Proposal of hybrid data mining method for early detection of depression

JUN SAWAMOTO*, KOTA WATANABE**, HIROSHI YAJIMA**
*Regional Cooperative Research Center
Iwate Prefectural University
152-89 Sugo, Takizawa, Iwate 020-0611, JAPAN
**Department of Information Systems and Multimedia Design
Tokyo Denki University
5 Senju-asahicho, Adachi-ku, Tokyo 120-8551, JAPAN
*sawamoto@iwate-pu.ac.jp

Abstract: - In Japan, neuropsychiatric disorders are estimated to contribute to 24.6% of the global burden of disease. The government launched a suicide/depression countermeasure project team, conducted interviews and discussions with experts, analyzed data on suicide such as demographic statistics, and actively examined measures based on the actual situation of suicide. As we can see from these current situations, we think that the problem of suicide in the current situation in Japan is an issue that must be actively addressed and aimed at improvement. In this paper, we focus on "depression" which is counted as the most common cause and motive among suicides. There are many people who have not had a medical examination even if they are suffering from depression, and they tend to be aggravated when they see a problem and have a consultation with a specialist. We consider early detection of disease as an issue.

Key-Words: - Hybrid data mining, Wearable sensors, Detection of depression, Neuropsychiatric disorders, QIDS-J, LOC, SRS-18

1 Introduction

In the Global Health Observatory (GHO) data [1], suicide mortality rate (per 100,000 population), by WHO region, 2016, are Japan 18.5, USA 15.3, France 17.7, Germany 13.6, UK 8.9. The suicide of young people in Japan was serious [2]. Japan's 15 to 34-year-old suicide mortality rate (2014) was 17.8, which was about 2.6 times higher than the accident mortality rate (6.9). In the UK, the number of accidents (12.1) was nearly twice that of suicide (6.6), and six countries in the United States, France, Germany, Italy, and Canada, the accident exceeded suicide.

In Japan, neuropsychiatric disorders are estimated to contribute to 24.6% of the global burden of disease (WHO, 2008) [3]. The government launched a suicide/depression countermeasure project team, conducted interviews and discussions with experts on five occasions, analyzed data on suicide such as demographic statistics, and actively examined measures based on the actual situation of suicide. As we can see from these current situations, we think that the problem of suicide in the current situation in Japan is an issue that must be actively addressed and aimed at improvement.

Depression is a type of mood disorder, which causes a decrease in mental activity, depressive mood, lack of interest, anxiety/irritability, suppression or agitation of psychomotor activity, loss of appetite, insomnia, etc. It is a psychiatric disorder that causes pain and dysfunction. According to the Ministry of Health, Labor and Welfare, depression has the following characteristics.

(1) There are many people who have not had a medical examination even if they are suffering from depression, and they tend to be aggravated when they see a problem and have a consultation with a specialist.
(2) Unlike other diseases, this is a disease in which no clear abnormality can be found in blood tests or imaging tests, and a slight change in the diagnostic criteria leads to a considerable difference in the number of patients.
(3) Depression is an important disease for which early detection is important.

In this paper, we focus on "depression" which is counted as the most common cause and motive among suicides. We consider early detection of disease as an issue.

The rest of this paper is organized as follows: In section 2, we propose our method to cope with the problem of early detection of depression. In section 3, we explain the hybrid information evaluation system for the elderly watching as an example of hybrid data mining. In section 4, we will explain in detail the various data that we collect and use in the

proposed method. In section 5, experiment to clarify the degree of influence and correlation of depression to various data is reported. Finally, the paper is concluded in section 6.

2 Proposed method

Many of the problems of depression were caused by missing the time of early detection. In this study, we focus on the early detection of depression. We try to discover features and the law for estimating the level of the stage of depression without making the user feel the burden, using the data obtained from wearable sensors and various questionnaires without using the direct questionnaire for depression, and verify its reliability and effectiveness.

Specifically, a simple questionnaire (QIDS-J) [4], which measures how much depression a respondent suffers is collected as a reference data. It is said that features such as “cognition”, “emotion” and “action” are particularly intimately related to depression [5]. We collect questionnaire data related to these three features and at the same time collect biological data using wearable sensors.

Multiple data mining (hybrid data mining) combines depression with various elements that make up depression, in particular, questionnaire data related to depression and various data that can be collected on a daily basis using wearable sensors. We discover the relationship between each data by multiple regression analysis to estimate depression and verify the usefulness and its accuracy.

