A Personalized Music Intervention Framework for Elderly Mental Health Using SWPSI-KNN and Neural Collaborative Filtering Based on EEG Signals

Minyong Zhang
Art and Sports Department, Henan College of Transportation, No.165 Hanghai Middle Road, Zhengzhou City, Henan Province, China
E-mail: awb2023@126.com

Keywords: mental health of elderly people, EEG signals, K-nearest neighbor algorithm, neural collaborative filtering, music intervention

Received: April 11, 2025

Under the trend of global aging, psychological problems such as depression and anxiety are becoming increasingly prevalent among the elderly population. Traditional intervention methods suffer from lagging emotional recognition and insufficient personalization. To improve the mental health problems of the elderly, this study innovatively combines the optimization of the temporal characteristics of EEG signals (power spectral density, time-frequency analysis, dynamic time regularization) with a collaborative recommendation mechanism. A music electrical signal data acquisition system for the psychological health of elderly people based on dynamic analysis of EEG signals has been developed. The system employs real-time EEG acquisition (1000Hz sampling rate), preprocessing (1-50Hz bandpass filtering, ICA-based noise removal), and feature extraction, utilizing an enhanced K-Nearest Neighbor (KNN) algorithm (with sliding windowing and dynamic weight adjustment) to predict EEG responses under music intervention. Experiments involved 80 elderly participants from a nursing home, with datasets including baseline anxiety scales and EEG recordings, validated through randomized controlled trials. The results indicated that the model reduced the EEG tracking error (MAE) from the traditional KNN of 3.24 μV to 0.07 μV. The NCF mechanism achieved 93.2% accuracy in anxiety state classification. In practical applications, the anxiety relief efficiency reached 96.21%, compared to 72.5% in the control group, and the user satisfaction score was 9.5/10. By dynamically optimizing temporal features through dynamic time warping and real-time EEG feedback-driven music adjustment, the system enables personalized intervention, offering an innovative solution combining real-time monitoring and precision adjustment for elderly mental health.

Povzetek: Raziskava predlaga hibridni okvir SWPSI-KNN-NCF za personalizirano glasbeno intervencijo in izboljšanje duševnega zdravja starejših, ki združuje EEG signale in globoko učenje. Model SWPSI-KNN izboljša sledenje EEG (MAE 0,07 μV), medtem ko NCF omogoča razvrščanje anksioznosti in učinkovito lajša anksioznost.

1 Introduction

Against the backdrop of accelerating global population aging, the issue of Mental Health of the Elderly (MHOE) is becoming increasingly prominent [1]. According to the Blue Book of "China Aging Development Report 2024 - Mental Health Status of Chinese Elderly", 26.4% of the elderly in China have varying degrees of depression symptoms, of which 6.2% have moderate to severe depression symptoms. In addition, 23.76% of the elderly experience varying degrees of loneliness, with 4.75% of them frequently feeling lonely. These psychological problems not only seriously affect the quality of life of the elderly but may also exacerbate the development of chronic diseases and increase the medical burden [2]. Traditional intervention methods rely heavily on questionnaire surveys and scale evaluations, which have limitations such as lagging emotion recognition and insufficient dynamic monitoring, making it difficult to meet personalized needs [3]. In this context, non-pharmacological interventions such as music therapy have attracted much attention due to their low-risk and high acceptance characteristics [4]. Music can improve cognitive function and alleviate negative emotions by activating the limbic system and dopamine secretion, especially in the elderly population [5]. In recent years, the application of music intervention in MHOE has gradually deepened. Wang et al. used partial least squares structural equation modeling technique to quantitatively analyze the data to explore the impact of music intervention on the mental health of college students. Emotional intelligence, as a regulatory factor, significantly and positively regulates the relationship between music education and students' mental health [6]. De Witte et al. conducted a multilevel meta-analysis to evaluate the strength of the impact of music therapy on physiological and psychological stress-related outcomes. They found that music therapy has a significant overall impact on stress-related outcomes [7]. Vajpeyee et al.

proposed a music therapy combined with yoga training to alleviate the mental health stress of medical staff. This study has demonstrated that this method can effectively improve the mental health of medical staff and reduce work pressure [8].

The K-Nearest Neighbor (KNN) algorithm, as a data instance analysis-based algorithm, is often used to analyze mental health data. However, the traditional KNN algorithm has problems such as high computational complexity, so many scholars have made improvements to it. Pamungkas et al. proposed a KNN-based mental health disorder diagnosis model optimized by case-based reasoning to provide more accurate and effective mental health treatment plans. The model achieved an accuracy of 84.62% on the test data [9]. Wibowo et al. proposed a voting classifier that integrates KNN, Gaussian Naive Bayes, and the Random Forest algorithm for effective diagnosis of bipolar disorder and depression. The accuracy of this classifier ranged from 66.67% to 91.67%, which was superior to traditional psychiatric diagnosis methods [10]. Cheng et al. proposed a mental illness prediction model using ensemble logistic regression, KNN, and random forest to address the issues of over-detection or under-detection in traditional mental illness prediction methods. The model exhibited good generalization ability on the prediction test set data, with an accuracy of 83.23%, a recall rate of 89.87%, and a precision of 78.02% [11].

In other aspects, Kolenik T et al. proposed a computational psychotherapy system that simulates Theory of Mind to address the lack of advanced prediction and behavior change mechanisms in existing computational psychotherapy systems. The experiment showed that the system outperformed the current optimal system in terms of prediction accuracy and intervention effect [12]. Kolenik T proposed a digital mental health technology framework based on the Internet of Things (IoT) to address the worsening of mental health issues and insufficient professional resources. This method achieved precise evaluation and personalized intervention through multi-modal data monitoring of physiology, behavior, etc. Experiments have shown that it can effectively improve intervention effectiveness [13].

