Artificial Intelligence-Enabled Identification and Path Solutions for Psychological Development Challenges in Rural Left-Behind Children During the Big Data Era

Xinhua Zhang[1,2*]

[1]College of Preschool Education, Suzhou Early Childhood Education College, Jiang Su, China, 2150002
[2]Rajamangala University of Technology Tawan-Ok, Bangkok, Thailand, 999003

Corresponding author: Xinhua Zhang, zdgxheb125@163.com

Abstract. To explore an effective method to solve the psychological development problems of left-behind children in rural areas, when identifying the psychological problems of left-behind children in rural areas, this paper uses ECG signals to carry out intelligent identification of students' mental health. Moreover, this paper uses the backward selection algorithm to filter the features, remove redundant and irrelevant features, and obtain the optimal feature combination. In addition, this paper extracts linear time domain features, linear frequency domain features, nonlinear features, and moment features to construct an intelligent model to identify the psychological development problems of rural left-behind children and formulate solutions. Finally, this paper combines experiments to evaluate the effect of the model and combines experiments to verify the effectiveness of the method in this paper.

Keywords: big data; rural areas; left-behind children; psychological development; problem identification; Artificial Intelligence.
DOI: https://doi.org/10.14733/cadaps.2024.S24.49-59

1 INTRODUCTION

1.1 Related Work

The narrowness of the research content is mainly manifested in that, from the perspective of the content of the theory of mind research, given the actual needs and the convenience of experimental operation, the current research on the theory of mind mainly focuses on the content defined by the narrow theory of mind (i.e., intentions, Wishes and beliefs and other mental states), and as mentioned above, beliefs are the main research content. Many experiments have used the classic "false belief" understanding task as the only indicator of whether children have acquired theory of mind ability, which undoubtedly narrows the scope of the theory of mind research. Although the results of the meta-analysis in this field in the literature[2] also acknowledge the fact that false belief understanding plays a vital role in the development of children's theory of mind, they recognize that false belief is not the only indicator for judging whether children have acquired theory of mind ability, and Claims that it will do better if you use a sequence of tasks. Literature [6] is directly questioning, "Focusing only on when children understand 'false beliefs' is often misleading because beliefs are just one of many mental states that children understand and use in their day-to-day interactions with others. Children may have A stable understanding of false beliefs, which does not develop until age 4. Still, they travel earlier on the developmental path of a common-sense psychological understanding of people." literature [1] also believes that "there is no evidence that children's meta-representational understanding of intentions is related to or develops with their meta-representational understanding of beliefs and that desires and beliefs are in determining people's understanding. Behavioural is equally important." Literature [5] argues that preschoolers already have a substantial understanding of perception, desire, and intention at an age when they still have only a shaky understanding of misbelief. This is indeed very easy to understand. We know that in a complete mental state, beliefs occupy a relatively high position, but just having beliefs cannot fully and effectively help us speculate and explain the behaviors caused by this mental state. Reference [3] believes that because children and adults often disagree on understanding the instruction language during the dialogue process of the theory of mind test, the children may not understand the real meaning or intention of the experimenter's question, which may underestimate the ability of young children. Because the experimenter's question has a specific scientific purpose, and the child makes a general conversational explanation, this does not mean that the child must have some psychological theory defect. This also reflects the importance of intention and understanding belief from another aspect.

Examining the ecological validity of the theory of mind research has also yielded some exciting findings [8]. A study [9] found that preschool children with more siblings to interact with performed better on false beliefs than those with fewer or no siblings. A study [7] showed that deaf children of parents with normal hearing but not fluent gestures performed worse on false belief tasks than those whose parents were also deaf but proficient in sign language. There are also some unexplainable phenomena from the research on particular groups of children's theory of mind. The study in the literature [10] shows that the theory of mind of autistic children has significant defects. Still, many of these defects only exist in social cognitive processing and some loneliness. The IQ level of children with schizophrenia may be higher than that of ordinary children, so it is difficult to explain with general mental retardation. Literature [11] believes that the above evidence constantly implies that in the current research process of theory of mind, some core mechanisms behind these seemingly "contradictory" experimental phenomena and conclusions may still be unknown. These mechanisms play a vital role in children's theory of mind development. If the tip of the iceberg of these mechanisms can be effectively revealed, the original debates can be quickly resolved [4]

1.2 Objectives

In the context of a significant data era, it is necessary to study the psychological education of rural left-behind children deeply. Society should give more care to left-behind children, create a good learning environment and living environment for them, pay attention to children's psychological development, correctly guide behavioral norms, cultivate healthy psychology, help left-behind children rightly view the problems they encounter, make them eliminate their inferiority complex, and treat life with a positive and sunny attitude.

