Prediction of Mental Health: Heuristic Subjective Well-Being Model on Perceived Stress Scale Using Machine Learning Algorithms

Ahmet Karakuş*, Akif Can Kılıç, and S. Emre Alptekin

***Abstract*—More research is being done to find out how well-being can be predicted using well-designed models. To create a workable Subjective Well-Being (SWB) model, it is vital to look at the backgrounds of characteristics. From the SWB literature, we have chosen variables that are appropriate for real-world data instructions. The objective of this work is to assess the model's performance on a real dataset by giving it SWB determinants and then classifying stress levels using machine learning techniques. Although it is a multiclass classification problem, we have nevertheless managed to obtain meaningful metric scores that can be considered for a particular assignment.**

***Index Terms*—Machine learning, multiclassification, subjective well-being, perceived stress scale**

I. Introduction

According to studies, maintaining one's well-being is crucial for keeping people healthy and effective [1].

Subjective Well-Being (SWB), one type of well-being, is broken down into three defining characteristics [2]. The first characteristic is a person's subjective viewpoint that is gained by their experience [3]; as a result, it is not imposed by any other external sources [4]. Since the goal is not just to look for the negative components, the second feature of SWB is that there are also positive measures. Third, both cognitive and affective well-being components are included in SWB assessments [5]. In more depth, affective well-being refers (AWB) to a person's mood, whereas the cognitive well-being (CWB) involves assessments of one's life as a whole and contentment with certain life domains [6]. As a result, the SWB's structure has not yet been established, although it can be presented as depicted in Fig. 1 below [7].

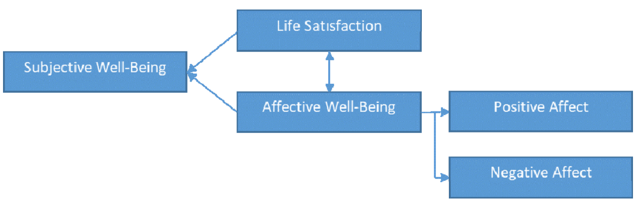

Fig. 1. Structure of subjective well-being (SWB).

The purpose of this study is to test the model by feeding it SWB variables and using machine learning methods to categorize the stress levels. There are several different factors that can influence the degree of SWB [8]. The variables of social support and work stressor may be characteristics that play a significant influence in SWB, according to the evidence [9].

The Fig. 2 illustrates the stress process model and the connection between stress and SWB [10]. The primary and secondary stressors, resources, status, and outcomes [11–13] are among the terms included in the model. According to this theoretical model, there are objective and subjective stressors among the main stressors. Based on the individual's evaluation of the objective stressors, which can be seen in either a good or negative light, the subjective stressors are identified. This factor has a healing effect on both harmful stressors and SWB, in accordance with the terms of resource [13, 14]. Resources can improve a person's SWB and help them manage with stress [15].

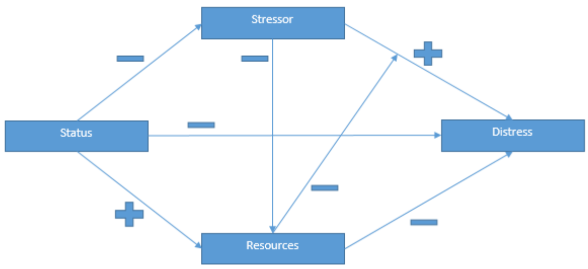

Fig. 2. Process of stress.

The remainder of this work is organized as follows: in Section II, a review of the literature gives background information on the heuristic model that is mentioned in relation to each of the selected SWB determinants, and then a broad outline of the heuristic model is presented. Section III analyzes the Perceived Stress Scale (PSS) as a target, and Section IV summarizes the dataset to explain the preprocessing that was applied to the data. The performance measurements are discussed in Section V, presented and assessed in Section VI, and finally, our findings are drawn in Section VII.

II. Subjective Well-Being (SWB) Determinants

A growing body of research has been done [16, 17] that focuses on employing well-designed models to predict well-being. Research is done on the backgrounds of the variables in order to create a solid SWB model. We have chosen the relevant variables from the Subjective Well-Being

Manuscript received July 5, 2022; revised August 22, 2022; accepted January 11, 2023.

A. Karakuş is with the Galatasaray University, Logistics and Financial Management, 34349, Istanbul, Turkye.

A. C. Kılıç is with İstanbul Kültür University, Department of Industrial Engineering, 34158, Istanbul, Turkye.

S. E. Alptekin is with the Galatasaray University, 34349, Istanbul, Turkye.

*Correspondence: akarakus94@gmail.com (A.K.)

doi: 10.18178/ijml.2023.13.3.1138

(SWB) literature that are suitable for actual data. In other words, we did not include the selected SWB determinant in the heuristic model if there were only a few patterns in the real data. Since the SWB determinants have been heuristically divided into the three groups, certain determinants may belong to different groups.

The three categories that we use to categorize the SWB determinants are: physical well-being, affective well-being, and cognitive well-being. Next, we discussed the rationale for choosing the SWB determinants (features).

A. Physical Well-Being (PWB)

Positive benefits on general health are produced by exercise, sleep, and healthy lifestyle choices [18]. For this reason, the PWB part's options for sleep and physical activity are chosen.

1) Sleep

Lack of sleep slows reaction times, changes mood, and worsens cognitive and perceptual impairment [19]. It is known that sleep deprivation has an influence on both AWB and CWB. When compared to CWB, AWB is more affected by sleep deprivation than CWB [20]. Mood and sleep have a direct relationship [21].

In this study, Pittsburgh Sleep Quality Index (PSQI) was employed [22]. It is a well-known self-report questionnaire that assesses the general effectiveness of sleep and the frequency of interruptions during particular intervals. Sleep Quality, Sleep Latency, Sleep Duration, Sleep Efficiency, Sleep Disturbances, Sleep Drug Use, and Daytime Impairments are some of the subjective aspects of sleep that the PSQI helps to understand [23]. We chose PSQI question 4 (average sleep per night) from the self-reported data because our goal is to create a robust heuristic model and sleep could be a useful SWB variable.