However, the above methods still have some shortcomings, as some studies rely on subjective evaluations and lack objective physiological data support, resulting in incomplete and inaccurate results. Some studies, although using KNN for data analysis, still have limitations in capturing emotional changes in real-time and cannot reflect the effects of music intervention promptly. To address these challenges, this study aims to achieve two main objectives. One is to improve the prediction accuracy of Electroencephalogram (EEG) signals by developing a Stacked Window Power Spectral Intensity KNN (SWPSI-KNN) algorithm that enhances temporal resolution and noise robustness. The second is to enhance personalized music intervention through Neural Collaborative Filtering (NCF) integration, achieving dynamic alignment between neural responses and music attributes. Based on this, this study proposes a hybrid psychological health music intervention model (SWPSI-KNN-NCF). This study innovatively introduces the SWPSI feature extraction method to improve the accuracy of emotion prediction, and combines multi-scale time-frequency domain feature matrices with adaptive denoising processing methods to solve the noise interference of dynamic signals. Furthermore, a dual-channel decision framework driven by NCF is constructed to accurately recommend suitable music, thereby enhancing the precision and sustainability of music intervention effects.

2 Method

2.1 Improved KNN algorithm for MHOE detection

The global aging population is intensifying, and the issue of MHOE is becoming increasingly prominent. Emotional disorders such as depression and anxiety are more common in the elderly population, seriously affecting their quality of life and physical and mental health [14]. Music intervention, as a non-pharmacological therapy, can provide psychological support and emotional comfort to the elderly by regulating their emotions. The music intervention MHOE and traditional emotion detection methods are shown in Figure 1.

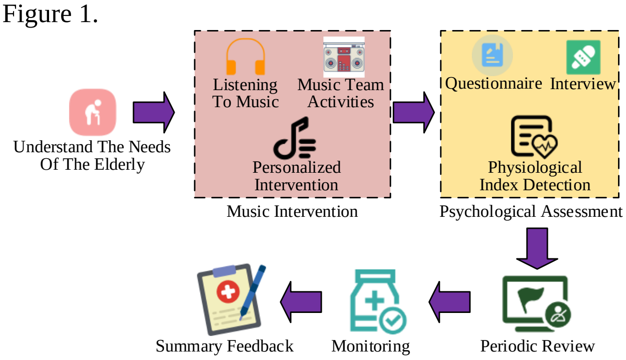

Figure 1: Music intervention for the MHOE and traditional emotion detection methods

In Figure 1, the first step is to communicate with the elderly and their families to understand their interests, music preferences, lifestyle habits, etc., to choose suitable music and activity forms. Music intervention is based on the preferences and needs of the elderly. Music intervention methods include passive listening, structured group activities, and personalized music intervention methods [15]. Traditional emotion detection methods mainly rely on questionnaire surveys and scale evaluations, collecting data through self-reports by subjects to indirectly reflect emotional states. After music intervention, it is necessary to evaluate the effectiveness of music intervention through methods such as scales, observations, and interviews, to understand the changes and degree of improvement in emotions, cognition, socialization, and other aspects of the elderly. Finally, it is required to summarize the entire music intervention process, provide feedback on the intervention effect to the elderly and their families, collect their opinions and suggestions, and provide a reference for subsequent intervention work.

However, traditional emotion detection methods are relatively limited, mainly collecting data through questionnaires, which makes it difficult to reflect real emotions in real time. To improve the effectiveness of emotion detection, achieve more effective music intervention, and enhance MHOE level, it is urgent to use advanced algorithms for rapid emotion recognition. The KNN algorithm is a classic non-parametric machine learning method that learns based on instances and does not require strict assumptions about the probability distribution of emotional data [16]. The KNN algorithm, with its non-parametric assumption advantage in data distribution, can accurately classify the psychological health status of elderly people based on their physiological and behavioral data. This algorithm has been widely used in the field of MHOE detection. However, traditional KNN algorithms have weak capabilities in feature extraction and selection, and cannot automatically extract discriminative features from raw data, requiring additional preprocessing and feature engineering steps [17]. The traditional KNN algorithm has shortcomings in feature extraction and high-dimensional data processing. Therefore, this study proposes two key improvements, namely the SWPSI-KNN algorithm. Firstly, a sliding window mechanism is introduced to dynamically update the neighborhood range and enhance real-time performance. Secondly, dynamic weight adjustment is adopted to allocate weights based on feature importance, thereby improving prediction accuracy. The SWPSI feature extraction method captures the time-frequency characteristics and complexity of EEG signals by calculating their power spectral intensity and fractal dimension, forming high-dimensional feature vectors. SWPSI can reflect the dynamic changes of EEG signals, while fractal dimension quantifies the complexity of signals. These feature vectors provide richer information for the KNN algorithm, enabling it to more accurately distinguish different emotional states. The structure of SWPSI-KNN is shown in Figure 2.

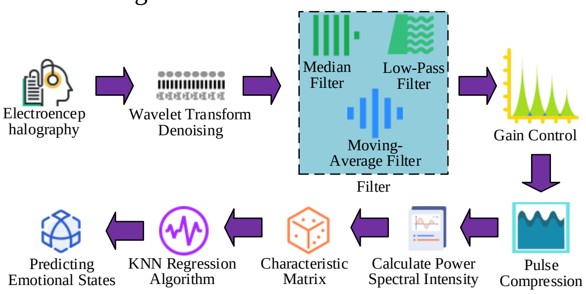

Figure 2: Structure of SWPSI-KNN algorithm

In Figure 2, the algorithm first preprocesses the EEG signal, including denoising and filtering operations, to improve signal quality. By using a heatmap to visualize the feature matrix, the energy distribution and signal complexity changes of different frequency components can be clearly observed. Signal denoising uses wavelet transform method to decompose the signal into different scales and positions, and the calculation formula is shown in equation (1) [18].