2 METHODOLOGY

This paper uses the ECG signal to identify the students' mental health intelligently.

2.1 ECG Signal Preprocessing and Feature Extraction

The acquisition of physiological signals in real scenarios requires a portable module. This paper uses Shimmer3.0 to collect ECG data of left-behind children. However, due to the scene and equipment, the data needs to be preprocessed, and then the R wave can be accurately located, the RR interval can be obtained, and various features can be extracted.

Figure 1 shows the preprocessed signal, and Figure 2 shows the location of the R wave after preprocessing.

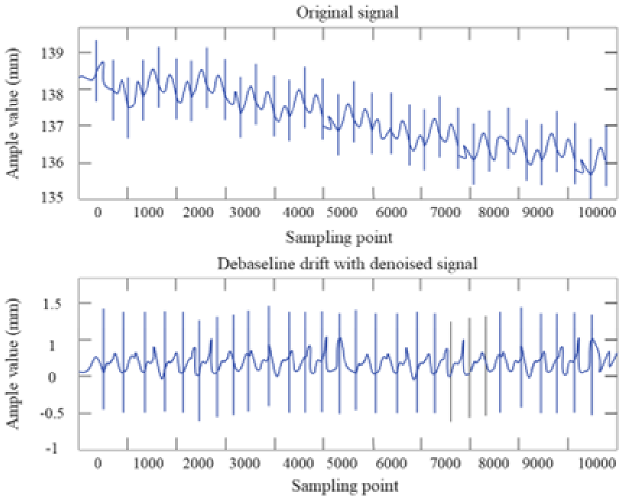

Figure 1: De-baseline drift and De-noising of ECG signals.

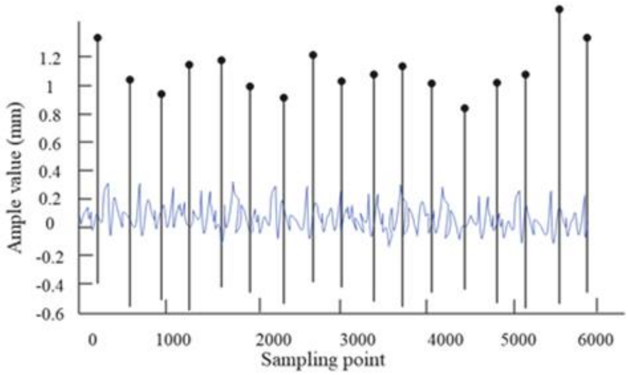

Figure 2: R wave localization.

ECG RR interval sequence, also known as heart rate variability signal, contains valuable information about the cardiovascular system and is commonly used to identify emotions by collecting ECG signals. As shown in Figure 2, the R-wave positioning of this signal is correct.

Since the psychological test data analysis later in this paper is based on the RR interval sequence, for each subject, it is necessary to manually check whether there is an error in the positioning of the R wave and whether there is a signal loss. An example of some signals lost due to equipment reasons is shown in Figure 3. Due to equipment problems, a part of the waveform is lost, which will directly lead to an error in the calculation of the RR interval. The process in this paper is to delete the wrong interval manually.