2) Physical Activity (PA)

Data from the NetHealth Project covering the years 2016 and 2019 were gathered for the study by [24]. Participants in the study who had favorable trends in physical activity (PA) show enhanced self-image, self-esteem, and health. Participants with negative PA trends, on the other hand, show a higher risk of anxiety and depression. A different phrase is employed in a different study [25]: quality of life (QOL). Positive and negative life perspectives make up quality of life. In the 1980s, a new term known as health-related quality of life (HRQoL) emerged [25]. The study's findings demonstrate that poor physical activity regularly and independently has a negative influence on QOL and HRQoL, and this relationship is reciprocal. Due to the fact that PA may be a desirable SWB feature and is thus linked to decreased stress and increased wellbeing, we opted using self-reported data to gauge an individual's level of activity throughout the course of the semesters.

B. Cognitive Well-Being (CWB)

Birth, death, retirement, and marriage are just a few examples of life events that are present in CWB and have a stronger impact on it than they do on AWB [26, 27]. As noted in the introduction, SWB is composed of two aspects: 1) AWB, which focuses on an individual's mood and might change daily, and 2) CWB, which focuses on external factors (such as money, employment position, or recent life events) [26]. Therefore, we have heuristically chosen the external factors that are social relations, mother's age, and income level of both their parents and networks.

1) Social relations

SWB is influenced by strong social ties [28–30]. For instance, parents who receive assistance from their social network after catastrophic occurrences adjust better [31]. All age groups can experience loneliness, which is one of the key determinants of social wellbeing [32]. Loneliness can be defined as having bad feelings of missing relationships. For this reason, we have developed a new scale that combines the network's degree of closeness with the frequency of meetings.

2) Mother's age

The mother's age of an individual significantly influences SWB and stress level. The power of time is pressing on the human race. The feeling, known as "time famine," affects people from all walks of life, including working parents and those with high or low incomes [33–36]. Because people who feel the pressure of time are less likely to be helpful, active, and physically healthy, time famine causes stress and has a detrimental impact on SWB [36, 37]. On the other hand, coming to terms with the fact that time is a finite resource may help someone gain insight that benefits SWB by helping them to appreciate daily activities more [38, 39].

Additionally, when mother age rises, a child's psychological health also improves. Additionally, as maternal age increases, less verbal and physical punishment is used [40]. The perspective of time and the abundance of resources accessible to older moms, which allow them to be emotionally stable, may be the cause [41–43].

3) Individual parent's and network's income level

SWB has an impact on the national and cultural levels, and it has been noted that nations with larger purchasing power exhibit higher levels of well-being [6, 28]. Income may have functional features [44] and may help people in two ways: a) as a resource to protect them from unfavorable life occurrences (medical bills, necessities, etc.); b) to satisfy their needs by buying goods and services [45, 46]. One's spending preferences in their living environment can be influenced by their income, so residing in a neighborhood with both high- and low-income groups may have a negative impact on their SWB [47].

C. Affective Well-Being (AWB)

Since AWB focuses more on an individual's emotional state than CWB does, personality traits and other factors (such as self-esteem) that are related to an individual's affective state have larger relationships with AWB [48]. Due to hereditary causes, the AWB dimension (positive & negative affect) may be linked to personality traits (Big Five) and self-esteem [48–50]. We discussed the Big Five and self-esteem as SWB aspects in the AWB section.

1) Big five

Many psychologists agree that there are five personality dimensions: extraversion, neuroticism, openness, agreeableness, and conscientiousness [51]. Table I provides brief descriptions of the Big Five Personality Traits.

TABLE I: DESCRIPTION OF PERSONALITY TRAITS [52]

Traits	Low	High
Extraversion	Shy	Active
Neuroticism	Stable	Moody
Openness	Commonplace	Imaginative
Agreeableness	Cold	Soft hearted
Conscientiousness	Careless	Organized

The best predictor of SWB is personality traits (individual differences) [53]. Many studies have been done in an effort to explain some of the Big Five's aspects. Extroverts, according to Lucas' theory [54], are more receptive to rewards because they enjoy and value social interactions more. Openness, conscientiousness, and agreeableness demonstrated high relationships with SWB in addition to extraversion and neuroticism personality traits [48]. All of the qualities are linked to both positive and negative affect, with extraversion being the key factor in positive affect and neuroticism being strongly linked to bad affect [48, 55]. The major five were chosen for these reasons: there is a connection between the big five and AWB, and we are aware that AWB is a crucial part of SWB.

2) Self-Esteem

Self-esteem is a general evaluation of one's value, and those who have high self-esteem may believe they are competent and deserving of rewards [56]. Numerous studies have found a significant link between well-being and self-esteem [57, 58]. Self-esteem and life happiness were found to be positively correlated in one study that included participants from 31 different nations [59]. American culture can be characterized as individualistic [60], whereas self-esteem and life happiness are less correlated in collectivistic nations [56]. According to a meta-analysis of 77 research, there is a substantial negative relationship between self-esteem and sadness and anxiety; as a result, self-esteem is predicted to reduce depression [61]. Self-esteem has been included in our heuristic SWB model since it is strongly correlated with SWB and has an impact on psychological suffering [62].

III. PERCEIVED STRESS SCALE

Perceived Stress Scale (PSS) has been set as the target in this study, and PSS is predicted by the SWB determinants. Before going into the specifics of the PSS structure, it is important to emphasize the reasons why PSS was selected as the goal. In our study, we looked into practical SWB determinants that are compatible with the data that were obtained, and we chose PSS as our primary aim (output) since it was one of the scales that was most closely related to measuring subjective well-being among the scales that were gathered. As noted in the section on SWB determinants, which also includes an affective and cognitive component, there are two causes for this target preference. The first reason is that 1) research shows a substantial inverse relationship between the Perceived Stress Scale (PSS) and the Satisfaction with Life Scale (SWLS) [63, 64]. A cognitive evaluation of subjective well-being (SWB) is considered to be the SWLS [65]. The second explanation is that PSS shows a strong correlation with both positive affect (PA) and negative affect (NA), which are terms for affective well-being [66].

