# WiP: Prediction of QoL in healthy older adults using non-motor information from smart devices

**Abstract**

Although global interest in well-being and QoL is increasing, continuous awareness of one's QoL in daily life remains challenging due to the need for repeated questionnaire responses. In this study, we evaluate the performance of a prediction model for QoL in healthy older adults using a Garmin Venu 3S fitness tracker and a FonLog data collection application to collect non-motor information and QoL data, and predict QoL using a support vector machine (SVM). The results of the prediction using a SVM showed that the Accuracy was approximately 0.96 and the F1-Score for each class was approximately 0.88 or higher. These results suggest the effectiveness of the QoL prediction model using non-motor information. In the future, we plan to improve the processing and prediction in real time, and to evaluate the accessibility, usability, and effectiveness of the system for a wider range of users through experiments with non-motor subjects.

**Contribution of the Paper:** Since smartwatches can be worn easily and are not difficult to use on a daily basis, the development of a prediction model is expected to be an easy way to measure QoL.

*Keywords:* well-being, QoL, non-motor, smart device, machine-learning



*ISSN: 2186-1390 (Online)*
*http://cennser.org/IJCVSP*

*Article History:*
*Received: 1 July 2018*
*Revised: 1 September 2018*
*Accepted: 20 September 2018*
*Published Online: 25 September 2018*

## 1. INTRODUCTION

There is a growing interest in well-being and quality of life (QoL) around the world. Well-being encompasses not only physical health, but also mental and social dimensions, and refers to happiness throughout life. It encompasses both individual and societal prosperity, and in economically advanced countries, the concept of well-being is being emphasized as a way of thinking about happiness. For example, Oxford University's Center for Well-Being Studies, in collaboration with the United Nations, publishes the annual World Happiness Report, which assesses the QoL in each country [1].

Typically, QoL is assessed through multi-question surveys that are impractical for continuous daily monitoring. In this study, we evaluate the performance of a QoL prediction model based on non-motor information collected from healthy older adults; non-motor information refers to data such as heart rate and step count, which can be obtained by simply wearing a commercially available smartwatch. Therefore, if a predictive model can be established, QoL can be easily measured.

## 2. RELATED WORKS

Victorino et al. conducted a study of Parkinson's disease using non-motor information as input to a predictive model [2]. This study aimed to predict the wear-off phenomenon, which is a recurrence of symptoms in Parkinson's patients, and demonstrated the feasibility of accurate prediction using non-motor data. We have also developed a system that can collect data from smartwatches, allowing easy input of non-motor information.

Non-motor information can be obtained by anyone who can wear a smartwatch, and can be used as an indicator of health status and life rhythm, so it can be used to predict various events, not only Parkinson's disease. In this study, QoL was used as a prediction target so that it can be used by healthy people who do not suffer from the disease, and it was decided to evaluate how accurate the results actually are.

## 3. METHOD

This study uses a Garmin Venu 3S smartwatch, the FonLog app, and a prediction model. This is adapted from the model developed by Victorino et al, but the questions in the FonLog app are changed to QoL indices [2]. The

Garmin Venu 3S acquires the patient's non-motor information such as heart rate, and the FonLog app records the questionnaire about QoL.

The Garmin Venu 3S is a fitness tracker from Garmin that can measure a variety of data from the device's onboard sensors. It weighs 40.0 g and measures 41 x 41 x 12 mm. It is also waterproof, making it suitable for data collection [3]. In this study, heart rate, stress score, number of steps, and sleep data will be collected.

The Garmin Venu 3S data set is measured by a photoplethysmograph (PPG) and accelerometer sensor. the PPG sensor estimates heart rate by detecting changes in the intensity of reflected light using light shone on the skin. The changes in reflected light are due to contraction and expansion of arteries and arterioles caused by pulsating blood pressure. In addition to heart rate, Venu 3S uses heart rate variability (HRV) to estimate stress levels. Sleep stage, on the other hand, is estimated from a combination of heart rate, heart rate variability, and accelerometer data. Finally, Venu 3S uses accelerometer data to estimate the number of steps taken.

Each data set available on the Garmin Venu 3S has a different time interval: heart rate, expressed in beats per minute (bpm), is provided every 15 seconds. Steps taken are accumulated every 15 minutes, and stress scores are reported every 3 minutes. Scores range from 0 to 100, with 100 indicating the highest stress. Other values include "-1" for insufficient data to estimate stress and "-2" for too much movement. Each sleep phase is measured along with its start and end times.