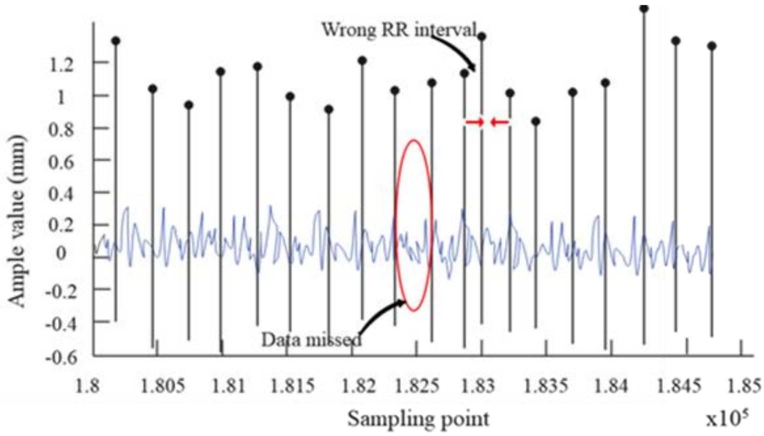

Figure 3: Data processing errors caused by device problems.

2.2 Feature Extraction

In terms of statistical features, this paper extracts the mean value μ_x, standard deviation σ_x, difference value of one interval (2Diff _mean), difference value of second interval (3Diff _mean), root mean square of adjacent differences (RMSSD), beats greater than 50ms (NN50), and the proportion of beats greater than 50ms to total beats (pNN50).

The calculation method of feature extraction is as follows:

1. Mean

$$\mu_x = \frac{1}{N}\sum_{n=1}^{N} X_n \tag{1}$$

2. Standard deviation

$$\sigma_x = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(X_n - \mu_z)^2} \tag{2}$$

3. The difference of one interval

$$2Diff\_mean = \frac{1}{N-2}\sum_{n=1}^{N-2}[X_{n+2} - X_n] \tag{3}$$

4. The difference between two intervals

$$3Diff\_mean = \frac{1}{N-3}\sum_{n=1}^{N-3}[X_{n+3} - X_n] \tag{4}$$

5. Root mean square of adjacent differences

$$\sigma_x = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(X_n - \mu_z)^2} \tag{5}$$

6. NN50: The difference between adjacent RR intervals is more significant than 50ms of heartbeats in all RR intervals.

7. pNN50: NN50 is divided by the total number of RR intervals and multiplied by 100.

Human HRV power spectrum range is generally 0~0.5Hz, and within this range, it can be divided into the following four frequency bands.

This paper extracts LF, HF, and HF/LF in the frequency domain. Figure 4 shows the HRV power density spectra of the two subjects under pressure.

Among them, the high-frequency power reflects the regulating function and activity of the vagus nerve and is related to respiratory sinus arrhythmia. Low-frequency power was associated with baroreflex modulation.

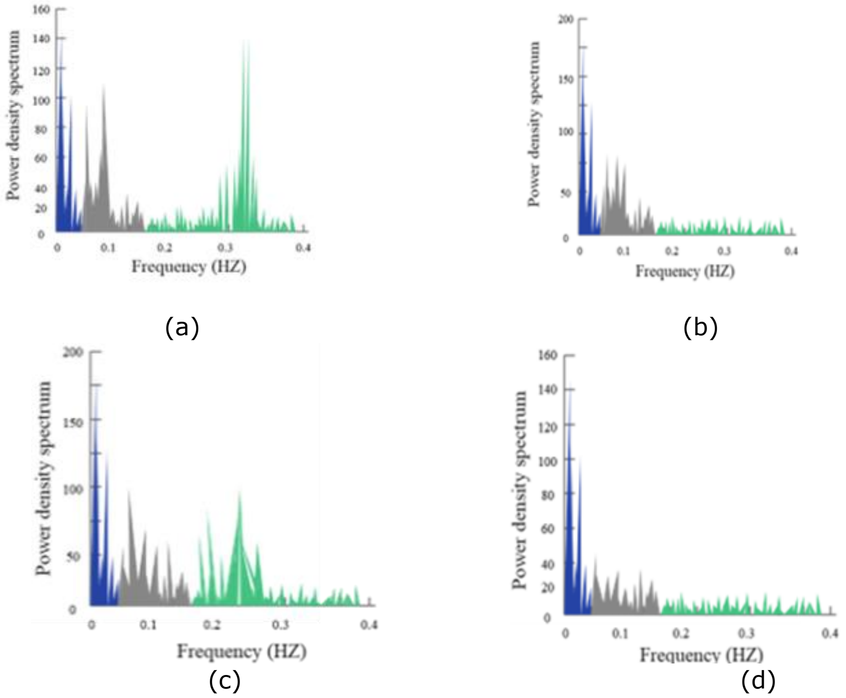

Figure 4:HRV Power Density Spectra of Two States: (a) Subject 1 is in a Calm State; (b) Subject 1 is in a Stressed State; (c) Subject 2 is in a Calm State; (d) Subject is in a Stressed State.