The PSS has a 14-item scale and a four-item scale, both of which show high reliability and validity [67]. There are numerous studies that examine the reliability and validity of national PSS-10 versions [68–71].

Even though PSS-10 may not be a scale of psychological symptomatology, it may nonetheless be useful to researchers in spotting early indications of particular clinical psychiatric problems in students or at study locations like workplaces and universities [67]. As a result, participants in the NetHealth Project complete the ten-item PSS. Four and fourteen item PSS are inferior to the ten item scale PSS [77].

IV. MODEL DESIGN

A. Dataset Description

In this study, machine learning classifiers are utilized to predict the amount of stress using SWB characteristics. Approximately 700 college students engaged in the data collection process between 2015 and 2019 using sensors and self-reports [72]. The initiative is referred to as NetHealth Study, and the scope of the collected data includes the information presented in Table II.

TABLE II: NETHEALTH DATASET BRIEF DESCRIPTIONS

Codebooks	Details
Basic Survey Codebook [73]	(Big 5 Personality traits), (Rosenberg Self-Esteem Scale), (Trust), (self-reports for anxiety, depression, stress), (demographics)
Network Survey Codebook [74]	(Self-report filled by the participants and contains wide range information about the network of the individuals)
Communication Events Codebook [75]	(Communication types for each individual such as WhatsApp, SMS etc.)
Fitbit Sleep and activity [76]	(Steps, Bed time & duration, Floor, Mean heart rate, Calories burned etc.)

B. Missing Data

TABLE III: DATA SELECTION FROM RELATED SEMESTERS

	# of person	---	---	209	252	139
SWB Determinants	*Semesters*	W1	W2	W4	W6	W8
CWB	Mother's Age	X				
	Parent's Income	X				
	Network's Income	X				
AWB	Social Relations			X	X	X
	Big Five		X		X	X
	Self-Esteem			X	X	X
PWB	Physical Activity			X	X	X
	Average sleep			X	X	X

The cause for missing data samples is provided in this study, thus we used the individual information from semesters 1, 2, 4, 6, and 8 along with 12 features (determinants), 1 target (output), and a total of 600 samples. The problem with using self-reports from each semester is that some of the reports, such Big Five and self-esteem, were not gathered for each semester and it is presented in Table III. With the exception of the Big Five personality traits, the required determinants and objective are located in semester 4. Therefore, personality traits showed a slight mean level shift over four years [78], and semester 2 Big Five scores are used.

Since semesters 2 and 4 are separated by less than a year, the alternate method is favored. As a result, we only use semester 1 for the general demographic data that was already obtained, and after creating semesters 4, 6, and 8, we have 600 total samples.

C. Determinant Analysis

The analysis of SWB determinants yields the following conclusions: (a) the data are tiny, (b) there are outliers, and (c) the data have ordinal and nominal parameters, allowing both parametric and non-parametric methodologies to be used. Pearson correlation searches for linear correlation, and Spearman correlation examines monotonic relationships. The tables of correlation are shown below.

It is challenging to identify a linear link between the target and the 12 features. Because not all features are linearly and monotonically associated, some of them may have polynomial or other types of correlations as a result. When we examine the correlation values shown in Table IV, we find that the spearman and Pearson correlations are different for each determinant. Additionally, not all p values are significant for every target (P values which are less than 0.1 are italic.)

TABLE IV: CORRELATION AND P VALUE SCORES FOR EACH DETERMINANTS

SWB features	P value	Pearson	Spearman
Physical activity	0.30	-0.17	-0.16
Trust Frequency	*0.10*	-0.03	-0.02
Average sleep	*0.00*	-0.25	-0.24
Parent's income	0.24	-0.01	-0.02
Networks Parent's income	0.44	-0.04	-0.03
Mother's age	*0.00*	-0.09	-0.09
Self-Esteem	*0.00*	-0.59	-0.59
Extraversion	0.11	-0.20	-0.18
Agreeableness	0.89	-0.28	-0.27
Conscientiousness	0.28	-0.35	-0.34
Neuroticism	*0.00*	0.57	0.56
Openness	0.17	-0.08	-0.06

D. Feature Design

Preprocessing is necessary in the NetHealth project because some datasets need to be converted into the correct format and presented in Table V. The preprocessing of the data is briefly discussed in the sections that follow.

1) Cognitive well-being (global judgements of life satisfaction)

1) There is no requirement for conversion because the mother's age is given in numerical format.
2) For the parent's income level, raw data is transformed into numerical groups before being presented in text groups for the parent's and network's income. Additionally, the level of revenue for networks is translated into numbers.

2) Cognitive well-being (satisfaction with specific life domains)

1) Social Relations: Network's Trust Level and Meeting Frequency: Participants in the network survey rate the level of trust in their network and provide information about how frequently they meet. Both of them are translated into numerical values and presented in the category format. Finally, both are multiplied to create a single scale.

3) Affective well-being

1) A 44-question Big Five Personality questionnaire was used to compute the scores for each personality attribute. These results from the fundamental survey are used.
2) Ten items make up the self-esteem questionnaire, and researchers compute the self-esteem score. The basic survey's score is used.

4) Physical well-being

1) Sleeping: The average bedtime according to the Pittsburgh Sleep Quality Index (PSQI) item 4 is based on self-reports. This SWB feature doesn't need to be converted.
2) Individual activity levels are displayed numerically. This SWB feature doesn't need to be converted.

TABLE V: DATA SELECTION FROM RELATED SEMESTERS

SWB Determinants	Semesters	Raw Data Format	Converted into
CWB	Mother's Age	Numeric	---
	Parent's Income	Groups in text	Numerical Groups
	Network's Income	Groups in text	Average of Numerical Groups
	Social Relations	Groups in text	Numerical Scale
AWB	Big Five	Numeric	---
	Self-Esteem	Numeric	---
PWB	Physical Activity	Numerical groups	---
	Average sleep	Numerical groups	---

E. Design of Target

The basic survey's PSS consists of 10 questions, which are categorized and given in text style. Text with categories is transformed into numbers and presented in Table VI.