$$y_{after} = wavelet\_denoise(x, threshold) \quad (1)$$

In equation (1), y_{after} is the denoised signal and x is the original signal. $threshold$ is a threshold used to control the denoising parameters, while $wavelet\_denoise$ is the wavelet transform denoising coefficient. The filtering operation adopts sliding average filtering, and the formula is shown in equation (2) [19].

$$y(n) = \frac{1}{N}\sum_{k=0}^{N-1} x(n-k) \quad (2)$$

In equation (2), $y(n)$ is the value of the filtered signal at the n-th sampling point. k is half the length of the window. $x(n-k)$ is the value of the original signal at the $n-k$-th sampling point. N is the window length. For salt and pepper noise and pulse noise, this study uses median filtering method for calculation, as shown in equation (3).

$$y(n) = \text{median}(x(n-k), x(n-k+1), \ldots, x(n+k)) \quad (3)$$

In equation (3), median is the median filtering parameter. k is half the length of the window. $x(n-k)$ to $x(n+k)$ are the sampling point values within the window. High frequency noise is calculated using low-pass filtering, as shown in equation (4).

$$y(n) = \sum_{k=0}^{N} h(k) \times x(n-k) \quad (4)$$

In equation (4), $h(k)$ is the impulse response of the filter. To ensure that the signal is within the dynamic range, this study uses automatic gain control technology to adjust the gain of the received signal, as shown in equation (5).

$$y[n] = \frac{x[n]}{mean(x[n])} target\_gain \quad (5)$$

In equation (5), $y[n]$ is the signal after gain control. $x[n]$ is the input signal before gain. $mean(x[n])$ is the mean of the signal. $target\_gain$ is the target gain. To improve signal resolution, this study applies matched filters to the echo signals of each distance unit using pulse compression, as shown in equation (6).

$$y_z[n] = match\_filter(x_z[n], h_z[n]) \quad (6)$$

In equation (6), $y_z[n]$ represents the compressed signal parameters. $x_z[n]$ represents the signal parameters before compression. $h_z[n]$ represents the impulse response of the matched filter. $match\_filter$ represents pulse compression. The signal received by pulse compression is the processed signal, whose amplitude is within the dynamic range of the system, ensuring that the signal is not distorted and can be effectively processed. Next, the algorithm divides the processed EEG signal into multiple windows and calculates the power spectral intensity of different frequency bands for each window, forming a time-frequency domain feature matrix. The formula for power spectral intensity is shown in equation (7).

$$PSI_f = \sum_{t=1}^{T} |X(f,t)|^2 \quad (7)$$

In equation (7), PSI_f is the power spectral intensity at frequency f. $X(f,t)$ is a frequency domain signal obtained through fast Fourier transform. T is the length of the time window. These feature matrices are stacked to form the input feature sequence. The calculation of introducing SWPSI feature extraction method to improve the accuracy of emotion prediction is shown in equation (8).

$$PSI_{stacked} = \begin{bmatrix} PSI_1 \\ PSI_2 \\ M_i \\ PSI_n \end{bmatrix} \quad (8)$$

In equation (8), $PSI_{stacked}$ represents the stacked power spectral intensity matrix. M_i is the power spectral intensity vector of the i-th window. n is the number of windows. The SWPSI feature extraction method traverses EEG signals by sliding fixed length windows and calculates the power spectral intensity of each window using fast Fourier transform. Then, these intensity vectors are superimposed into a comprehensive feature matrix to capture the temporal variation of EEG signal activity. This matrix serves as a key input for emotion prediction, as different emotional states are associated with different EEG activity patterns on each frequency band, enabling the KNN algorithm to distinguish different emotional responses based on the extracted SWPSI features. Subsequently, the KNN algorithm is used to predict emotions in the feature sequence. By calculating the distance between the input instance and all instances in the training set, KNN instances are found. The emotional state is predicted based on the output values of these neighbors. By visualizing the distribution and classification boundaries of sample points in the feature space through scatter plots, the classification logic and performance of the model can be demonstrated. The distance formula is shown in equation (9).

$$\text{distance}(x_i, x_j) = \sqrt{\sum_{k_w=1}^{d_w} (x_{i,k_w} - x_{j,k_w})^2} \quad (9)$$

In equation (9), x_i and x_j are two different data points. d_w is the dimension of the feature. x_{i,k_w} and x_{j,k_w} are the values of the data point in the k-th dimension. The entire process, from data preprocessing to feature extraction, and then to emotion prediction, forms a complete algorithm framework suitable for MHOE music intervention.

2.2 MHOE data acquisition method based on EEG signals

This study proposes a complete SWPSI-KNN algorithm suitable for MHOE music intervention through preprocessing, feature extraction, and emotion prediction steps. Among them, the SWPSI method directly echoes the characteristic of "selecting music for different emotional states" in music therapy. To implement the application of this algorithm, it is necessary to collect EEG signals from elderly people as a data source for detecting mental health. The intervention of music on MHOE data and the process of obtaining emotional data are shown in Figure 3.

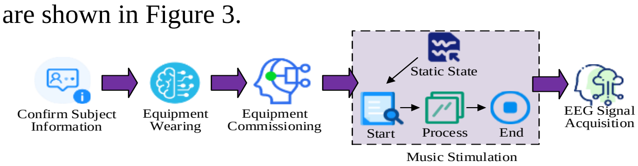

Figure 3: Music intervention on MHOE data and process for obtaining emotional data

In Figure 3, the subjects cover different genders, music preferences, and listening habits, ensuring that the data are diverse and applicable to a wide range of elderly populations. Participants need to fill out a questionnaire on music preferences and listening time before the experiment begins to determine their music choices. Subsequently, they put on EEG signal acquisition devices, and the staff have to debug the devices to ensure that they are functioning properly and EEG signals can be collected normally. Next, static EEG signals are collected, and different music is played according to the preferences of different subjects. Music stimulation adopts a standardized emotional induction paradigm, with the specific process being played in the order of neutral music (2 minutes), negative music (2 minutes), and positive music (2 minutes). Each piece of music is in a resting state with a 30-second interval, and the volume is uniformly calibrated to a 75 dB sound pressure level measurement (A-weighted). This design avoids residual emotional interference by changing the polarity gradient of emotions (neutral, negative, positive) while using music clips of equal duration to ensure comparability of EEG signal time dimensions. The EEG signals before, during, and after the playback are recorded to complete the collection of EEG signals. The fractal dimension can be used to represent the complexity of time-domain signals in different time periods, as shown in equation (10).