3 Hybrid data mining

Hybrid data mining is not an idea from the technology side, but an idea from the application side. It is defined "a method to extract useful information by combining various mining methods from miscellaneous (hybrid) data" [6]. We consider selecting multiple target data to be mined according to the problem to be solved, applying various mining methods suitable for it, and solving the problem with a combination of them. For example, consider applying the hybrid information evaluation system as shown in Fig. 1 to the elderly watching.

The hybrid information evaluation system will be constructed by combining the following technical elements (Fig. 1).

- Technology combining hybrid data according to the purpose

We develop techniques for collecting subjective information such as text information and voice information, information on living space, time series numerical information from biosensor etc. efficiently according to the purpose and interest of the user and performing preprocessing

- Technology that selects and combines mining methods suited to the purpose for various forms and various information sources

We develop mining methods that can deal with diverse formats and various kinds of information sources and has flexibility to cope with versatility and changes in circumstances. In particular, we focus on mining method of semi-structured data represented by text information and structured data represented by sensor data, and powerful exception discovery technology for detecting situation changes.

- Technology to evaluate and interpret multiple mining results according to purpose

We develop a mining system for specific problem areas and technologies (display method of knowledge, evaluation method, and effective feedback method to users) for making the excavated knowledge useful for users.

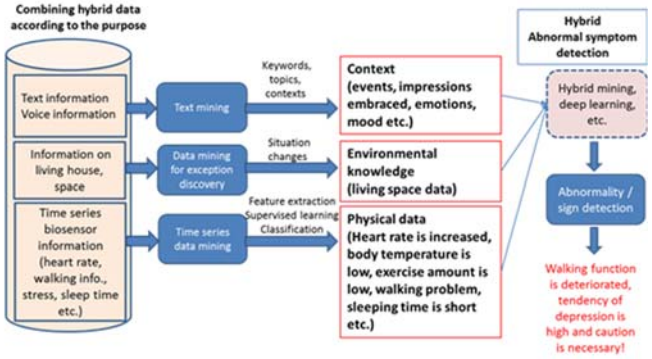

Fig.1 Overall schematic description of hybrid information evaluation type watching system for the elderly to improve QOL

4 Collected data

We will explain in detail the various data that we mentioned to collect and use by the proposed method.

4.1 Simple Depression Symptom Scale (QIDS-J)

The Quick Inventory of Depressive Symptomology – Japanese (QIDS-J) is a self-administered evaluation scale of 16 items, which can evaluate the severity of depression and is also used to evaluate the American Psychiatric Association DSM-IV diagnostic criteria. It has the feature that it corresponds to the diagnostic criteria of major depressive disorder (core depression).

In the QIDS-J scoring method, items for sleep (items 1 to 4), items for appetite/weight (items 6 to 9), and two items for psychomotor status (items 15 and 16), the item with the highest score is selected and scored. For the other items (items 5, 10, 11, 12, 13, and 14), each score is written out. For the severity of depression, sleep, appetite/body weight, psychomotor activity, and other 6 items are evaluated and the total score (0 to 27) of 9 items is evaluated. Since each item corresponds to the symptoms of major depressive disorder, In addition to being used for evaluation screening, it is possible to see the change in depression status by calculating the total points. If it is 6 points or more, there is a possibility of depression and it is considered better to consult with a medical institution.

Table 1 QIDS-J evaluation table

Score	Evaluation	Score	Evaluation
0 ~ 5	Normal	16 ~ 20	Severe
6 ~ 10	Mild	21 ~ 27	Extremely severe
11 ~15	Medium		

4.2 Scale to measure cognition "Locus of Control"

It is a theory on the occurrence of human behavior and the cognition of its reinforcement, which Rotter (1996) [7] advocated based on its own social learning theory. It is a theory of recognizing whether the reinforcement of one's own behavior is acquired by one's own ability or internal force, or as a result of good luck, chance, or bad luck unrelated to one's actions or ability. The former is called internal control and the latter is called external control. In this study, data collection is performed using the scale with the following features created by Kanbara et al based on this LOC concept [8].

(1) Exclude social and political events and measure the general LOC for personal events.
(2) Unify the form of questionnaire items to reflect the LOC for oneself.
(3) As for external assignees, prepare items of various types that can be preliminary prepared, and select item groups with high single factority.
(4) Homogenize subjective value by using events that can be assumed to be of value to everyone, such as "to be happy", and abstract events such as value-neutral, such as "your own life".
(5) A scale that comprehensively reflects various aspects of behavioral strategies such as "I would like to leave it to follow its course" and "I should do it systematically", with a focus on cognition and sense of control.