TABLE VI: DATA CONVERSION PROCESSES FOR TARGET

Targets	Raw Data Format	Converted into
PSS	Groups in text	Numerical Groups

1) Class thresholds

Target is categorized into three classes depending on their quartile scores (0, 1, and 2). Table VII and fig.3 below show quartile scores and class distributions. Fig. 3 demonstrates that although low stress level (class 0) and high stress level (class 2) have roughly the same number of rows (patterns), moderate stress levels (class 1) have a greater number of rows. As a result, there is an issue of class imbalance, which will be mentioned further on in this paper.

TABLE VII: DATA CONVERSION PROCESSES FOR TARGET

Count	600.00
Mean	15.89
Standard Deviation	6.49
Minimum	0.00
25%	11.75
50%	16.00
75%	20.00
Maximum	36.00
Scale Range	0 to 40

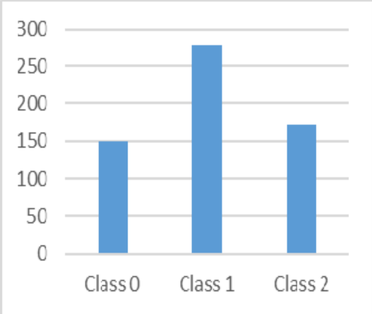

Fig. 3. Class distribution for PSS.

V. Methodology for Prediction

This study's goal is to use SWB factors to categorize three different types of stress level.

The dataset for PSS has a natural structure, which causes the distribution of classes to be uneven. The output must be classified into one class from the non-overlapping classes because of this, making it a multi-classification problem [79]. Numerous strategies have been put out in an effort to improve the performance of pertinent measures for multiclass classification. These strategies can be broken down into the following three basic categories which are data level, algorithmic level and cost sensitive [80].

These techniques could be the subject of additional research. A heuristic SWB model is presented in this study, and predictions from the model are made using some of the most well-known machine learning techniques, such as decision tree classifier or random forest, which may be useful tools for establishing a proper relationship between SWB features and the target (PSS) [81]. Examples of ensemble learning include random forest [82], which is effective in lowering variance bias. Internally, algorithms are developed through the creation of new ones or through revamping current ones [83]. From the standpoint of inductive bias, decision trees [84] and support vector machines with various penalty constants can both have their probabilistic estimation at the tree leaf altered. AdaBoost is one of the effective boosting method examples for cost-sensitive learning, and it may also be minimizing bias [85].

A. Evaluation Metrics

Overall accuracy as a metric might be useful for binary classification, but due to the many misclassification costs, it is insufficient for multiclass classification and is thus better used in conjunction with other metrics [86]. Confusion matrices are used to get some of the evaluation measures [79].

The following metrics [79, 87–89] are taken into account in this study:

Precision $= \dfrac{TP}{(TP+FP)}$: of all estimations, how many are correctly estimated.

Recall $= \dfrac{TP}{(TP+FN)}$: of all true positive class, how many are correctly classified.

F1 score $= 2 \times \dfrac{(Precision \times Recall)}{(Precision \times Recall)}$: the harmonic mean calculation by using precision and recall.

Accuracy $= \dfrac{(TP+TN)}{(TP+FP+TN+FN)}$: Overall efficiency of a classifier.

ROC AUC analysis: True positive rate on the y-axis and false positive rate on the x-axis represents a probabilistic score. When the true positive rate increases meaning that graph approaches to left corner. Area under the curve (AUC) is a single measure for the classification.

VI. Experiment and Results

The Jupyter Notebook used to calculate the results has an i7vpro processor. Technical details are shared in Table VII.

TABLE VIII: Technical Preferences

Technical Aspects (Brief Description)	
Laptop processor	i7vpro
Interface	Jupyter notebook
Missing values	are dropped with Dropna function
Outliers	After application 614 rows reduced to 600 rows
Standardization	Applied
Train test ratio	70% Train, 30% Test
5-fold stratified cross validation	Applied
Applied Machine Learning Algorithms Without Parameter Setting	
AdaBoost, (Base Estimator = Decision Tree)	
Support Vector Machine	
Decision Tree	
K Nearest Neighbors	
Logistic Regression	
Random Forest	

A. Results

The tables below provide the performance characteristics of our heuristic model for the PSS.

The SWB heuristic model has been applied in this study to predict PSS classes (0, 1, and 2). Despite the fact that it is a multiclass classification problem, we have obtained significant metric scores that may be taken into account for a particular challenge. We discovered significant relationships between SWB determinants and PSS in terms of p values and correlation scores as we examined the determinants of the heuristic SWB model in the dataset description section. Additionally, all PSS metrics scores show at least 50% for a particular machine learning algorithm for each class level (low, moderate, high).

The methods section stressed that the cost of misclassification could vary depending on the specific issue. For instance, PSS has three classes, with class 2 having a larger misclassification cost than classes 0 and 1. The classification of people as being in class 2 (high stress) because they have higher levels of stress may be an early marker of psychiatric symptomatology. As is well known, countries bear a significant financial burden related to mental health [90]. That is why misclassifying class 2 could result in increased costs associated with mental illness. (Please refer to Table IX and Table X for the metric scores and best classifiers.)