FonLog is an Android smartphone application used primarily as a data collection tool for human activity recognition in nursing services [4]. In this study, EuroQol's EQ-5D-5L was used to assess and label each participant's QoL [5].EQ-5D-5L contains the following five items:

1. Degree of mobility
2. Personal care
3. Daily activities
4. Pain / discomfort
5. Anxiety / distraction

For each of these items, the degree of problem is recorded on five levels: no problem, minor problem, moderate problem, serious problem and extreme problem. Each item is recorded on a 5-point scale: no problem, minor problem, moderate problem, serious problem, and extreme problem. In addition to the EQ-5D-5L, the participants also answer a questionnaire in which they write their impressions of their participation in the experiment.

System users were provided with a Garmin Venu 3S. The Garmin Connect app is developed by Garmin and allows users to check their data transmitted from the Venu 3S to their smartphones via Bluetooth. The Garmin Connect app is developed by Garmin, Inc. Meanwhile, the Venu 3S dataset is automatically sent to a server system on Amazon Web Services (AWS) [6].Data collected by the FonLog app is also automatically sent to the FonLog server for storage and retrieval.

There are no strict limitations or restrictions during the data collection period, but system users are asked to wear the Garmin Venu 3S while bathing and sleeping as much as possible. It is also emphasized that users should use the FonLog application to record their data approximately three times per day. Users can modify and review their responses over time.

The server system that receives the data from the Garmin Health API stores the data in a database and makes it available for downloading after the experiment is complete. When downloading data, the server system applies the user ID assigned to each system user and the time period of the data to be retrieved to SQL statements, and retrieves the matching data from the database.

After data extraction, pre-processing is performed. First, the raw sleep data is converted to match other datasets. Datasets such as heart rate are reported for each date and time. As an example, the raw heart rate dataset is shown in Table 1.

Table 1: Raw heart rate dataset from Garmin Health API

| Time | Heart Rate |
|---|---|
| 2021/3/23 1:51:15 | 84 |
| 2021/3/23 1:51:30 | 84 |
| 2021/3/23 1:51:45 | 84 |

On the other hand, the raw sleep dataset is reported for each calendar day and each sleep category. An example is shown in Table 2.

Table 2: Raw sleep dataset from Garmin Health API (dates omitted)

| Start Time | End Time | Sleep Type |
|---|---|---|
| 2021-02-23 02:24:00 | 2021-02-23 02:32:00 | Light |
| 2021-02-23 02:32:00 | 2021-02-23 02:33:00 | Awake |
| 2021-02-23 02:33:00 | 2021-02-23 02:36:00 | Light |

Since the prediction model requires a unified dataset, all input data were merged, each sleep stage is converted to a minute feature and distributed across calendar days to match the other datasets that were resampled at a specific interval. An example of the transformed sleep data is shown in Table 3.

Table 3: Converted sleep dataset

| Calendar Date | Awake | Deep | Light | REM |
|---|---|---|---|---|
| 2021-02-23 | 2.0 | 0.0 | 150.0 | 27.0 |
| 2021-02-25 | 0.0 | 150.0 | 66.0 | 54.0 |
| 2021-02-23 | 0.0 | 83.0 | 55.0 | 0.0 |

Next, several sleep features were computed from the transformed sleep dataset.

$$\text{Total non-REM duration} = \text{Deep sleep duration} + \text{Light sleep duration} \quad (1)$$

$$\text{Total sleep duration} = \text{Total non-REM duration} + \text{REM sleep duration} \quad (2)$$

$$\text{Total non-REM percentage} = \frac{\text{Total non-REM duration}}{\text{Total sleep duration}} \quad (3)$$

$$\text{Sleep efficiency} = \frac{\text{Total sleep duration}}{\text{Total sleep duration} + \text{Total awake duration}} \quad (4)$$

Total non-REM duration is the state in which the body and brain are at rest and is the sum of Deep sleep duration and Light sleep duration (equation (1)). Total sleep duration is the sum of total non-REM duration and REM sleep duration (equation (2)). Total non-REM percentage is the ratio of the Total non-REM duration to the Total sleep duration (equation (3)) [7]. Sleep efficiency is the ratio of sleep duration to total sleep duration, including awake duration (equation (4)) [8].

After processing the sleep datasets, cleanup and resampling were performed on all raw datasets. First, a “-1” was assigned to the missing values before resampling. The “-1” value was used according to the way the Garmin Health API reports stress scores that could not be estimated. In addition, a value of “-1” indicates that the fitness tracker was not worn. Next, we resampled each dataset to match the intervals. The data set was resampled at 15 minute intervals and missing values were forward filled. This approach simulates real-time streaming from fitness trackers, where future data are unavailable at the time of prediction.

