataset that are considered to be of the utmost significance. The k features with the highest scores are the ones that are kept by the algorithm and ultimately employed in the machine learning model.

3.4 Training for Models and Their Classification

After the data preprocessing was finished, a technique called 5-fold cross-validation was used to reduce the amount of bias that was linked with the dataset. Using this method, the dataset is split up into k separate folds that are all the same size, with each fold comprising the same number of samples. After that, the

classification model is trained and put to the test k times. Every time, the first k-1 subsets are used for training, and the subset that is left over is used for testing.

This research made use of three different classification algorithms, including Logistic Regression, Balanced Random Forest Classifier, and Neural Network. Logistic Regression was selected as the model to analyze because of its capacity to make accurate predictions regarding binary outcomes by utilizing only early observations of the dataset. The Neural Network uses neurons to learn new information and backpropagation techniques to minimize loss, whereas the Balanced Random Forest Classifier uses an undersampling strategy to balance the dataset. Both methods are employed by the Balanced Random Forest Classifier.

Both the Logistic Regression and the Balanced Random Forest Classifier underwent hyperparameter tuning so that the overall performance of the model may be improved.

3.5 Model Evaluation and Performance Metrics

Following the training of the models, their performance was evaluated using a variety of evaluation measures, including accuracy, precision, recall, and F1-score, among others. These parameters enabled a complete assessment of the efficacy of the algorithms in forecasting the study and work-life balance of students in the higher education system in South Korea.

Results and Discussion

The study examined student work-life balance and the best machine learning methods for predicting it. Survey responses from 231 Hannam University students in South Korea were obtained. After data preprocessing and feature selection, Logistic Regression, Balanced Random Forest Classifier, and Neural Network were trained and assessed using multiple performance measures.

Table 1- Performance of the machine learning algorithms

Model Name	Precision	Recall	F1 Score	Accuracy
Logistic Regression	0.82	0.75	0.78	0.76
Balanced Random Forest	0.68	0.82	0.74	0.78
Neural Network	0.77	0.79	0.78	0.80

The logistic regression model has 0.82 precision, 0.75 recall, 0.78 F1 score, and 0.76 accuracy. Balanced random forest classifier precision was 0.68, recall 0.82, F1 score 0.74, and accuracy 0.78. The neural network model had improved precision, recall, F1 score, and accuracy of 0.77, 0.79, 0.78, and 0.80. These measures show how well models anticipate student work-life balance. Overall the neural network algorithm perform best among all the other algorithm