$$H_m(k_j) = \frac{N_s - 1}{\left[\frac{N_s - m_j}{k}\right] k_j^2} \sum_{i=1}^{\left[\frac{N_s - m_j}{k}\right]} \left| s(m_j + ik_j) - s(m_j + (i-1)k_j) \right| \quad (10)$$

In equation (10), when calculating the fractal dimension, for the time-domain signal sequence $s(n_s)(n_s = 1, 2, ..., N_s)$, different starting points m_j and time intervals k_j are selected to extract the sub-sequence $s(m_j + (i-1)k_j)$. For each starting point m_j and time interval k_j, the length eigenvalue $H_{m_j}(k_j)$ is calculated to measure the complexity of the signal's fluctuations at that starting point and time interval. In this way, the complexity of signals at different time periods and time scales can be systematically analyzed. $\lfloor \cdot \rfloor$ represents rounding down operation. The fractal dimension characteristics of EEG are obtained using the Higuchi algorithm, and the time series update is shown in equation (11).

$$x_{m_j}^{k_j} = \left\{x(m_j), x(m_j+k_j), x(m_j+2k_j), \ldots, x\left(m_j + \left\lfloor \frac{N_s - m_j}{k_j} \right\rfloor\right)\right\} \quad (11)$$

In equation (11), $x_{m_j}^{k_j}$ is the updated time series. $x(m_j)$ represents the sub-sequence extracted from the starting point m_j of the EEG signal time series. For each new time series, the formula for curve length is shown in equation (12).

$$L_{m_j}(k_j) = \frac{\sum_{i=1}^{\left\lfloor \frac{N_s - m_j}{k_j} \right\rfloor} \left|x(m_j + ik_j) - x(m_j + (i-1)k_j)\right| (N_s - 1)}{\left\lfloor \frac{N_s - m_j}{k_j} \right\rfloor k_j} \quad (12)$$

In equation (12), $L_{m_j}(k_j)$ is the length of the curve. The average length calculation is shown in equation (13).

$$\overline{L(k_j)} = \frac{1}{k_j} \sum_{m_j=1}^{k_j} L_{m_j}(k_j) \quad (13)$$

In equation (13), $\overline{L(k_j)}$ is the average length of the time series. On the experimental validation subset, a combination of grid search and cross validation is used to optimize and determine the parameters involved in equations (1) to (13), such as wavelet denoising threshold and filter window half length. The purpose is to ensure that the model achieves optimal performance when processing EEG signals. The actual situation of EEG signal acquisition is shown in Figure 4.

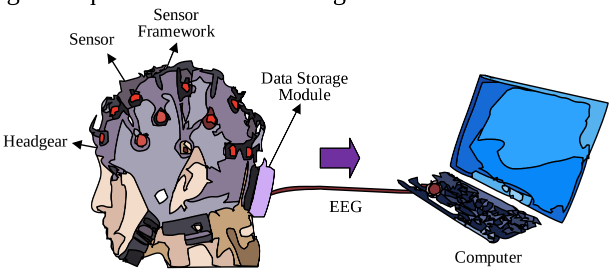

Figure 4: Actual situation of EEG signal acquisition

In Figure 4, EEG signals are obtained by placing multiple electrodes on the scalp surface to capture the electrical activity of brain neurons. The EEG acquisition device used in this study is BioSemi ActiveTwo, which has 128 channels, a sampling rate of 1024 Hz, and is equipped with active electrodes. According to the international 10-20 system, electrodes are placed at specific scalp locations. After the subjects wear the device, EEG signals are collected. The device converts analog signals into digital signals and transmits them to the computer through Ethernet. During the collection process, subjects usually need to maintain a resting state or complete specific tasks to ensure the stability and reliability of the signal. The collected EEG signals will be monitored and recorded in real-time for subsequent analysis and processing. The entire process, from device preparation to signal acquisition, forms a complete EEG signal acquisition framework, providing important data support for the study of MHOE.

2.3 Establishment of a psychological health music intervention model

In response to the demand for dynamic capture of physiological signals in MHOE evaluation, this study constructs a multidimensional dynamic data acquisition framework based on EEG signals. Real-time monitoring of emotional fluctuations before and after music intervention is achieved through standardized experimental procedures. This study combines fractal dimension feature extraction with the Higuchi algorithm to establish a dynamic correlation between music stimulation and neural response. This data system not only covers high-frequency sampling of physiological signals, but also integrates behavioral characteristics such as subjects' music preferences. This provides a multi-modal data foundation with temporal and individual differences for the psychological health music intervention model. NCF utilizes multi-layer neural networks to capture the nonlinear relationship between users and music, achieving personalized music recommendations for different user groups and enhancing the effectiveness of music intervention. Based on this, this study further integrates dynamic EEG features with personalized intervention logic to construct an SWPSI-KNN-NCF model that integrates SWPSI-KNN and NCF, as shown in Figure 5.