TABLE IX: METRIC SCORES FOR EACH MACHINE LEARNING ALGORITHMS

	Decision Tree				
Metrics	*Recall*	*Precision*	*F1*	*ACC*	*ROCAUC*
Class 0	0.54	0.50	0.52	0.75	0.68
Class 1	0.51	0.54	0.52	0.57	0.56
Class 2	0.52	0.53	0.52	0.73	0.67
Average	0.53	0.52	0.52	0.68	0.64
	AdaBoost				
Metrics	*Recall*	*Precision*	*F1*	*ACC*	*ROCAUC*
Class 0	0.54	0.49	0.51	0.74	0.68
Class 1	0.49	0.53	0.51	0.57	0.56
Class 2	0.56	0.54	0.55	0.74	0.68
Average	0.53	0.52	0.52	0.68	0.64
	Random Forest				
Metrics	*Recall*	*Precision*	*F1*	*ACC*	*ROCAUC*
Class 0	0.54	0.66	0.59	0.82	0.85
Class 1	0.72	0.60	0.65	0.64	0.67
Class 2	0.58	0.71	0.64	0.81	0.84
Average	0.61	0.66	0.63	0.76	0.79
	Support Vector				
Metrics	*Recall*	*Precision*	*F1*	*ACC*	*ROCAUC*
Class 0	0.47	0.66	0.55	0.81	0.84
Class 1	0.73	0.56	0.63	0.61	0.66
Class 2	0.49	0.66	0.57	0.78	0.84
Average	0.56	0.63	0.58	0.73	0.78

For each class, the best machine learning algorithms are shown below. (Ml: Machine Learning, Svm: Support vector machine, Ada: AdaBoost, Dt: Decision Tree, Rf: Random Forest)

TABLE X: BEST MACHINE LEARNING ALGORITHMS AT CLASS LEVEL

Metrics	**Recall**	**Precision**	**F1**	**ACC**	**ROCAUC**
Class 0	Ada, Rf, Dt	Rf, Svm	Rf	Rf	Rf
Class 1	Svm	Rf	Rf	Rf	Rf
Class 2	Rf	Rf	Rf	Rf	Svm, Rf
Average	Rf	Rf	Rf	Rf	Rf

VII. CONCLUSION

To the best of our knowledge, this paper is the first study to predict stress (PSS) from a Subjective Well-Being (SWB) perspective by the help of machine learning classifiers. We've hypothesized our approach based on the SWB literature and we've reached high metric scores that can be chosen for the right objectives.

In this study, we predicted PSS using SWB determinants as model input. First, we get rid of any outliers that might be present, then we normalize the data, and finally, we use stratified cross validation, which allows us to take samples that are evenly distributed across the train-test population. This study is a multi-classification problem and it is hard to predict each classes with appropriate metric scores. There are well-designed studies for binary classification problems specifically psychological assessments. Metric scores of these studies are approximately %80 accuracy and %70 precision and recall scores [16, 91, 92]. With the random forest ensemble machine learning algorithm, we were able to achieve an accuracy score of 76% along with a precision and recall score of approximately 60% using this method.

A. Threats to Validity

Happiness is only one aspect of well-being, which is why the two terms should not be used interchangeably. Many well-being definitions have been found to be incomplete after the multifaceted design (AWB, CWB) was recognized [93]. That's why our model may fail, however these papers are steadily assisting in the appropriate definition of SWB. The literature review for subjective well-being can be used to identify practical therapies that could reduce stress in individuals. In other words, a person may be made aware of any SWB determinants that they are lacking.

B. Limitation

Time-based data is not appropriately given in the NetHealth data. In other words, number of patterns has decreased to 221 when we use heart rate, sleep time, steps as time-based input. That's why 400 rows are removed. In order to avoid this, we opted a static model with self-reports. Additionally, certain information is not made publicly available, such as content from Twitter and Facebook, and this kind of information can be used to forecast big five test scores, self-esteem, or any other self-reports without having to complete surveys, which is a significant time saver.

C. Future Works

Dynamic models may provide a better answer than static models for predicting the amount of stress, wellbeing, and mental health. Drawing a stronger framework for SWB determinants and targets as part of future work would help to give a comprehensive view. Using a time-based model may make it possible to gather more data on people's wellbeing on an hourly basis. Finally, preprocessing a dataset that contains data from mobile applications like Twitter, Facebook, and Instagram may allow for the identification of some significant correlations between self-reports of big five personality traits and app usage [94].

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

S. Emre Alptekin proposed the main idea related to this work and provided direction and oversight; Ahmet Karakuş and Akif Can Kılıç analyzed the data, implemented and ran the experiments; Ahmet Karakuş wrote the paper; and all authors approved the final version.

FUNDING

This research has been financially supported by Galatasaray University Research Fund, with the project number FBA-2022-1091.

REFERENCES

[1] K. V. Das, C. Jones-Harrell, Y. Fan *et al.*, "Understanding subjective well-being: perspectives from psychology and public health," *Public Health Rev.*, vol. 41, no. 25, 2020.

[2] E. Diener and E. Suh, "Measuring quality of life: Economic, social, and subjective indicators," *Social Indicators Research*, vol. 40(1–2), pp. 189–216, 1997.

[3] A. Campbell, "Subjective measures of well-being," *American Psychologist*, vol. 31, no. 2, pp. 117–124, 1976.

[4] E. Diener, "Subjective well-being," *Psychological Bulletin*, vol. 95, pp. 542–575, 1984.

[5] E. L. Deci and R.M. Ryan, "Hedonia, eudaimonia, and well-being: an introduction," *Journal of Happiness Studies*, vol. 9, pp. 1–11, 2006.

[6] E. Diener, "Subjective well-being. The science of happiness and a proposal for a national index," *Am Psychol.*, vol. 55, no. 1, pp. 34–43, Jan 2000.

[7] M.A. Busseri, and S.W. Sadava, "A review of the tripartite structure of subjective well-being: Implications for conceptualization, operationalization, analysis, and synthesis," *Personality and Social Psychology Review*, vol. 15, no. 3, pp. 290–314, 2011.

[8] G. Feist, T. Bodner, J. Jacobs, M. Miles, and V. Tan, "Integrating top-down and bottom-up structural models of subjective well-being: A longitudinal investigation," *Journal of Personality and Social Psychology*, vol. 68, pp. 138–150, 1995.

[9] C. Scheck, A. Kinicki, and J. Davy, "Testing the Mediating Processes between Work Stressors and Subjective Well-Being," *Journal of Vocational Behavior*, vol. 50, pp. 96–123, 1997.