The RR interval sequence is $\{RR_1, RR_2, \ldots\ldots, RR_M\}$ RRM, and M, representing the number of RR intervals. First, the sequence needs to be integrated:

$$f(k) = \sum\left(RR_i - \overline{RR}\right), k = 1,2,\ldots,M \tag{6}$$

The maximum Lyapunov exponent can effectively discriminate whether the time series is chaotic. This paper uses a small data volume method to calculate the foremost Lyapunov exponent, and the time series can be used for subsequent calculations after phase space reconstruction. In the reconstructed phase space, the distance between a trajectory x and its nearest x is:

$$d_j(0) = \min_{X_{\hat{j}}} \left\| X_j - X_{\hat{j}} \right\| \tag{7}$$

Among them, $\|X_j - X_j\|$ is the Euclidean distance, and $d_j(0)$ is the distance between the initial moment of xj and the nearest track.

$$d_j(i) = d_j(0) \times e^{\lambda_1(i \times \Delta t)} \tag{8}$$

We assume that the distance between $d_j(i)$ and $d_j(0)$ satisfies the exponential relationship, namely:

Taking logarithms on both sides of the above equation, we have:

$$ln\, d_j\,(i) = ln\, d_j\,(0) + \lambda_1(i \times \Delta t) \tag{9}$$

for each step i:

$$y(i) = \frac{1}{M \times \Delta t} \sum_{j=1}^{M} ln\, d_j\,(i) \tag{10}$$

In this paper, we use the Legendre moment and Krawtchouk moment as the features of the RR interval sequence to identify and classify stress.

1. Legendre moment

For the one-dimensional discrete signal g(xi) at point M, 1≤i≤M, the expression of the p-order one-dimensional Legendre moment is:

$$L_p = \frac{2p+1}{M-1} \sum_{i=1}^{M} P_p(x_i) d(x_i) \tag{11}$$

Among them, xi = (2i-M-1)/(M-1), and Pp(xi) is a one-dimensional Legendre polynomial of order p:

$$P_p(x) = \frac{1}{2^p} \sum_{j=0}^{p/2} (-1)^j \frac{(2p-2j)!}{j!(p-j)!(p-2j)!} x^{p-2j} \tag{12}$$

Among them, $x \in [-1,1]$. The following recursive formula obtains Legendre:

$$P_{p+1}(x) = \frac{2p+1}{p+1} \tag{13}$$

2. Krawtchouk moment

A Krawtchouk classical polynomial of order n is defined as a hypergeometric function of the form:

$$K_n(x; p, M) = \sum_{k=0}^{M} a_{k,n,p} x^k =_2 G_1 \left(-n, -x; -M; \frac{1}{p}\right) \tag{14}$$

G_1 is a hypergeometric function. To keep the polynomial stable, the weighted Krawtchouk polynomial is introduced:

$$\bar{K}_n(x; p, M) = K_n(x; p, M) \sqrt{\frac{\omega(x;p,M)}{r(n;p,M)}} \tag{15}$$

Among them, $\omega(x; p, M)$ and $r(n; p, M)$ are defined as:

$$\omega(x; p, M) = \binom{M}{x} p^x (1-p)^{M-x} \tag{16}$$

$$r(n; p, M) = (-1)^n \left(\frac{1-p}{p}\right)^n \frac{n\,!}{(-M)} \tag{17}$$

To calculate the weighted Krawtchouk polynomial simply:

$$K_{n-1}(x;p,M)=\left(1+\frac{n-np-x}{pM-pn}\right)K_n(x;p,M)-\frac{n-np}{pM-pn}K_{n-1}(x;p,M) \tag{18}$$

$$\omega(x+1;p,M)=\frac{\omega(x;p,M)p(M-x)}{x+1-p-xp} \tag{19}$$

The initial conditions of the above two formulas are $K_0(x;p,M)=1$, $K_1(x;p,M)=1-x/(Mp), \omega(0;p,M)=(1-p)^M$.