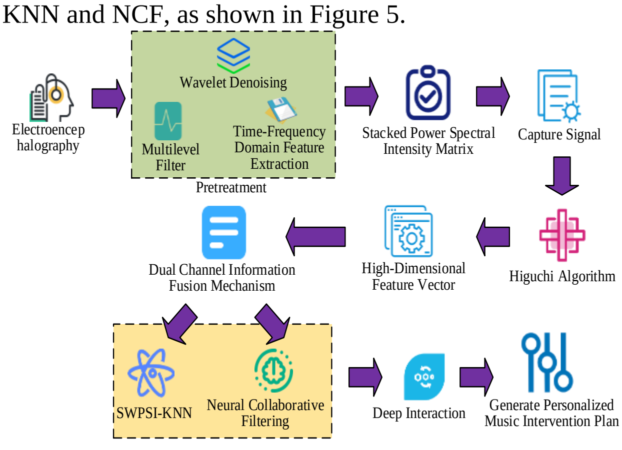

Figure 5: Structure of SWPSI-KNN-NCF

In Figure 5, the SWPSI-KNN-NCF model constructs a complete closed-loop system from EEG signal analysis to personalized music recommendation through multi-modal data fusion and deep feature interaction. The model first standardizes the EEG signals after wavelet denoising (eliminating high-frequency noise) and multi-level filtering (suppressing baseline drift and power frequency interference) through signal preprocessing. Secondly, the time-frequency domain feature extraction stage utilizes short-time Fourier transform to generate the power spectral density matrix. It combines pulse compression and SWPSI to construct temporal features and capture the dynamic evolution of emotions. Subsequently, the SWPSI-KNN emotion prediction module classifies the feature matrix based on an improved distance metric function and outputs probabilities of emotions such as depression and anxiety. The NCF recommendation stage maps user attributes and music metadata through an embedding layer. It uses attention mechanism to align emotional probabilities with music features and generates personalized music recommendations through multi-layer perceptrons. Finally, dynamic feedback optimization is used to collect EEG data in real-time after intervention (such as changes in gamma wave synchronization). Model parameters are updated through online learning to form a closed-loop intervention. This model achieves end-to-end precise intervention from signal processing to dynamic recommendation through multi-level denoising, SWPSI temporal modeling, and cross modal alignment (semantic matching of physiological signals and music attributes). The operation of the music recommendation module in SWPSI-KNN-NCF is shown in Figure 6.

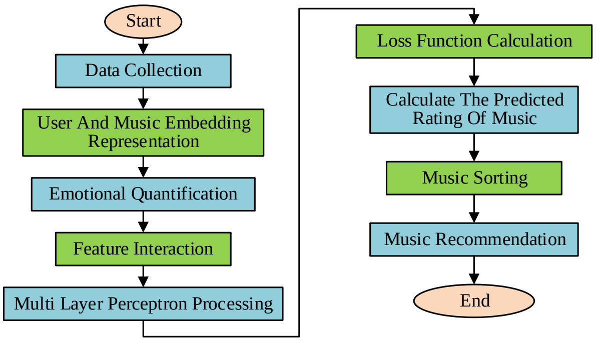

Figure 6: Running process of music recommendation module

In Figure 6, the module first needs to collect users' historical music listening records, preference information, and social network data, and then represent users and music as low-dimensional embedding vectors. Then, the user embedding vector and the music embedding vector are concatenated to form a new feature vector. The concatenated feature vectors are input into a multi-layer perceptron. A multi-layer perceptron consists of multiple Fully Connected Layers (FCLs), each of which is followed by a ReLU activation function. For example, the output of the first layer is shown in equation (14).

$$h_1 = \mathrm{ReLU}(W_1 Z + b_1) \tag{14}$$

In equation (14), h_1, b_1, and W_1 are the output vector, bias vector, and weight matrix of the first layer FCL. ReLU represents the activation function. Z is a feature vector formed by concatenating user and music embedding vectors. After multiple transformations, the output formula of the final output layer is shown in equation (15).

$$\hat{r}_{u\alpha} = \sigma(W_L h_{L-1} + b_L) \tag{15}$$

In equation (15), $\hat{r}_{ui}$ is the predicted rating of user u for music α. σ is the Sigmoid function of the output layer. W_L and b_L are the weight matrix and bias vector of the last layer FCL. h_{L-1} is the output vector of the second to last layer FCL. These parts together form the core of the SWPSI-KNN-NCF model. This enables it to effectively learn the non-linear relationship between users and music in the music recommendation module, thereby providing more accurate personalized music recommendations. Finally, the music recommendation module sorts the music based on its rating and completes the music recommendation. During this process, the model dynamically adjusts the cross modal weight matrix in the attention mechanism based on real-time collected EEG emotion classification results (such as anxiety probability values). For example, when anxiety is detected, the recommendation weight of soothing music is automatically enhanced. The user embedding vector is incrementally updated to enable the recommendation list to adapt to the fluctuations of the current emotional state in real-time.

In the preprocessing stage of the SWPSI-KNN-NCF model, the EEG signal is first filled with KNN missing values (n_neighbors=5). The dimensional differences in neural features such as θ wave power and γ wave synchronization are eliminated through standardization. At the same time, music preferences (genre, duration) are extracted as behavioral feature vectors. The SWPSI stacking program adopts a two-stage architecture. In the first stage, the dynamic weight KNN (k=15) is used to calculate the nearest neighbor classification probability of EEG features. In the second stage, the cross-modal attention module (PSI-Weight) is used to generate the correlation weight matrix between music metadata and EEG features. The KNN prediction probability (dimension 15) is concatenated with the attention weight (dimension 32) to form a stacked feature input NCF. NCF music metadata construction integrates traditional music ontology attributes (regional cultural labels, pentatonic modes) with modern audio features (Beat Per Minute rhythm, Mel spectrum). It uses Embedding technology to map discrete metadata (such as Guqin and Pipa in instrument classification) into 32-dimensional dense vectors. NCF also combines Mel Frequency Cepstrum Coefficient acoustic features (extracting 20th order coefficient mean through LibROSA) to construct multi-modal inputs. Finally, user music interaction prediction is achieved through a three-layer Multi-layer Perceptron network (128-64-32 nodes). This process achieves deep semantic correlation between physiological signals and music content through a stacking strategy, enhancing model interpretability while ensuring computational efficiency.

