

An Estimation Method of Intellectual Concentration State by Machine Learning of Physiological Indices

Kaku Kimura¹, Shutaro Kunimasa¹, You Kusakabe¹, Hirotake Ishii¹,
and Hiroshi Shimoda¹

¹ Graduate School of Energy science, Kyoto University, Kyoto, Japan
{kaku,kunimasa,kusakabe,hirotake,shimoda}@ei.energy.kyoto-u.ac.jp

Abstract. Although recent information society has improved the value of intellectual work productivity, its objective and quantitative evaluation has not been established. It is suggested that intellectual productivity can be indirectly evaluated by estimating intellectual concentration states when giving cognitive load. In this study, therefore, the authors have focused on physiological indices such as pupil diameter and heart rate which are supposed to be closely related to cognitive load in office work, and an estimation method of intellectual concentration states from the measured indices has been proposed. Multiple patterns of classification learning methods such as Decision Tree, Linear Discrimination, SVM, and KNN were employed as the estimation method. Based on the estimation method, an evaluation experiment was conducted where 31 male university students participated and the measured psychological indices were given to the classification learning estimators.

Keywords: Intellectual Concentration State · Machine Learning · Physiological Indices

1 Introduction

Recent information society has improved the value of intellectual work and companies have been tackling improvement of intellectual work productivity by improving office environment. The improvement of intellectual productivity can introduce not only their own profit but also social benefit in total. Although the intellectual productivity is very important in this modern society, its objective and quantitative evaluation method has not been established. Considering the objective and quantitative evaluation method of intellectual productivity, most of the actual office work is simple intellectual work and its work efficiency is closely related to intellectual concentration which is also related to cognitive load while working. This means that intellectual productivity can be indirectly evaluated by estimating intellectual concentration states in office work.

In this study, therefore, the authors have focused on physiological indices which are supposed to be closely related to cognitive load in office work [1][2], and an estimation method of intellectual concentration states from physiological indices has been proposed. Pupil diameters and heart rate variability were employed as the indices which

are supposed to be affected by their cognitive load. In addition, 4 types and 11 patterns of classification machine learning methods such as Decision Tree, Linear Discrimination, SVM (Support Vector Machine) and KNN (K-Nearest Neighbors) were employed as the estimation methods and the concentration states were estimated by the estimator with the highest classification performance among them.

The concentration states to be estimated were one of three states in this study when giving three kinds of cognitive loads which were high, medium and low to artificially generate high, medium and low concentration states, respectively. If it is confirmed that this estimation method is effective, real-time estimation of the intellectual concentration state in the actual office environment can be realized, and an application can be expected as one of the objective and quantitative evaluation method of intellectual productivity.

2 Estimation Method of Intellectual Concentration State

2.1 Physiological Indices during Cognitive Task

2.2 Pupil diameter and heart rate are employed in this study as measurement indices for extraction of feature value. Both the pupil diameter and the heart rate are known as indicators which easily vary when giving cognitive load [3][4]. Regarding the heart rate, the electrocardiogram waveform is measured by a polygraph, Polymate AP216, and the four values, HR(Heart Rate), LF(Spectrum of Low Frequency), HF(Spectrum of High Frequency), LF / HF are extracted. 60 seconds where the power spectrum of the LF band contains for at least 3 cycles is defined as one section, and this section is shifted every 30 seconds, and the average value of each section excluding the start of 30 seconds is extracted as the feature value of the section. The pupil diameter is detected by a measurement device with an infrared camera, FaceLab 5, and the left and right pupil diameters are measured. Also, the extraction of the feature value is the average diameter of each time section. Thus, Totally six values which are HR, LF, HF, LF / HF, left and right pupil diameters are used as explanatory variables for estimation. On the other hand, the objective variable is one of three intellectual concentration states set by changing their cognitive load. The control of the cognitive load is realized by changing answering method of a cognitive task, Receipt- Classification Task. The method of answering tasks is as follows; Task A: High pace, Task B: Slow pace, Task C: Click (do not solve the tasks). Then, the concentration state corresponding to each task's answering method is defined as high, middle, and low concentration state respectively. In this study, the answer time of each task is set to 5 minutes which is considered to keep the same concentration state. Therefore, the total number of data for each variable in one task is 24.

2.3 Machine Learning

It is supposed that physiological response is greatly depending on the individuals, even though it shows common tendency. In this study, estimators are individually generated by machine learning of measured physiological indices data for three kinds of

intellectual concentration states. For the machine learning, Four classification learning methods which are decision tree, linear discrimination, SVM, KNN were employed and realized with MATLAB [5] application. Table 1 shows the classification learning method of all four types and 11 patterns. Regarding SVM, in order to realize multi-class classification with binary classifier, ECOC [6] is used.