4.3 Scale to measure emotion "Multifaceted emotional state scale"

A scale created by Terasaki et al. [9] was to measure the subjective state of emotion using Japanese adjectives or words similar to that. The multifaceted emotional state scale has as many fixed and reproducible emotional state factors as possible in comparison with other emotional scale: depression/anxiety, hostility, malaise, activeness, non-activeness, affinity, concentration, and startle were judged to be handy for collecting data on a daily basis, and excellent for collecting "emotional" data used in this study. A total of 40 short questionnaires version were used when collecting the data.

4.4 Scale to measure behavior "Psychological Stress Response Scale (SRS-18)"

Psychological stress responses are emotional, cognitive, and behavioral changes caused by various every day stressors. And the degree of expression of psychological stress responses is a factor that has a great influence on the physical and mental health status. Here, we will use a simple and easy-to-use scale that can measure psychological stress responses that are frequently experienced in a wide range of age groups created by Suzuki et al. [10].

Table 2 Degree of stress reaction (college standard) [11]

The degree of stress reaction		weak	usual	slightly high	high
Total score	Male	~ 7	8 ~ 19	20 ~ 31	32 ~
	Women	~ 10	11 ~ 21	22 ~ 32	33 ~

4.5 Biological sensor data

Biological sensor data is data such as physical characteristics and behavior patterns of people. In this research, wearable device “Smartband 2” by Sony is used as daily data collection. In addition to the measurement of sleep and calories consumed by the accelerometer, recording of the heart rate information by the heart rate sensor is possible. Log information recorded can be easily visualized with the dedicated application.

The data items that can be collected are as follows.

(1) Pulse (minimum pulse, average pulse, maximum pulse)
(2) Stress level (recovered, low, medium, high)

(3) Sleep (time, depth, shallowness)
(4) Activity (steps, walking, running, calorie consumption, exercise time)

5 Experiment

Based on the idea of Hybrid data mining, students' data of "QIDS-J", "psychological stress response scale SRS-18", "Locus of Control", "multifaceted emotional state scale", "biometric data" are collected for one month in the first half of the experiment. After an appropriate processing for each data, multiple regression analysis is applied with the score of depression determined by “QIDS-J” as the objective variable and we clarify the degree of influence and correlation of depression to various data.

In the second half of the experiment, for verifying the usefulness of the mathematical formula of the multiple regression analysis obtained in the first half, we calculate score of depression using four kinds of data, "psychological stress response scale SRS-18", "Locus of Control, "a multifaceted emotional state scale" and "biometric data" and verify the difference from "QIDS-J" score.

5.1 Experiment (first half)

Data was collected for students. The data collected on one day is treated as one case, and the results are shown in Tbale3. The results are summarized for each depression score obtained from “QIDS-J”.

Table 3 Depression score breakdown

Score	Number of data points
1	70
2	6
3	15
4	17
5	24
6	12
7	18
8	12
9	6
10	0
11	2
12 ~	1

Of the data collected, the number excluding defects and outliers was 183. When counted according to the severity of depression, normal:132, mild:48, medium:3 cases, and as a whole, a large number of data collected was not depression.

5.1.1 Cluster Analysis

In order to incorporate biological data including qualitative data into multiple regression analysis, it is necessary to convert it into quantitative data, so perform cluster analysis to classify data by features and assign a cluster number to each data to perform multiple regression. However, since it is not realistic to incorporate the combination pattern of all data items that can be collected from wearable devices into cluster analysis, we pick up data items that are closely related to "depression" as shown in Table 4. The analysis is performed based on this combination .

Table 4 Combination of bio-data

Average pulse
Sleep time
Sleep depth
Stress recovery
Low stress
Middle stress
High stress

In the cluster analysis, the number of clusters is left to the analyst. In this paper, we analyze all four to six, and use the one with the best result.

5.1.2 Multiple Regression Analysis

The multiple regression analysis is an analytical method to calculate the regression equation representing the score of depression as the objective variable y using questionnaire data as multiple explanatory variables xi (i = 1, 2, 3, ...). The importance of each factor to estimate the score of depression is calculated and the equation is obtained.

As shown in Table 5, the results of the multiple regression analysis of the determination coefficient and the adjusted determination coefficient show high correlation. Focusing on the data of cluster number 6, which was excellent in the analysis result, we estimate the score of depression.

Table 5 Results of Multiple Regression Analysis

No. of Clusters	Multiple R-Squared	Adjusted R-Squared
4	0. 9636	0. 9409
5	0. 9619	0. 9401
6	0. 9758	0. 9594

5.1.3 Analysis results

Based on the results of the previous multiple regression analysis, cluster number is set to 6 and bio-sensor data items are average pulse, sleep time and depth, stress (recovery, low, medium, high). Table 6, Figure 2 and 3 show the number of data in each cluster, its scatter chart, radar chart showing features respectively.