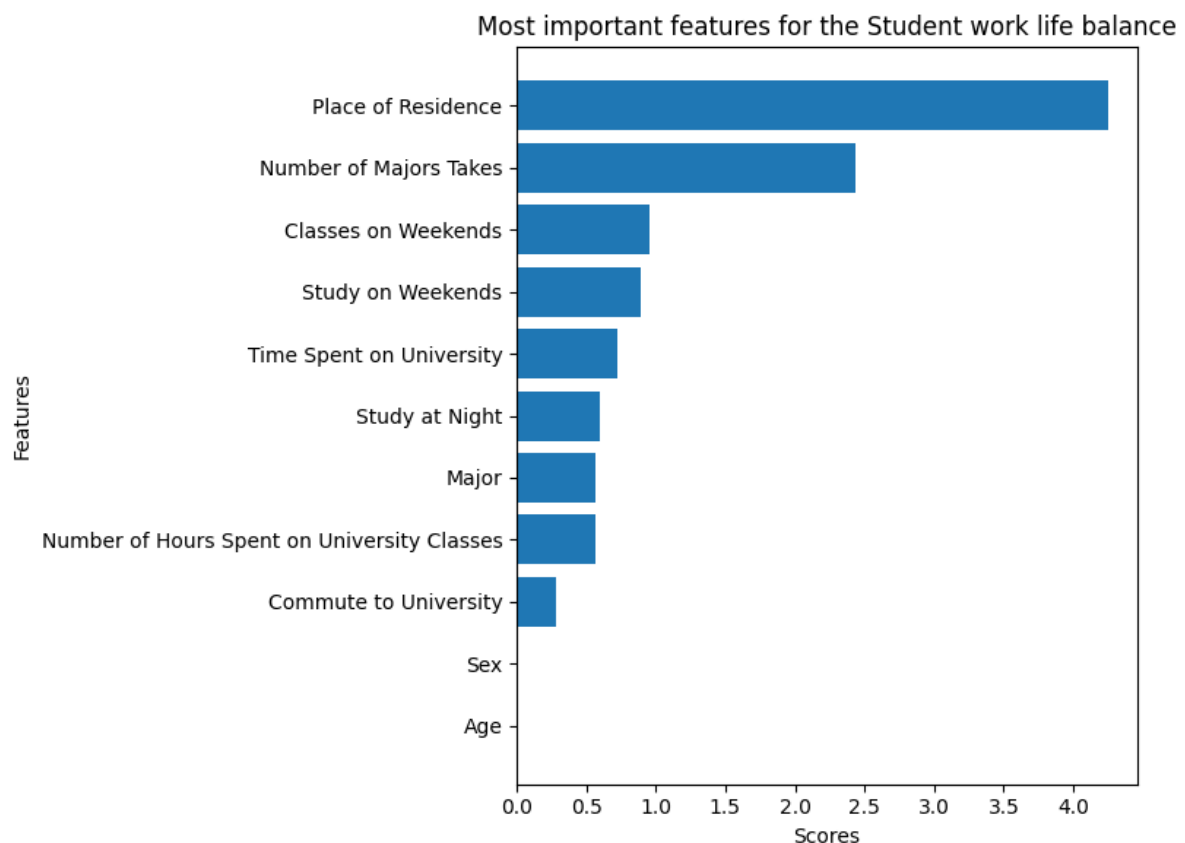


Figure 1- Most important features for the work life balance

Work-life balance was predicted using SelectKBest. The top 15 characteristics were "Place of residence," "Number of Majors takes," "University time," "Weekly university classes," "Weekend classes? Yes/no," "Weekend study? Yes/no," "Study at night?" "Yes/no," "Do you commute to university? Yes/no," "Do you work? Yes/no," "SLB_1," "SLB_2," "SLB_3," "SLB_4," "SLB_5," and "PSD_5". These features reveal important elements affecting students' work-life balance.

Logistic regression, which can accurately predict binary outcomes with minimal observations, did well. The balanced random forest classifier, which under sampled to balance the dataset, had improved recall and F1 score values. The balanced random forest classifier appears to better identify pupils with poor work-life balance. The neural network model, which used neurons and backpropagation, had the best precision, recall, F1 score, and accuracy of the three models.

The findings illuminate higher education students' work-life balance. Stakeholders in education and work can learn how to help students by identifying key criteria. Understanding how elements like majors, university activities, and weekend classes affect academic programmes might help institutions create better balanced programmes. Employers can also help student workers achieve a good work-life balance by considering job engagement and commute.

This study has limitations. The dataset came from a single South Korean institution, which may restrict its applicability to other student demographics or cultures. Self-reported survey replies might be biased and inaccurate. To improve robustness and generalizability, future study might increase sample size and include data from numerous colleges or countries. This study sheds light on student work-life balance and compares machine learning techniques for predicting it. The neural network model beat the logistic regression model and balanced random forest classifier in precision, recall, F1 score, and accuracy. Stakeholders may build interventions and policies that improve students' work-life balance by identifying relevant qualities. Work-life balance research advances and informs education and employment stakeholders.

CONCLUSION

This study investigated student work-life balance and the best machine learning techniques for predicting it. We gained important insights from survey data from 231 Hannam University students in South Korea. The findings showed that several factors affect student work-life balance. "Place of residence," "Number of Majors Taken," "Time Spent on University," and "Classes on Weekends" were the most important factors. These findings help education and employment stakeholders build interventions and policies that promote a healthy balance between academic and personal duties.

Logistic Regression, Balanced Random Forest Classifier, and Neural Network were compared. The Neural Network method has the best precision, recall, F1 score, and accuracy. Based on the detected features, the Neural Network model may predict student work-life balance. This research is important. Universities can improve work-life balance by studying the elements that affect it. Employers can accommodate students' duties and help them manage work and life. This improves student well-being, productivity, and achievement.

This study has limitations. The data came from one South Korean institution, which may restrict its applicability to other student demographics or cultures. Self-reported survey replies may be biased and inaccurate. Future studies might integrate data from many colleges or nations to better understand student work-life balance. Interviews and focus groups can also reveal students' work-life balance opinions.

In conclusion, this research advances student work-life balance understanding and has practical implications for education and employment stakeholders. Institutions may promote students' well-being and achievement by recognising relevant aspects and using machine learning algorithms, resulting in a more harmonious integration of academic and personal life.