For a one-dimensional discrete signal f(x) of length N, the weighted Krawtchouk moment $\bar{Q}_n$ is defined as:

$$\bar{Q}_n=\sum_{i=1}^{M}\bar{K}_n(i-1;p,M-1)f(x_i)\,, x_i=i-1 \tag{20}$$

Among them, the selection of parameter p is based on which part of the signal recording time is of interest. In this paper, p=0.5 is finally selected.

2.3 Feature Selection

The backward selection algorithm (SBS) used in this paper refers to removing one remaining feature from the feature set each time to obtain the optimal evaluation function. The classifiers used for feature selection by the SBS algorithm in this paper are the BP neural network, KNN classifier, and SVM classifier.

KNN is the most direct method of classifying unknown data. It can be understood through Figure 5 and a simple text introduction. n.

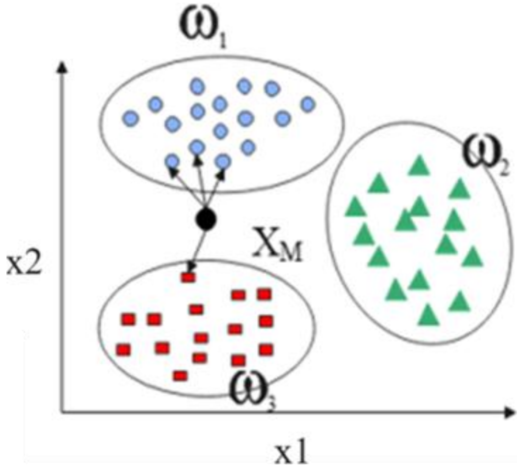

Figure 5: K-NN classification principle.

The integrated view of the left-behind children's psychological theory explanation model is shown in Figure 6.

This article needs to perform various three-category analyses in the follow-up, so the sum of the error rates of the three categories is used to evaluate the quality of the feature combination. The SBS algorithm performs three-category feature selection for resting state, simulated psychological test, and end-of-term psychological test, and the results are counted.

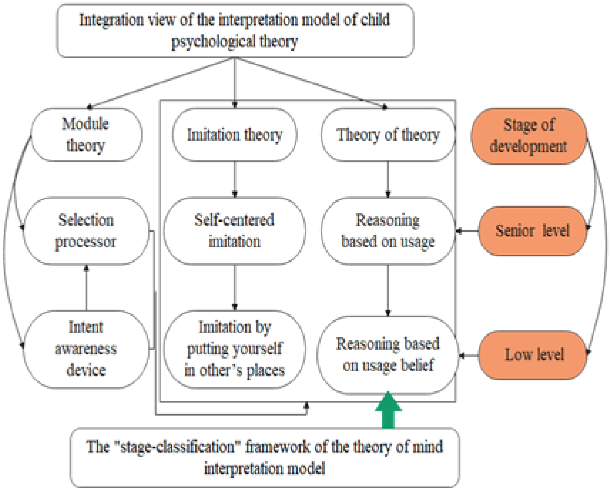

Figure 6: The integrated view of left-behind children's psychological theory explanation model.

Based on constructing an intelligent model, this paper evaluates the effect of identifying rural left-behind children's psychological development problems and analyzing their solutions.

3 RESULTS

The SBS algorithm selects three categories of features for resting state, simulated psychological test, and end-of-term psychological test, and the results are counted. The results are shown in Figure 7. In the recognition results, the minimum error rate obtained by the KNN classifier was 1.035, the minimum error rate obtained by the BP classifier was 0.9213, and the minimum error rate obtained by the SVM classifier was 0.6667. Therefore, SVM classifiers are used for subsequent data classification and identification.