[10] L.I. Pearlin, "The Stress Process Revisited," in *Handbook of the Sociology of Mental Health*, C. S. Aneshensel and J. C. Phelan, Eds. Springer, Boston, MA, 1999.

[11] L. Pearlin, E. Menaghan, M. Lieberman, and J. Mullan, "The stress process," *Journal of Health and Social Behavior*, vol. 22, no. 4, pp. 337–356, 1981.

[12] L. Pearlin and I. Leonard, "The sociological study of stress," *Journal of Health and Social Behavior*, vol. 30, no. 3, pp. 241–256, 1989.

[13] S. Folkman, "Stress: Appraisal and coping," *Encyclopedia of Behavioral Medicine*, Springer New York, 2013, pp. 1913–1915.

[14] R.S., Lazarus and S. Folkman, *Stress, Appraisal, & Coping*, New York: Springer.pe, 1984.

[15] S.Y. Rueger, C. K. Malecki, and M. K. Demaray, "Relationship between multiple sources of perceived social support and psychological and academic adjustment in early adolescence: Comparisons across gender," *Journal of Youth and Adolescence*, vol. 39, no. 1, p. 47, 2010.

[16] S. Liu, F. Vahedian, D. Hachen, O. Lizardo, C. Poellabauer, A. Striegel, and T. Milenković, "Heterogeneous Network Approach to Predict Individuals' Mental Health," *ACM Trans. Knowl. Discov. Data*, vol. 15, no. 2, Article 25, p. 26, April 2021.

[17] P. Robles-Granda *et al*., "Jointly predicting job performance, personality, cognitive ability, affect, and well-being," *IEEE Computational Intelligence Magazine*, vol. 16, no. 2, pp. 46–61, May 2021.

[18] G. C. Vagetti, V. C. B. Filho, N. B. Moreira, V. de Oliveira, O. Mazzardo, and W. de Campos, "Association between physical activity and quality of life in the elderly: A systematic review, 2000–2012," *Revista Brasileira De Psiquiatria*, vol. 36, no. 1, pp. 76–88, 2014.

[19] O. P. Krueger, "Sustained work, fatigue, sleep loss and performance: A review of the issues," *Work Stress*, vol. 3, pp. 129–141, 1989.

[20] J. J. Pilcher and A. J. Huffcutt, "Effects of sleep deprivation on performance: A meta-analysis," *Journal of Sleep Research & Sleep Medicine*, vol. 19, no. 4, pp. 318–326, 1996.

[21] C. Stoica, "Sleep, a predictor of subjective well-being," *Procedia - Social and Behavioral Sciences*, vol. 187, pp. 443–447, 2015.

[22] J. Li, A. Lepp, and J. E. Barkley, "Locus of control and cell phone use: Implications for sleep quality, academic performance, and subjective well-being," *Computers in Human Behavior*, vol. 52, pp. 450–457, 2015.

[23] D. J. Buysse, C. F. III Reynolds, T. H. Monk, S. R. Berman, and D. J. Kupfer, "The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research," *Psychiatry Research*, vol. 28, no. 2, pp. 193–213, 1989.

[24] L. Faust, R. Purta, D. Hachen, A. Striegel, C. Poellabauer, O. Lizardo, and N. V. Chawla, "Exploring Compliance: Observations from a Large Scale Fitbit Study," in *Proc. the 2nd International Workshop on Social Sensing*, Association for Computing Machinery, New York, NY, USA, pp. 55–60, 2017.

[25] N. P. Gothe, D. K. Ehlers, E.A. Salerno, J. Fanning, A. F. Kramer, and E. McAuley, "Physical activity, sleep and quality of life in older adults: Influence of physical, mental and social well-being," *Behavioral Sleep Medicine*, vol. 18, no. 6, pp. 797–808, 2020.

[26] M. Luhmann, L. C. Hawkley, M. Eid, and J. T. Cacioppo, "Time frames and the distinction between affective and cognitive well-being," *Journal of Research in Personality*, vol. 46, no. 4, pp. 431–441, 2012.

[27] M. Luhmann, W. Hofmann, M. Eid, and R. E. Lucas, "Subjective well-being and adaptation to life events: a meta-analysis," *Journal of Personality and Social Psychology*, vol. 102, no. 3, pp. 592–615, March 2012.

[28] D. G. Myers, "The funds, friends, and faith of happy people," *American Psychologist*, vol. 55, no. 1, pp. 56–67, 2000.

[29] N. Steverink, G. J. Westerhof, C. Bode, and F. Dittmann-Kohli, "The personal experience of aging, individual resources, and subjective well-being," *J Gerontol B Psychol Sci Soc Sci*, vol. 56, no. 6, pp. 364–373, November 2001.

[30] G. Windle and R. Woods, "Variations in subjective wellbeing: The mediating role of a psychological resource," *Ageing and Society*, vol. 24, no. 4, pp. 583-602, 2004.

[31] D. N. McIntosh, R. C. Silver, and C. B. Wortman, "Religion's role in adjustment to a negative life event: Coping with the loss of a child," *Journal of Personality and Social Psychology*, vol. 65, pp. 812–821, 1993.

[32] J. D. J. Gierveld and T.V. Tilburg, "A 6-Item Scale for Overall, Emotional, and Social Loneliness: Confirmatory Tests on Survey Data," *Research on Aging*, vol. 28, no. 5, pp. 582–598, 2006.

[33] L. A. Perlow, "The time famine: Toward a sociology of work time," *Administrative Science Quarterly*, vol. 44, no. 1, pp. 57-81, 1999.

[34] D. S. Hamermesh and J. Lee, "Stressed out on four continents: Time crunch or yuppie kvetch?" *The Review of Economics and Statistics*, vol. 89, no. 2, pp. 374–383, 2007.

[35] R. E. Goodin, J. M. Rice, M. Bittman, and P. Saunders, "The time-pressure illusion: Discretionary vs. free time," *Social Indicators Research*, vol. 73, no. 1, pp. 43–70, 2005.