References

1. Saif, A., Rehman, T., & Alam, S. (2020). Contours of Emerging Technology on Work Life Balance and Economic Well-Being. *GMJACS*, 10(1), 14-14.
2. Radha, K., & Rohith, M. (2021). An Experimental Analysis of Work-Life Balance Among The Employees using Machine Learning Classifiers. *arXiv preprint arXiv:2105.07837*. doi: 10.14445/22312803/ijctt-v69i4p108.
3. Byrne, U. (2005). Work-life balance: Why are we talking about it at all?. *Business Information Review*, 22(1), 53-59. doi: 10.1177/0266382105052268.
4. Matulevicius, S. A., Kho, K. A., Reisch, J., & Yin, H. (2021). Academic medicine faculty perceptions of work-life balance before and since the COVID-19 pandemic. *JAMA network open*, 4(6), e2113539-e2113539. doi: 10.1001/jamanetworkopen.2021.13539.
5. Lewis, S., & Humbert, A. L. (2010). Discourse or reality? "Work-life balance", flexible working policies and the gendered organization. *Equality, Diversity and Inclusion: An International Journal*, 29(3), 239-254. doi: 10.1108/02610151011028840.
6. Bellmann, L., & Hübler, O. (2021). Working from home, job satisfaction and work-life balance—robust or heterogeneous links?. *International Journal of Manpower*, 42(3), 424-441. doi: 10.1108/IJM-10-2019-0458.
7. Hamberger, L. K. (2007). Preparing the next generation of physicians: medical school and residency-based intimate partner violence curriculum and evaluation. *Trauma, Violence, & Abuse*, 8(2), 214-225. doi: 10.1177/1524838007301163.
8. Yadav, M. (2017). A Comparative Study of Work Life Balance among Higher Education Students in ASEM Courses. *International Journal of Engineering and Management Research (IJEMR)*, 7(3), 37-41.
9. Ryan, H. and Hwang, Y.-S. (2019) 'Work life balance practices and the link to innovation and productivity: A

- comprehensive literature review', *International Convergence Management Association*, 7(1), pp. 26–38. doi: 10.20482/jemm.2019.7.1.26.
10. Natarajan, N. et al. (2021) 'Work From Home Situation and Work Life Balance of It Employees During Covid-19 Pandemic Period-a Case Study Approach', *Article in Journal of Huazhong University of Science and Technology*, (June). Available at: <https://www.researchgate.net/publication/352443243>.
11. Choi, Y., & Choi, J. W. (2021). A study of job involvement prediction using machine learning technique. *International Journal of Organizational Analysis*, 29(3), 788-800. doi: 10.1108/IJOA-05-2020-2222.
12. Grace, P. J., & Banu, N. N. (2015). Machine Learning on Emotional Intelligence and Work Life Balance. *International Journal of Computer Applications*, 116(10). doi: 10.5120/20376-2597.
13. Bucea-Manea-Țoniș, R., Bucea-Manea-Țoniș, R., Simion, V. E., Ilic, D., Braicu, C., & Manea, N. (2020). Sustainability in higher education: The relationship between work-life balance and XR e-learning facilities. *Sustainability*, 12(14), 5872. doi: 10.3390/su12145872.
14. Mansor, N., Sani, N. S., & Aliff, M. (2021). Machine learning for predicting employee attrition. *International Journal of Advanced Computer Science and Applications*, 12(11). doi: 10.14569/IJACSA.2021.0121149.
15. Saradhi, V. V., & Palshikar, G. K. (2011). Employee churn prediction. *Expert Systems with Applications*, 38(3), 1999-2006. doi: 10.1016/j.eswa.2010.07.134.
16. Felstead, A., & Henseke, G. (2017). Assessing the growth of remote working and its consequences for effort, well-being and work-life balance. *New Technology, Work and Employment*, 32(3), 195-212. doi: 10.1111/ntwe.12097.
17. Alao, D. A. B. A., & Adeyemo, A. B. (2013). Analyzing employee attrition using decision tree algorithms. *Computing, Information Systems, Development Informatics and Allied Research Journal*, 4(1), 17-28. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.101.2.2947&rep=rep1&type=pdf>.
18. Pawlicka, A., Pawlicki, M., Tomaszewska, R., Choraś, M., & Gerlach, R. (2020). Innovative machine learning approach and evaluation campaign for predicting the subjective feeling of work-life balance among employees. *PloS one*, 15(5), e0232771. doi: 10.1371/journal.pone.0232771
19. Shanafelt, T. D., Boone, S., Tan, L., Dyrbye, L. N., Sotile, W., Satele, D., ... & Oreskovich, M. R. (2012). Burnout and satisfaction with work-life balance among US physicians relative to the general US population. *Archives of internal medicine*, 172(18), 1377-1385. doi: 10.1001/archinternmed.2012.3199.