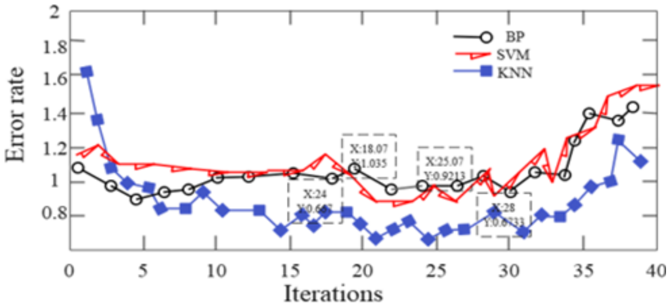

Figure 7: Backward selection results.

This paper evaluates the identification of rural left-behind children's psychological development problems and analyzes their solutions. The results are shown in Tables 1 and 2.

Number	*Psychological problem identification*	*Number*	*Psychological problem identification*	*Number*	*Psychological problem identification*
1	*62.11*	*11*	*73.57*	*21*	*70.39*

2	*65.68*	*12*	*70.01*	*22*	*62.27*
3	*71.04*	*13*	*69.29*	*23*	*71.03*
4	*74.37*	*14*	*72.36*	*24*	*62.19*
5	*73.79*	*15*	*62.00*	*25*	*67.81*
6	*70.17*	*16*	*62.60*	*26*	*65.78*
7	*67.99*	*17*	*75.84*	*27*	*74.03*
8	*73.52*	*18*	*70.63*	*28*	*66.46*
9	*72.60*	*19*	*75.17*	*29*	*61.94*
10	*66.60*	*20*	*70.29*	*30*	*65.04*

Table 1: The identifying effect of psychological development problems in rural left-behind children.

Number	*Psychologically improve decision-making*	*Number*	*Psychologically improve decision-making*	*Number*	*Psychologically improve decision-making*
1	*72.690*	*11*	*65.597*	*21*	*72.174*
2	*65.484*	*12*	*61.133*	*22*	*61.685*
3	*67.681*	*13*	*71.943*	*23*	*67.182*
4	*63.226*	*14*	*60.690*	*24*	*67.668*
5	*67.223*	*15*	*61.283*	*25*	*66.496*
6	*64.264*	*16*	*61.946*	*26*	*63.032*
7	*64.946*	*17*	*60.035*	*27*	*64.303*
8	*70.570*	*18*	*69.026*	*28*	*68.719*
9	*68.361*	*19*	*63.204*	*29*	*63.148*
10	*63.587*	*20*	*64.752*	*30*	*71.686*

Table 2: The analysis effect of the solutions to the psychological development problems of rural left-behind children.

4 DISCUSSION

It can be seen from Figure 6 that when the SVM classifier is used as the discriminant function for backward selection, the sum of the error rates of the three categories is the smallest at the 25th iteration. The features selected in the 25th iteration are Feature 4, Feature 10, Feature 11, Feature 14, Feature 15, Feature 22, Feature 24~Feature 27, Feature 30, Feature 33, Feature 38, a total of 13 features. The selected features are Feature 4, Feature 10, Feature 11, Feature 14, Feature 15, Feature 22, Feature 27, Feature 30, Feature 33, Feature 38, a total of 10 features. The error rate of the 28th iteration is only higher than that of the 24th iteration, and the selected features are three fewer. The eight features chosen by the 28th iteration are used

in the subsequent recognition.

Tables 1 and 2 verify that the method and model proposed in this paper can play an essential role in identifying rural left-behind children's psychological development problems and the path to solution. Moreover, it has a specific role in solving these problems.

5 CONCLUSIONS

Left-behind children in rural areas are a phenomenon that will exist for a long time. To fundamentally solve this problem, we need the attention and coordination of all parties, including families, schools, communities, and the government. This paper uses intelligent algorithms and models based on the background of big data to identify the psychological problems of left-behind children in rural areas. This paper mainly uses ECG signals to determine students' mental health intelligently. At the same time, this paper improves the identification effect of rural left-behind children's psychological development problems through algorithm improvement. It builds an intelligent model to identify rural left-behind children's psychological development problems and formulate solutions. Finally, this paper evaluates the effect of the model based on experiments. The statistical results of the data verify that the method and model proposed in this paper can play an essential role in identifying rural left-behind children's psychological development problems and the path solution. The application of artificial intelligence (AI) to tackle psychological development challenges faced by rural left-behind children during the Big Data era presents an opportunity to provide personalized and targeted support for this vulnerable demographic.