Table1. List of classification methods

Classification Method	Remarks
1. Decision Tree	
2. Linear Discrimination	
3. Quadratic Discrimination	
4. Linear SVM	
5. Quadratic SVM	
6. Cubic SVM	
7. Fine Gaussian SVM	$\sigma = 0.6$
8. Middle Gaussian SVM	$\sigma = 2.4$
9. Row Gaussian SVM	$\sigma = 9.8$
10. Fine KNN	$k=1$
11. Row KNN	$k=10$

2.4 Evaluation Method of Estimation Accuracy

As described in section 2.1, a total of 24 training data are extracted from measuring physiological indices while the cognitive task is performed. 11 patterns of estimators as shown in Table 1 are generated by applying classification learning. The generated estimators are evaluated its classification performance by cross validation of 24 division. Finally, one with the highest classification performance among the 11 patterns is selected as the best estimator and apply it to the estimation of the unknown data. The unknown data means test data measured at different time from the training data. In this way, the correct estimation rate of the intellectual concentration state is evaluated and the average value of all the correct estimation rates is taken as the estimation accuracy.

3 Evaluation Experiment

3.1 Method of Experiment

An evaluation experiment was conducted to show the estimation accuracy of the intellectual concentration state with the estimation method described in Chapter 2. The participants in the experiment were 31 male university students (age: 21.4 ± 1.9). One set of repetition of 1-minute resting time and 5-minute task time was set as one set, and two sets of data for training and test were obtained in total. In each set, each task was conducted for 5 minutes with three kinds of answering methods as described in 2.1. In order to cancel ordering effect, the order of task answer methods was random for each participant.

3.2 Result and Discussion of the Experiment

In the experiment, The data of 6 participants were excluded in the later analysis because of to data loss or so on. The physiological indices data were measured for 2 sets in total, so that estimation accuracy could be calculated when set 1 is as training data, set 2 as test data and vice versa, and the average value of both accuracies was taken as the estimation accuracy.

Table 2. The classification method applied to each participant and the classification performance of the estimator

Participant	Set 1		Set 2	
	Classification method	Correct rate (Number)	Classification method	Correct rate (Number)
p2	Fine KNN	0.88 (21)	Linear Discrimination	0.75 (18)
p3	Linear Discrimination	0.96 (23)	Cubic SVM	1.00 (24)
p4	Linear Discrimination	1.00 (24)	Linear Discrimination	1.00 (24)
p5	Linear Discrimination	1.00 (24)	Decision Tree	0.88 (21)
p6	Middle Gaussian SVM	0.88 (21)	Decision Tree	0.83 (20)
p7	Linear Discrimination	0.75 (18)	Middle Gaussian SVM	0.79 (19)
p8	Quadratic SVM	1.00 (24)	Decision Tree	0.96 (23)
p11	Linear Discrimination	0.83 (20)	Middle Gaussian SVM	0.88 (21)
p12	Decision Tree	0.96 (23)	Linear Discrimination	1.00 (24)
p13	Cubic SVM	0.71 (17)	Cubic SVM	0.83 (20)
p15	Middle Gaussian SVM	1.00 (24)	Linear Discrimination	1.00 (24)
p16	Quadratic SVM	0.96 (23)	Linear SVM	0.92 (22)
p18	Decision Tree	0.79 (19)	Roe Gaussian SVM	0.92 (22)
p19	Linear Discrimination	1.00 (24)	Middle Gaussian SVM	0.92 (22)
p20	Middle Gaussian SVM	1.00 (24)	Linear Discrimination	0.92 (22)
p21	Linear SVM	1.00 (24)	Decision Tree	1.00 (24)
p22	Linear Discrimination	0.96 (23)	Quadratic SVM	0.96 (22)
p23	Linear Discrimination	0.83 (20)	Row Gaussian SVM	0.63 (15)
p24	Quadratic SVM	0.83 (20)	Middle Gaussian SVM	0.71 (17)
p25	Linear Discrimination	1.00 (24)	Linear Discrimination	1.00 (24)
p26	Linear Discrimination	1.00 (24)	Linear Discrimination	1.00 (24)
p27	Middle Gaussian SVM	0.96 (23)	Middle Gaussian SVM	1.00 (24)
p28	Linear Discrimination	0.96 (23)	Linear Discrimination	0.96 (23)
p29	Linear Discrimination	1.00 (24)	Linear Discrimination	0.96 (23)
p30	Linear Discrimination	0.79 (19)	Linear Discrimination	0.71 (17)

Performance Evaluation of Training Data by Classification Learning.

In this section, results and discussions of performance evaluation of estimators are described. Table 2 shows the classification learning methods applied to each participant and the generalization performance of the estimators by the cross validation method. The average of the classification performance of all participants' estimators was 91.1%, and it was confirmed that the concentration states were almost correctly classified by choosing the optimum estimator. Among the 2 sets of measurement data for all 25 participants, the highest classification performance was Linear Discrimination (18 data), and the second was Middle Gaussian SVM (9 data). In the physiological indices field, SVM is often applied as an effective method in classification performance. However, in this study, since the number of training data for the test data used for cross validation

was as many as 23, it can be supposed that some estimation error occurred due to excessive learning when predicting the remaining one piece of test data. On the other hand, with respect to Linear Discrimination, the number of variables necessary to determine the categorical plane is smaller than that of SVM, and the possibility of excessive learning when predicting test data was low. As a result, estimation errors were less likely to occur and it seems that it became a method with high classification performance. It was also confirmed that the performance of Decision Tree and KNN was low compared with the Linear Discrimination unless they are relatively simple methods and their training speed is fast.