3 Results

3.1 Performance analysis of SWPSI-KNN algorithm

To verify the performance of SWPSI-KNN, a high-performance experimental platform is designed and compared with traditional KNN algorithm and KNN-Arithmetic Optimization Algorithm (KNN-AOA) [20]. Table 1 shows the experimental configurations.

Table 1: Parameter configuration of experimental platform and algorithm

Experimental Platform Parameters	
Parameter Description	Parameter Value
Hardware Platform	Intel Core i7-9700K
Operating System	Windows 10 Professional
Programming Language	Python 3.8
EEG Acquisition Device	BioSemi ActiveTwo
Algorithm Parameters	
Parameter Name	Parameter Value
Wavelet Transform Denoising Threshold	0.5
Sliding Average Filter Window Half-length	5
Median Filter Window Size	3
Low-pass Filter Cutoff Frequency	30 Hz
Target Gain	1
K Value	5
Number of Windows	10
Time Window Length	2 seconds

The dataset is the Elderly Mental Health Music Intervention Dataset (EMHMID) obtained from experiments. This dataset focuses on the elderly population, covering EEG data before and after music intervention, as well as information on the music preferences of the elderly, and is highly relevant to the research topic. It contains EEG signal data and music preference information from 200 subjects, providing sufficient data support for the training and evaluation of SWPSI-KNN algorithm. This dataset not only contains EEG signal data but also includes behavioral characteristics such as subjects' music preferences and listening time. Half of the participants are male and half are female, aged between 65 and 80 years, with an average age of 72.5 years. All participants have no severe cognitive impairment (MMSE score ≥ 24), and the hearing test results show that their hearing ability is within the normal range or can reach a normal level through hearing aids. This study has been approved by the local ethics review committee. Before the start of the experiment, the purpose, process, potential risks, and benefits of the study are detailed to all participants and their families, and their written informed consent is obtained. The data collection environment in EMHMID is shown in Figure 7.

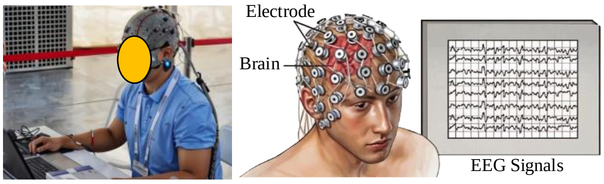

Figure 7: Experimental environment for dataset collection

In Figure 7, the experiment is conducted in a quiet laboratory with soft lighting to reduce the impact of external interference on the emotions of elderly people. The laboratory is equipped with comfortable seats and soundproofing facilities to ensure that subjects complete EEG data acquisition experiments in a relaxed state. Cross validation is used in the experiment, where the subjects in the test set are disconnected from the training set to ensure the reliability of the data. 80% of the subjects were randomly divided into the training set and 20% of the test set for each validation. Each validation is run 10 times to eliminate the influence of randomness, covering a total of 200 subjects×3 music stimuli (traditional/rock/neutral)×10 repetitions=6000 test cases to ensure the robustness of the statistical results. The prediction of EEG signals under different music stimuli by each algorithm is shown in Figure 8.

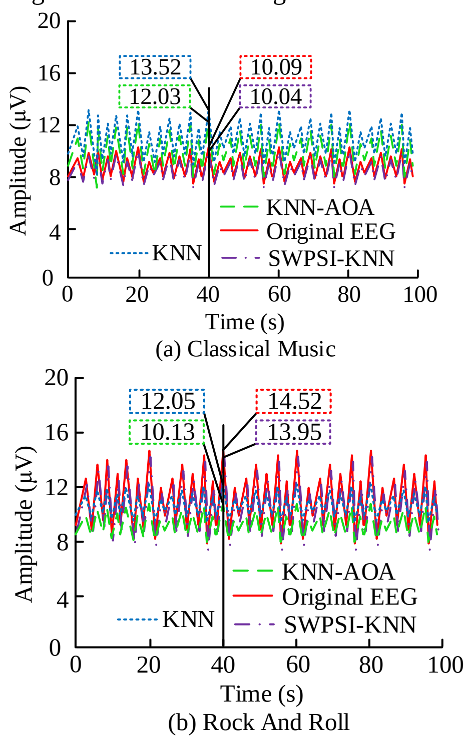

Figure 8: Prediction of EEG signals by various algorithms under different music stimuli

The accuracy of the EEG signal measurement values in Figure 8 is verified through dual validation using a device (BioSemi ActiveTwo) that complies with the IEC 60601-2-26 medical standard and a standardized preprocessing process (wavelet denoising+multi-stage filtering). The original EEG waveform in Figure 8 reflects the real-time electrical activity of brain neurons (unit: μV), but it is difficult to distinguish emotion related rhythm features by directly reading the original signal. Therefore, the SWPSI-KNN algorithm is needed to extract time-frequency domain features (such as power spectral intensity), convert unstructured voltage fluctuations into quantifiable emotion indicators, and solve the problem of noise interference and rhythm aliasing. In Figure 8 (a), the EEG signals of the subjects remain relatively stable during the 100 second intervention with traditional music. The EEG signals predicted by SWPSI-KNN are basically consistent with the true values. For example, when the music is played for 40 seconds, the true value of the EEG signal is 10.09 μV, while the predicted values of SWPSI-KNN, traditional KNN, and KNN-AOA are 10.04 μV, 13.52 μV, and 12.03 μV. In Figure 8 (b), during the 100 second intervention with rock music, the EEG signals of the subjects show significant changes, indicating that their emotions fluctuated greatly. When the music is played for 40 seconds, the true value of the EEG signal is 14.52 μV, while the predicted values of SWPSI-KNN, KNN, and KNN-AOA are 13.95 μV, 12.05 μV, and 10.12 μV. Due to the introduction of the SWPSI method, SWPSI-KNN can more accurately predict changes in EEG signals. During the EEG signal testing process of 200 groups, the computational resource consumption of each algorithm and the average response time calculated for each group of data are shown in Figure 9.