[36] L. Strazdins, A. L. Griffin, D. H. Broom, C. Banwell, R. Korda, J. Dixon, and J. Glover, "Time scarcity: Another health inequality?" *Environment and Planning A*, vol. 43, no. 3, pp. 545–559, 2011.

[37] C. Mogilner, Z. Chance, and M. I. Norton, "Giving time gives you time," *Psychological Science*, vol. 23, no. 10, pp. 1233–1238, 2012.

[38] L. L. Carstensen, D. M. Isaacowitz, and S. T. Charles, "Taking time seriously: A theory of socioemotional selectivity," *American Psychologist*, vol. 54, no. 3, pp. 165–181, 1999.

[39] P. J. Cozzolino, K. M. Sheldon, T. R. Schachtman, and L. S. Meyers, "Limited time perspective, values, and greed: Imagining a limited future reduces avarice in extrinsic people," *Journal of Research in Personality*, vol. 43, no. 3, pp. 399–408, 2009.

[40] T. Trillingsgaard and D. Sommer, "Associations between older maternal age, use of sanctions, and children's socio-emotional development through 7, 11, and 15 years," *European Journal of Developmental Psychology*, vol. 15, no. 2, pp. 141–155, 2018.

[41] A. Camberis, C. A. McMahon, F. L. Gibson, and J. Boivin, "Age, psychological maturity, and the transition to motherhood among English-speaking Australian women in a metropolitan area," *Developmental Psychology*, vol. 50, pp. 2154–2164, 2014.

[42] A. Camberis, C. A. McMahon, F. L. Gibson, and J. Boivin, "Maternal age, psychological maturity, parenting cognitions, and mother-infant interaction," *Infancy*, vol. 21, pp. 396–422, 2016.

[43] L. L. Carstensen, B. Turan, S. Scheibe, N. Ram, H. Ersner-Hershfield, G. Samanez-Larkin, and J. R. Nesselroade, "Emotional experience improves with age: Evidence based on over 10 years of experience sampling," *Psychology and Aging*, vol. 26, pp. 21–33, 2011.

[44] L. Tay, M. Zyphur, and C. L. Batz, "Income and subjective well-being: Review, synthesis, and future research," in *Handbook of Well-Being*, E. Diener, S. Oishi, and L. Tay, Eds. Salt Lake City, UT: DEF Publishers, 2018.

[45] E. Diener, L. Tay, and S. Oishi, "Rising income and the subjective well-being of nations," *Journal of Personality and Social Psychology*, vol. 104, no. 2, pp. 267–276, 2013.

[46] L. Tay and E. Diener, "Needs and subjective well-being," *Journal of Personality and Social Psychology*, vol. 101, pp. 354–365, 2011

[47] E. F. P. Luttmer, "Neighbors as Negatives: Relative earnings and well-being," *The Quarterly Journal of Economics*, vol. 120, no. 3, pp. 963–1002, August 2005.

[48] P. Steel, J. Schmidt, and J. Schultz, "Refining the relationship between personality and subjective well-being," *Psychological Bulletin*, vol. 134, no. 1, pp. 138–161, 2008.

[49] S. Lyubomirsky, L. King, and E. Diener, "The benefits of frequent positive affect: Does happiness lead to success?" *Psychological Bulletin*, vol. 131, no. 6, pp. 803–855, 2005.

[50] C. Zanon, M. R. Bastianello, J .C. Pacico, and C. S. Hutz, "Relationships between positive and negative affect and the five factors of personality in a brazilian sample," *Paidéia (Ribeirão Preto)*, vol. 23, no. 56, pp. 285–292, 2013.

[51] R. R. McCrae and O. P. John, "An introduction to the five-factor model and its applications," *Journal of Personality*, vol. 60, no. 2, pp. 175–215, June 1992.

[52] O. P. John and S. Srivastava, "The big five trait taxonomy: History, measurement, and theoretical perspectives," *Handbook of personality: Theory and Research*, pp. 102–138, 1999.

[53] R. E. Lucas, "Exploring the associations between personality and subjective well-being," in *Handbook of Well-Being*, E. Diener, S. Oishi, and L. Tay, Eds., Salt Lake City, UT: DEF Publishers, 2018.

[54] E. Diener, E. M. Suh, R. E. Lucas, and H. L. Smith, “Subjective well-being: Three decades of progress,” *Psychological Bulletin*, vol. 125, no. 2, p. 276, 1999.

[55] M. S. M. Yik, and J. A. Russell, “Predicting the big two of affect from the big five of personality,” *Journal of Research in Personality*, vol. 35, pp. 247–277, 2001.

[56] S. Margolis, and S. Lyubomirsky, “Cognitive outlooks and well-being,” in *Handbook of Well-Being*, E. Diener, S. Oishi, and L. Tay, Eds., Salt Lake City, UT: DEF Publishers, 2018.

[57] S. Lyubomirsky, C. Tkach, and M. R. DiMatteo, “What are the differences between happiness and self-esteem?” *Social Indicators Research*, vol. 78, pp. 363–404, 2006.

[58] A. W. Paradise and M. H. Kernis, “Self-esteem and psychological well-being: Implications of fragile self-esteem,” *Journal of Social and Clinical Psychology*, vol. 21, pp. 345–361, 2002.

[59] E. Diener and M. Diener, “Cross-cultural correlates of life satisfaction and self-esteem,” *Journal of Personality and Social Psychology*, vol. 69, pp. 851–864, 2009.

[60] D. Oyserman, H. M. Coon, M. Kemmelmeier, “Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses,” *Psychological Bulletin*, vol. 128, no. 1, pp. 3–72, 2002.

[61] J. F. Sowislo and U. Orth, “Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies,” *Psychological Bulletin*, vol. 139, pp. 213-240, 2013.

[62] R. W. Robins, H. M. Hendin, and K. H. Trzesniewski, “Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg self-esteem scale,” *Personality and Social Psychology Bulletin*, vol. 27, no. 2, pp. 151–161, 2001.