The survey results were similarly complemented (forward fill) with the previous data and merged with the resampled Garmin data. This data set was then fed into the prediction model.

Although there are no restrictions on the algorithm used for the prediction model, we use a support vector machine (SVM), which is currently considered one of the best learning models in pattern recognition due to its high discriminative performance on untrained data [9]. SVM aims to learn effective separating hyperplanes for classification problems in high-dimensional spaces, and in particular to improve generalization performance by finding decision boundaries that maximize the margin.

SVM can be classified using simple hyperplanes when data is linearly separable, but when linear separation is difficult, it can be applied to nonlinear data at low computational cost by using the kernel trick. The kernel trick transforms data that are not linearly separable in the original space into a linearly separable form by mapping the data to a higher-dimensional space.

Typical kernel functions include linear kernel, Gaussian kernel (RBF kernel), polynomial kernel, and sigmoid kernel. Since the choice of kernel function greatly affects classification accuracy, it is important to select an appropriate kernel according to the distribution and characteristics of the data.

SVM is characterized by (1) high classification accuracy even with small amounts of data, (2) the ability to handle nonlinear problems through kernel tricks, and (3) relative robustness against outliers. In particular, SVM is known for its good classification performance even with small sample sizes and high-dimensional data, and is widely used in fields such as image recognition and bioinformatics. SVM was also employed in this study because (1) the data sample was not that large, (2) classification based on non-motor information was expected to be a nonlinear problem, and (3) outliers due to non-wearing of smart devices, etc. were possible.

## 4. EXPERIMENT

In this experiment, we evaluated the accuracy of predicting QoL from non-motor information with the cooperation of 10 healthy older adults. QoL was recorded over a period of 21 days using FonLog with a Garmin Venu 3S. The recorded data sets were downloaded from each server to a local computer and evaluated using an SVM prediction model. Personal information, such as names, was collected to distinguish and contact each participant during the experiment, but was not included in the data set or in the predictive model.

First, the collected data were preprocessed; the Garmin dataset was preprocessed and then the QoL records obtained from the FonLog were merged into each record at 15-minute intervals. For each 15-minute interval, a label of 1 was assigned when the QoL was not at its highest state during that 15-minute period, and a label 0 when it was at its highest state. This labeling was done to account for classification bias, as many of the participants gave the highest ratings.

To deal with unbalanced data, we multiplied the input data by the inverse of the number of label occurrences relative to the total number of samples to align the ratio of weights. For the hyperparameters required for the model, we tried all combinations from the following list and selected the best model.

1. C parameter selected from 1, 10, 100
2. Kernel is RBF (Radial Basis Function kernel)
3. RBF parameter, gamma (kernel influence range) selected from 1, 0.1, 0.01

## 5. RESULTS

First, we confirmed the imbalance of the data collected in this study. While the total number of records for the 10 participants in the experiment was 15056, the number of records assigned label 1 was 4214, or slightly less than 30 percent of the records were assigned label 1.

After searching the list for the optimal combination of parameters, the best performance on the test data was obtained when C=100 and gamma=0.1. Using this optimal model, we calculated the Accuracy, Precision, Recall, and F1-score. The results are shown in Table 4.

Table 4: Classification Results for Healthy Older Data

| Class | Precision | Recall | F1-score |
|---|---|---|---|
| 0 | 0.995624 | 0.949896 | 0.972222 |
| 1 | 0.803279 | 0.980000 | 0.882883 |
| Accuracy | 0.955095 | 0.955095 | 0.955095 |
| Macro Avg | 0.899451 | 0.964948 | 0.927553 |
| Weighted Avg | 0.962403 | 0.955095 | 0.956792 |

The confusion matrix is shown in Figure 1, and the importance of each feature is shown in Figure 2.