Table 6 Number of data in each cluster

Cluster	1	2	3	4	5	6
Number of data	9	19	40	36	28	50

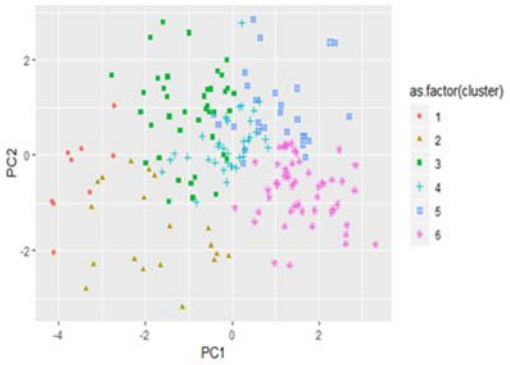

Fig. 2 Scatter chart with 6 clusters

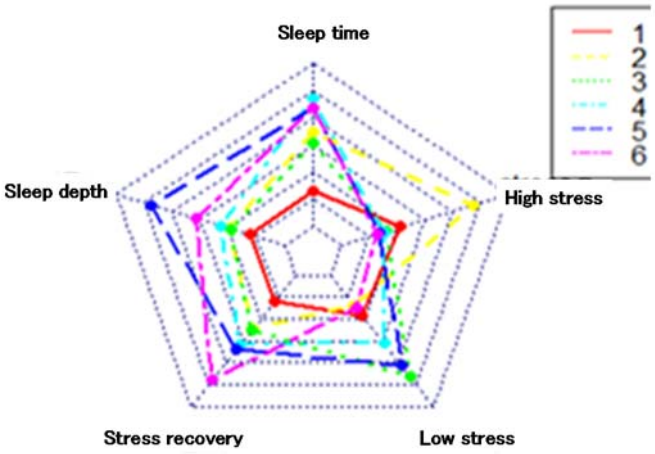

Fig. 3 Rader chart showing features for 6 clusters

The characteristics of each cluster that can be read from the radar chart are as follows.

Cluster1: Almost all items are lower than other clusters: The only item that exceeds clusters other than cluster2 is high stress, which is almost the same value as other clusters.

Cluster 2: Compared to other clusters, the value of high stress is overwhelmingly high. There are no other outstanding features.

Cluster 3: The highest low stress, no other noticeable features.

Clutser 4: The sleep time is the longest, but the recovery of stress which usually depends on the sleep time and low stress value is not low, similar to Cluster 5. But sleep quality is overwhelmingly lower than Cluster 5.

Cluster 5: Sleep depth and time and the low stress are overwhelmingly high. It is thought that they tend to sleep much usually and not to act too much.

Cluster6: Stress recovery is overwhelmingly high. Stress recovery is said to be performed during exercise and sleep. However, because the time of sleep does not differ much from other clusters, it is thought that they tend to exercise well.

Next, the partial regression coefficients obtained from the multiple regression analysis performed by incorporating the results of the previous cluster analysis are shown in Table 7. The standardized partial regression coefficients are calculated from the values obtained by standardizing the explanatory variables and the objective variables. The degree of influence can be calculated because the difference due to the variation of each explanatory variable is removed.

Table 7 Standardized partial regression coefficients

Data items	Coefficients
Intercept	−16. 16
LOC	0. 10
Psychological stress response scale SRS-18	0. 53
Depression / anxiety	0. 36
Hostility	0. 86
Fatigue	1. 40
Activeness	0. 65
Inactiveness	3. 10
Affinity	−0. 66
Concentration	−0. 42
Startle	−3. 95
Affiliate cluster	1. 05

The following equation is used to calculate the score for depression using standardized partial regression coefficients as follows (round off the third decimal place).

$$\begin{aligned}\text{Derpession score} = &-16.16 + 0.10 \times \text{LOC} \\ &+ 0.53 \times \text{Psychological stress response scale} \\ &+ 0.36 \times \text{Depression/anxiety} \\ &+ 0.86 \times \text{Hostility} + 1.40 \times \text{Fatigue} \\ &+ 0.65 \times \text{Activeness} + 3.10 \times \text{Inactiveness} \\ &- 0.66 \times \text{Affinity} - 0.42 \times \text{Concentration} \\ &- 3.95 \times \text{Startle} + 1.05 \times \text{Affiliate cluster}\end{aligned}$$

Data items with high positive values for standardized partial regression coefficients, which are the results of multiple regression analysis, are “fatigue” and “inactiveness” in the multifaceted emotional state scale, “affiliate clusters” that classify biological data by features and the data with a high negative value is the "startle" of the multifaceted emotional state scale, and these data are important factors to estimate the score of depression. And the "affiliate cluster" is directly linked to the biological data, suggesting that the biological data can be used to predict depression.
In addition, depression is considered to be very closely related to the inside of human beings such as emotions. In particular, it can be inferred that people who prefer to spend time alone are more likely to be depressed from the positive emotional factor "inactive" and the negative emotional factor "startle" of the standardized partial regression coefficient.