6 RECOMMENDATIONS

This article combines intelligent methods to strengthen the mental health education of left-behind children and help them "learn to be a man, learn to learn, and learn to live," which is the ultimate goal of their education. The model in this paper can play an essential role in identifying rural left-behind children's psychological development problems and path solutions.

Xinhua Zhang, https://orcid.org/0009-0007-9982-522X

ACKNOWLEDGEMENT

The key research topic of Jiangsu education science planning (Research on the collaborative governance mechanism underlying the quality improvement of inclusive private kindergartens from stakeholders' perspective: Taking Jiangsu Province as an example), No. B/2021/01/52.

REFERENCES

[1] Decker, J.; Phillips, T.; Davis, A.; Hehr, A. T.; Ochylski, S.: Big Data in the Veterans Health Administration: A Nursing Informatics Perspective, Journal of Nursing Scholarship, 53(3), 2021, 288-295. https://doi.org/10.1111/jnu.12631

[2] Huang, F.; Ding, H.; Liu, Z.; Wu, P.; Zhu, M.; Li, A.; Zhu, T.: How Fear and Collectivism Influence Public's Preventive Intention Towards COVID-19 Infection: A Study Based on Big Data from the Social Media, BMC Public Health, 20(1), 2020, 1-9. https://doi.org/10.1186/s12889-020-09674-6

[3] Jung, H.; Chung, K.: Social Mining-Based Clustering Process for Big-Data Integration, Journal of Ambient Intelligence and Humanized Computing, 12(1), 2021, 589-600. https://doi.org/10.1007/s12652-020-02042-7

[4] Mao, J. M.; Haupert, M. L.; Smith, E. R.: How Gender Identity and Transgender Status Affect Perceptions of Attractiveness. S, 2019. https://doi.org/10.1177/1948550618783716

[5] Nastro, F. F.; Croce, D.; Schmidt, S.; Basili, R.; Schultze-Lutter, F.: Insideout Project: Using Big Data and Machine Learning for Prevention in Psychiatry, European Psychiatry, 64(S1), 2021, S343-S343. https://doi.org/10.1192/j.eurpsy.2021.919

[6] Park, J.; Kang, U. G.; Lee, Y.: Big Data Decision Analysis of Stress on Adolescent Mental Health, Journal of The Korea Society of Computer and Information, 22(11), 2017, 89-96.
[7] Rouf, K.; Wainwright, T.: Linking Health Justice, Social Justice, and Climate Justice, The Lancet Planetary Health, 4(4), 2020, e131-e132. https://doi.org/10.1016/S2542-5196(20)30083-8
[8] Rudorfer, M. V.: Psychopharmacology in the Age of Big Data: The Promises and Limitations of Electronic Prescription Records, CNS Drugs, 31(5), 2017, 417-419. https://doi.org/10.1007/s40263-017-0419-y
[9] Seo, M.: Amplifying panic and facilitating prevention: Multifaceted Effects of Traditional and Social Media Use During the 2015 MERS Crisis in South Korea, Journalism & Mass Communication Quarterly, 98(1), 2021, 221-240. https://doi.org/10.1177/1077699019857693
[10] Trotman, G. P.; Veldhuijzen van Zanten, J. J.; Davies, J.; Möller, C.; Ginty, A. T.; Williams, S. E.: Associations Between Heart Rate, Perceived Heart Rate, and Anxiety During Acute Psychological Stress, Anxiety, Stress, & Coping, 32(6), 2019, 711-727. https://doi.org/10.1080/10615806.2019.1648794
[11] Tsoy, D.; Tirasawasdichai, T.; Kurpayanidi, K. I.: Role of Social Media in Shaping Public Risk Perception During Covid-19 Pandemic: A Theoretical Review, International Journal of Management Science and Business Administration, 7(2), 2021, 35-41. https://doi.org/10.18775/ijmsba.1849-5664-5419.2014.72.1005