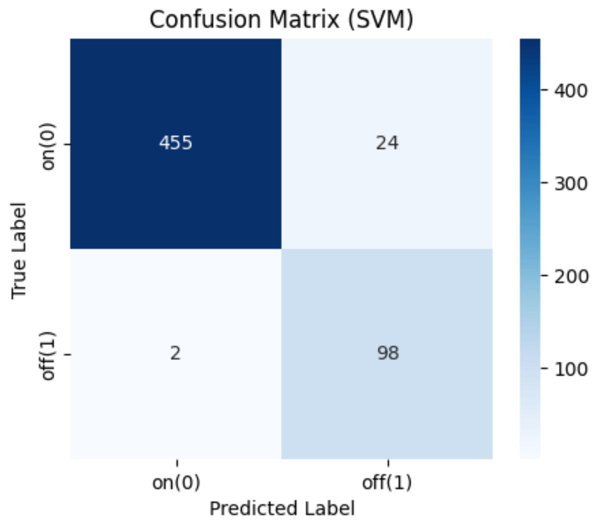


Figure 1: Confusion matrix of healthy data

The prediction results were 0.955095 for Accuracy, 0.882883 for Class 1 F1-Score, and 0.972222 for Class 0 F1-Score, all of which indicate strong model performance. This indicates that it may be effective to predict QoL from non-motor information using a prediction model. The confusion matrix in Figure 1 also shows that many of the data were correctly classified. In addition, Figure 2 highlights sleep-related metrics as key predictors of QoL.

## 6. CONCLUSION

In this study, QoL was predicted using non-motor information obtained from a smartwatch, and the model was evaluated. SVM was used to predict QoL for healthy older adults, with Class 1 being the case where QoL is not at its highest state. The results showed that the Accuracy was 0.955095, and the F1-Score for each class was high. The smartwatch is easy to wear and suitable for daily use, indicating its usefulness as a simple means of quantitative QoL monitoring.

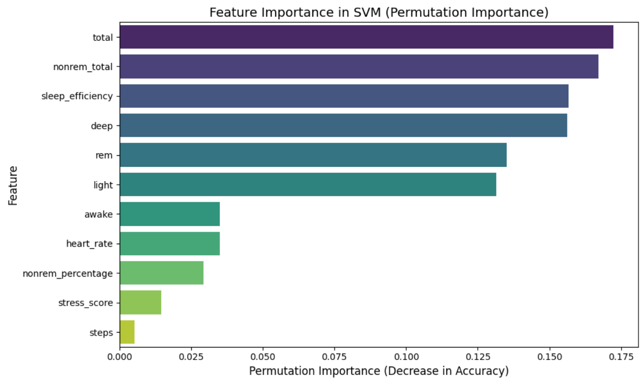


Figure 2: Permutation Importance in SVM

While the current study shows promising results, the effectiveness of the system for the general population is unknown due to the small sample size. Future studies will test the system on a larger and more diverse population to see if the system works for more than just healthy older adults. In particular, we plan to evaluate the accessibility, usability, and effectiveness of the system for a wider range of users by conducting experiments with non-healthy users, which were not addressed in this study. Currently, the collected data are processed and labeled together and input into a prediction model. Future work will focus on enabling real-time data processing and QoL forecasting. Specifically, we plan to achieve this by running the prediction model on a server and automatically collecting nonmotor information and QoL questionnaire results.

## References

[1] the Wellbeing Research Centre at the University of Oxford, Home — the world happiness report, [(Accessed on 04/17/2025)]. URL `https://worldhappiness.report/`

[2] J. N. Victorino, Y. Shibata, S. Inoue, T. Shibata, Predicting wearing-off of parkinson's disease patients using a wrist-worn fitness tracker and a smartphone: A case study, Applied Sciences 11 (16) (2021) 7354.

[3] Garmin, Garmin venu 3s, `https://www.garmin.com/en-US/p/873214`, (Accessed on 04/17/2025).

[4] S. Lab, Fonlog, `https://play.google.com/store/apps/details?id=jp.sozolab.gtolog&pli=1`, (Accessed on 04/17/2025).

[5] EuroQol, Eq-5d-5l - euroqol, `https://euroqol.org/information-and-support/euroqol-instruments/eq-5d-5l/`, (Accessed on 04/17/2025).

[6] G. Developer, Health api, `https://developer.garmin.com/gc-developer-program/health-api/`, (Accessed on 04/17/2025).

[7] L. A. Schroeder, O. Rufra, N. Sauvageot, F. Fays, V. nina Pieri, N. J. Diederich, Reduced rapid eye movement density in parkinson disease: A polysomnography-based case-control study, Sleep 39.12 (2016) 2133–2139.

[8] D. L. Reed, W. P. Sacco, Measuring sleep efficiency: What should the denominator be?, Journal of Clinical Sleep Medicine : JCSM : Official Publication of the American Academy of Sleep Medicine 12.2 (2016) 263–266.

[9] T. Kurita, Introduction to support vector machines, Brain Neuroinformatics Research Institute, National Institute of Advanced Industrial Science and Technology (AIST) 2002 (2002) 1.