5.2 Experiment (second half)

In addition to the first half data, we collected 50 new data. In order to verify the usefulness of the formula and the law found in the first half of the experiment. We calculate score of depression using four kinds of data, "psychological stress response scale SRS-18", "Locus of Control", "a multifaceted emotional state scale", "biometrics data", and verify whether there is a difference with the score of "QIDS-J". The number of data for each depression score collected for verification is shown in Table 8. And the average value of collected data, the predicted value derived from analysis, and the relative error are shown in Table 9.

Table 8 Depression score breakdown

Score	Number of data points
1	11
2	5
3	12
4	2
5	6
6	4
7	6
8	3
9	0
10	0
11	1
12 ~	0

Table 9 Verification of the score of "QIDS-J"

Actual value	Expected value	Relative error
(average)	(average)	(average)
3. 92	3. 90	0. 51%

Of the 50 newly collected data, the score for depression (average value) is 3.92, and the predicted value for score for depression (average value) calculated using the formula derived in the first half of the experiment is 3.90. The error was 0.02, and the relative error was 0.51%. There is a large number of data exceeding 10%, but on average it is a good forecast.

6 Conclusion

In this study, with the aim of early detection of depression, we analyzed the characteristics and mathematical expressions that can be detected simply on a daily basis, and verified the usefulness of our approach. The formula had the accuracy to estimate the score of depression, but most of the data we collected this time are "not depression" or "mild depression", so it seems that convergence of the data is likely to improve the accuracy of multiple regression analysis. In addition, although multiple regression analysis is described that the data accuracy improves as the number of data items increases. But, increasing the data blindly leads to making the weight of each data lighter. Efforts such as how to make predictions with few data items without degrading the accuracy of multiple regression analysis, is also an issue.
Based on the above consideration, it is necessary to improve the following points in the future study.

(1) Secure the number of data
(2) Acquisition of data with high scores on depression
(3) Reduce the number of data items in multiple regression analysis and explore how the accuracy changes
(4) Find effective actions for people with high scores of depression, and verify their effects.

References:

[1] Global Health Observatory (GHO) data 2016, WHO https://www.who.int/gho/mental_health/suicide_rates_crude/en/
[2] White paper on Suicide, Ministry of Health, Labor and Welfare, 2018 https://www.mhlw.go.jp/wp/hakusyo/jisatsu/17/index.html
[3] Mental Health Atlas 2011 - Department of Mental Health and Substance Abuse, World Health Organization

https://www.who.int/mental_health/evidence/atlas/profiles/jpn_mh_profile.pdf?ua=1&ua=1

[4] BERNSTEIN, Ira H., et al. Psychometric Properties of the Quick Inventory of Depressive Symptomatology in Adolescents. International journal of methods in psychiatric research, 2010, 19.4: 185.

[5] BEDROSIAN, RICHARD C.; BECK, AARON T. Principles of Cognitive Therapy. Psychotherapy Process: Current Issues and Future Directions, 2012, 127.

[6] SAWAMTO, Jun, et al. Research on Hybrid Information Evaluation Type Watching Technology for the Improvement of QOL of the Elderly. Procedia Computer Science, 2018, 126: 967-975.

[7] Rotter, J. B. 1966 Generalized expectancies for internal vs. external control of reinforcement. Psychological Monographs, 80 (Whole No.609), 1-28.

[8] Kanbara, M., Higuchi, K., & Shimizu, N. (1982). New locus of control scale: Its reliability and validity. Japanese Journal of Educational Psychology, 30, 302-307.）

[9] Terasaki, M., et al. Construction of a multiple mood scale. The Japan Journal of Psychology, 1992, Vol. 62, pp.350~356.

[10] Suzuki, S., Development of a New Psychological Stress Response Scale (SRS-18) and Investigation of the Reliability and the Validity, Japanese Journal of Behavioral Medicine, 1997, 4(1), p. 22-29.

[11] Shinichi, Hironori Kuwata, Yuji Sakano, Isamu Fukui, Makoto Hasegawa "Guide to Using SRS-18 (Stress Response Scale-18)" (2007), Heart net.