Table.3 Estimation Accuracy of test data for each participant

Participant	Correct rate (Correct number)		
	Set1 training, Set2 test	Set2 training, Set1 test	Average
p2	0.50 (12)	0.63 (15)	0.56
p3	0.88 (21)	0.79 (19)	0.83
p4	0.63 (15)	0.58 (14)	0.60
p5	0.46 (11)	0.46 (11)	0.46
p6	0.25 (6)	0.33 (8)	0.29
p7	0.46 (11)	0.38 (9)	0.42
p8	0.50 (12)	0.33 (8)	0.42
p11	0.71 (17)	0.67 (16)	0.69
p12	0.96 (23)	0.79 (19)	0.88
p13	0.21 (5)	0.25 (6)	0.23
p15	0.71 (17)	0.71 (17)	0.71
p16	0.33 (8)	0.50 (12)	0.42
p18	0.50 (12)	0.33 (8)	0.42
p19	0.63 (15)	0.66 (16)	0.65
p20	0.63 (15)	0.42 (10)	0.52
p21	0.83 (20)	0.92 (22)	0.88
p22	0.42 (10)	0.63 (15)	0.52
p23	0.29 (7)	0.29 (7)	0.29
p24	0.58 (14)	0.46 (11)	0.52
p25	0.96 (23)	0.67 (16)	0.81
p26	0.92 (22)	0.83 (20)	0.88
p27	0.54 (13)	0.50 (12)	0.52
p28	0.63 (15)	0.54 (13)	0.58
p29	0.58 (14)	0.79 (19)	0.69
p30	0.58 (14)	0.54 (13)	0.56
Average	0.587	0.560	0.573

Evaluation of Estimation Accuracy with Test Data in Different Time.

Table 3 shows the correct answer rate of intellectual concentration state of each participant. The average accuracy of all valid data was 57.3% which was significantly higher than random estimation ($p < 0.001$). However, there were variations depending on the participants from a correct answer rate of 90% to below 30%. As for the participants who had a high estimation accuracy such as p21, the differences in intellectual concentration state due to the differences in tasks clearly appeared in the difference in physiological responses, and they showed similar responses in both set 1 and set 2. On the other hand, as for the participants with a low estimation accuracy,

there were two types those who tended to have a low performance of the estimator like p7, and those who tended not to estimate correctly even though their performance of the estimator were high like p8. As for the former, due to differences in intellectual concentration state due to differences in tasks did not tend to appear as physiological responses, it was difficult to estimate by classification learning and estimation of test data could not be conducted correctly. On the other hand, with regard to the latter, it was supposed that different physiological responses appeared when measured at different times, so that some drift occurred in the physiological indices, and that an incorrect concentration state was made. Psychological burden such as familiarity with cognitive task, stress or fatigue may have caused the drift when measured at different times.

4 Conclusions

In this study, the authors have developed an estimation method of intellectual concentration state by machine learning of psychological indices. An evaluation experiment was conducted where 31 male university students participated and the measured psychological indices were given to the machine learning models. As the result of the evaluation by cross validation, the model which showed the highest classification performance was Linear Discrimination (18 data) and the second was Middle Gaussian SVM (8 data). On the other hand, the estimation accuracy of the test data which were not used as training data of machine learning was only 57.3% in average. There was considerable difference from the participant who had near 90% correct estimation rate to those who was less than 30%. In the future, it is necessary to explore additional indices which well-reflect their concentration states and are robust to other factors such as difference of time or fatigue in order to improve the estimation accuracy.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP17H01777.

References

1. Tryon, W. W. : Pupillometry: A survey of sources of variation. *Psychophysiology*, 12(1):90–93 (1975).
2. Jorna, P. G. A. M. : Spectral analysis of heart rate and psychological state: A review of its validity as a workload index. *Biological Psychology*, 34(2):237–257 (1992).
3. Eckhard H. Hess, J. M. P. : Pupil size in relation to mental activity during simple problemsolving. *Science*, 143(3611):1190–1192 (1964).
4. Mulder, G. and Mulder, L. J. M. : Information processing and cardiovascular control, psychophysiology. *Psychophysiology*, 18(4):392– 402 (1981).
5. Inc MathWorks. MATLAB, <http://www.mathworks.co.jp/products/matlab.html>. (Accessed May. 2018).
6. Dietterich, T. G. and Bakiri, G. : Solving multiclass learning problems via error-correcting output codes. *Journal of Artificial Intelligence Research*, 2:263–286 (1995).