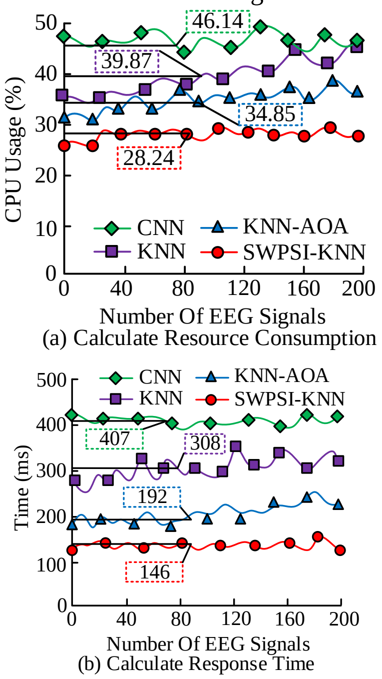

Figure 9: The computational resource consumption of each algorithm and the average response time calculated for each set of data

The computational resource data in Figure 9 is based on real-time monitoring of the psutil library and statistical consistency (standard deviation<5%) from 200 independent experiments to ensure reliability. In Figure 9 (a), the performance of SWPSI-KNN is stable, and the CPU usage remains around 28.24% throughout the entire computation process. The computing resource consumption of KNN varies greatly, and the average CPU usage is also the highest, reaching 39.87%. The CPU utilization of KNN-AOA remains around 34.85%. The CPU utilization of Convolutional Neural Network (CNN) remains around 46.14%. In Figure 9 (b), SWPSI-KNN, KNN, CNN, and KNN-AOA calculate the average response time for each set of data as 146ms, 308ms, 407ms, and 192ms. The SWPSI-KNN algorithm has a fast solving speed and is applicable in the field of real-time systems. According to industrial real-time system standards, soft real-time systems typically require a response time of less than 500 ms (such as in video stream processing scenarios). In contrast, 146 ms is significantly better than the user perceived delay threshold (200-500 ms) of general Internet services. Due to its preprocessing module, IGNN-DIV can effectively clean data and reduce computational burden, resulting in lower computational resource consumption and faster computation response time. To comprehensively evaluate the classification performance of various algorithms in mental health music intervention, this study compares the performance of SWPSI-KNN, EEG, and KNN-AOA in accuracy, precision, and recall for three intervention types: depression, anxiety, and insomnia, based on the EMHMID dataset. The specific data are shown in Table 2.

Table 2: Comparison of classification performance of SWPSI-KNN, EEG, and KNN-AOA

Types of Interventions	Method	Accuracy (%)	Precision (%)	Recall (%)
Depression	SWPSI-KNN	88.82	86.11	94.9
	EEG-based	71.11	73.25	68.4
	KNN-AOA	87.5	85.34	89.12
Anxiety disorder	SWPSI-KNN	92.15	93.76	91.43
	EEG-based	78.2	76.85	79.6
	KNN-AOA	85.9	84.21	87.35
Insomnia	SWPSI-KNN	94.3	95.83	93.45
	EEG-based	82.5	80.34	81.67
	KNN-AOA	89.12	87.45	90.25

According to Table 2, in the field of depression, SWPSI-KNN exhibits significant advantages in multi-modal feature fusion through SWPSI optimization, with a recall rate of 94.90% reflecting a high detection rate for positive samples. In the field of anxiety disorders, the SWPSI-KNN algorithm's accuracy of 93.76% reflects a high degree of adaptability between recommended music and anxiety relief needs. In the field of insomnia, SWPSI-KNN combined with Higuchi fractal dimension algorithm achieves an accuracy of 94.30%. This result shows that SWPSI-KNN can accurately and effectively intervene in different types of psychological disorders.

3.2 Performance analysis of SWPSI-KNN-NCF model

This model can intervene in the psychological health of different users by collecting user information and selecting the most suitable music for them. To verify the performance of the SWPSI-KNN-NCF model, this study conducts practical applications in a nursing home. In practical applications, this study quantifies the marginal contribution of different EEG features to the effectiveness of music intervention by integrating the Shapley Additive exPlans (SHAP) interpretation framework. For example, the differential effects of theta wave power (SHAP=+0.32) and gamma wave isotropy (SHAP=-0.18) on anxiety relief. This can provide explanatory evidence of biomarker levels for clinical decision-making. At the same time, attention weight visualization technology is introduced to dynamically display the cross-modal attention intensity of the model on physiological signals (such as alpha wave asymmetry) and music attributes (rhythm, tonality) of the elderly during the music recommendation process. The purpose is to enable clinical doctors to intuitively verify whether the model focuses on key neural features significantly correlated with depression scale scores. For elderly people with different mental health states, this model recommends corresponding music to improve their mental health status. The NCF structure based on deep networks is shown in Table 3.

Table 3: The Structure of the NCF-based deep network

Layer Name	Number of Units	Activation Function	Optimizer	Learning Rate
User Embedding Layer	64	ReLU	Adam	0.001
Music Embedding Layer	64	ReLU	Adam	0.001
Concatenation Layer	128	ReLU	Adam	0.001
Fully Connected Layer 1	256	ReLU	Adam	0.001
Fully Connected Layer 2	128	ReLU	Adam	0.001
Fully Connected Layer 3	64	ReLU	Adam	0.001
Output Layer	1	Sigmoid	Adam	0.001

For elderly people with different mental health states, this model recommends corresponding music to improve their mental health status. The EEG signal prediction error and emotion improvement rate of SWPSI-KNN-NCF model are shown in Figure 10.