[63] S. Yew, M. K. Lim, Y. Haw, and S. K. Gan, “The association between perceived stress, life satisfaction, optimism, and physical health in the Singapore Asian context,” *Asian Journal of Humanities and Social Sciences*, vol. 3, no. 1, pp. 1–11, 2015.

[64] S. Park and K. F. Colvin, “Psychometric properties of a Korean version of the Perceived Stress Scale (PSS) in a military sample,” *BMC Psychol*, vol. 7, article 58, 2019.

[65] I. Aslan, D. Ochnik, and O. Çınar, “Exploring Perceived Stress among Students in Turkey during the COVID-19 Pandemic,” *International Journal of Environmental Research and Public Health*, vol. 17, no. 23, p. 8961, 2020.

[66] A. Civitci, “The moderating role of positive and negative affect on the relationship between perceived social support and stress in college students,” *Kuram ve Uygulamada Eğitim Bilimleri*, vol. 15, no. 3, pp. 565–573, 2015.

[67] S. Cohen, T. Kamarck, and R. Mermelstein, “A global measure of perceived stress,” *Journal of Health and Social Behavior*, vol. 24, no. 4, pp. 385–396, 1983.

[68] C. D. Bastianon, E. M. Klein, A. N. Tibubos *et al*., “Perceived Stress Scale (PSS-10) psychometric properties in migrants and native Germans,” *BMC Psychiatry*, vol. 20, article 450, 2020.

[69] J. M. Taylor, “Psychometric analysis of the ten-item perceived stress scale,” *Psychol Assess*, vol. 27, no. 1, pp. 90–101, March 2015.

[70] F. X. Lesage, S. Berjot, and F. Deschamps, “Psychometric properties of the French versions of the perceived stress scale,” *Int J Occup Med Environ Health*, vol. 25, pp. 178–184, 2012.

[71] R. Khalili R, N. M. Sirati, A. Ebadi, A. Tavallai, and M. Habibi, “Validity and reliability of the Cohen 10-item Perceived Stress Scale in patients with chronic headache: Persian version,” *Asian J Psychiatr.*, vol. 26, pp. 136–40, 2017.

[72] University of Notre Dame. Sample. [Online]. Available: http://sites.nd.edu/nethealth/sample-2/

[73] University of Notre Dame. BASIC SURVEY DATA. [Online]. Available: http://sites.nd.edu/nethealth/basic-survey-data/

[74] University of Notre Dame. NETWORK SURVEY DATA. [Online]. Available: http://sites.nd.edu/nethealth/network-survey-data/

[75] University of Notre Dame. COMMUNICATION EVENT DATA. [Online]. Available: http://sites.nd.edu/nethealth/communication-event-data/

[76] University of Notre Dame. FITBIT DATA. [Online]. Available: http://sites.nd.edu/nethealth/fitbit-data/

[77] E. H. Lee, “Review of the psychometric evidence of the perceived stress scale,” *Asian Nurs Res (Korean Soc Nurs Sci)*, vol. 6, no. 4, pp. 121–127, December 2012.

[78] J. Specht, B. Egloff, and S. C. Schmukle, “Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five,” *Journal of Personality and Social Psychology*, vol. 101, no. 4, p. 862–882, 2011.

[79] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, 2009.

[80] A. Fernández, V. López, M. Galar, M. J. del Jesus, and F. Herrera, “Analysing the classification of imbalanced data-sets with multiple classes: Binarization techniques and ad-hoc approaches,” *Knowledge-Based Systems*, vol. 42, pp. 97–110, 2013.

[81] D. R. Cutler, T. C. Jr. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler, “Random forests for classification in ecology,” *Ecology*, vol. 88, pp. 2783–2792, 2007.

[82] L. Breiman, “Bagging predictors,” *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[83] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, “A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches,” in *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 4, pp. 463–484, July 2012.

[84] J. R. Quinlan, “Improved estimates for the accuracy of small disjuncts,” *Mach. Learn.*, vol. 6, pp. 93–98, 1991.

[85] Y. Freund and R. E. Schapire, “Experiments with a new boosting algorithm,” in: *Proc. the Thirteenth International Conference on Machine Learning*, The Mit Press, Cambridge, MA, Morgan Kaufmann, Los Altos, CA, 1996, pp. 148–156.

[86] Z. Weiwei, H. Guang-Bin, and C. Yiqiang, “Weighted extreme learning machine for imbalance learning,” *Neurocomput*, vol. 101, pp. 229–242, February 2013.

[87] T. Fawcett, “An introduction to ROC Analysis,” *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

[88] M. Kubat, R. Holte, and S. Matwin, “Machine learning for the detection of oil spills in satellite radar images,” *Mach. Learn.*, vol. 30, pp. 195–215, 1998.

[89] Y. Sun, A.K. Wong, M.S. Kamel, “Classification of Imbalanced Data: a Review,” *Int. J. Pattern Recognit. Artif. Intell.*, vol. 23, pp. 687–719, 2009.

[90] M. K. Christensen, C. Lim, S. Saha *et al.*, “The cost of mental disorders: A systematic review,” *Epidemiology and Psychiatric Sciences*, vol. 29, article 161, 2020.

[91] T. Yarkoni and J. Westfall, “Choosing prediction over explanation in psychology: Lessons from machine learning,” *Perspectives on Psychological Science*, vol. 12, no. 6, pp. 1100-1122, 2017.

[92] H. Rosenbusch, F. Soldner, A. M. Evans, and M. Zeelenberg, “Supervised machine learning methods in psychology: A practical introduction with annotated R code,” *Soc Personal Psychol Compass*, vol. 15, 2021.

[93] W. Tov, “Well-being concepts and components,” in *Handbook of Well-Being*, E. Diener, S. Oishi, and L. Tay, Eds., Salt Lake City, UT: DEF Publishers, 2018.

[94] E. Peltonen, P. Sharmila, K.O. Asare, A. Visuri, E. Lagerspetz, and D. Ferreira, “When phones get personal: Predicting Big Five personality traits from application usage,” *Pervasive Mob. Comput.*, vol. 69, article 101269, 2020.

Copyright © 2023 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).