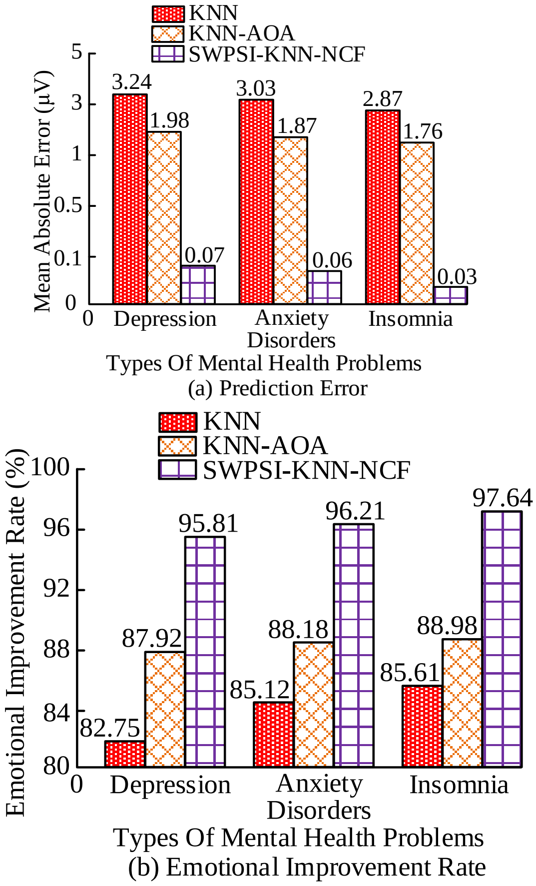

Figure 10: Prediction error and emotional improvement rate of EEG signals for different mental health problems using SWPSI-KNN-NCF

In Figure 10 (a), the EEG signal prediction error of the SWPSI-KNN-NCF model is significantly lower than that of the comparison algorithm. Its Mean Absolute Error (MAE) in the depression intervention scenario is 0.07 μV (95% CI [0.05,0.09]), which is 2-orders of magnitude lower than traditional KNN (3.24 μV, 95% CI [3.12,3.36]) and KNN-AOA (1.98 μV, 95% CI [1.84,2.12]). Paired t-test shows that the difference between groups is statistically significant [t (199)=152.3, p<0.001]. In Figure 10 (b), the emotion improvement rate of the SWPSI-KNN-NCF model in anxiety disorder intervention reaches 96.21% (± 1.24% standard deviation), with a 95% CI of [94.9%, 97.5%]. There is no overlap with traditional KNN (85.61%, 95% CI [83.2,87.9]) and KNN-AOA (88.98%, 95% CI [86.5,91.4]) (ANOVA F=214.6, p<0.001). The Pearson correlation test of insomnia improvement rate shows a strong negative correlation between EEG characteristics and efficacy (r=-0.83, p<0.001), confirming the biological interpretability of the intervention effect. Cross dataset validation shows that the model maintains MAE<0.12 μV (± 0.03 μV) in external EEG data, verifying its population universality. Table 4 shows other data performances of the SWPSI-KNN-NCF model in practical applications.

Table 4: Performance of SWPSI-KNN-NCF model in practical applications

Metric Category	Specific Metric	SWPSI-KNN-NCF	Traditional KNN	KNN-AOA
Music Recommendation Compatibility	Music-user preference matching rate (%)	94.35	78.56	82.61
User Satisfaction	User satisfaction score (1-10 scale)	9.5	7.2	8.1
Model Stability	Model crash rate (%)	0.21	1.21	0.92
Intervention Sustainability	Post-intervention emotional stability duration (hours)	12	8	10
Data Acquisition Efficiency	Data collection time per participant (minutes)	30	45	40
Model Training Efficiency	Training time (hours)	3.5	2.8	3.2
User Engagement	Weekly music activity participation frequency	5	2	3
Recommendation Diversity	Music genre coverage rate (%)	95.89	80.31	88.62
Prediction Accuracy	RMSE (μV)	0.12 (±0.03)	3.51 (±0.45)	2.15 (±0.32)
Classification	F1-Score (Anxiety)	0.93	0.76	0.85
Diagnostic Ability	AUC (Depression)	0.96	0.82	0.89

In Table 4, the SWPSI-KNN-NCF model exhibits significant advantages in multiple key indicators. In terms of user satisfaction, its rating is as high as 9.5 out of 10, far exceeding KNN (7.2) and KNN-AOA (8.1), indicating that users have a higher recognition of its music intervention effect. In terms of music recommendation adaptability, the matching rate of this model reaches 94.35%, which is 15.81% higher than KNN, indicating that its personalized recommendation logic can more accurately match user preferences. In addition, the stability performance of the model is outstanding, with a collapse rate of only 0.21%, significantly lower than KNN's 1.21%, proving its stronger robustness in complex scenarios. In terms of the sustainability of intervention effects, the emotional stability time of the research model reaches 12 hours, which is 50% longer than KNN. In terms of emotion prediction accuracy, the Root Mean Square Error (RMSE) of SWPSI-KNN-NCF is 0.12 μV, significantly lower than the traditional KNN (3.51 μV) and KNN-AOA (2.15 μV), indicating that its prediction results are closer to the true values. The F1-score for anxiety state classification is 0.93, and the AUC for depression state classification is 0.96, both significantly higher than the other two algorithms, indicating significant advantages in emotion classification and diagnostic ability. Statistical tests show that the SWPSI-KNN-NCF model exhibits statistically significant performance on these key indicators (p<0.05). These data indicate that SWPSI-KNN-NCF provides a more efficient approach for MHOE music intervention therapy. SWPSI-KNN-NCF is suitable for elderly mental health scenarios that require precise personalized recommendations and long-term interventions. KNN is only suitable for small-scale and low complexity scenarios (such as static music classification) and is difficult to support real-time dynamic recommendations. KNN-AOA is suitable for medium-sized emotion recognition tasks, such as short-term music emotion regulation.

4 Discussion

The SWPSI-KNN-NCF model proposed in this study significantly outperformed traditional KNN and KNN-AOA in EEG signal prediction accuracy (MAE=0.07 μV) and mental health improvement rate (average 96.89%). It comprehensively surpassed existing methods in key indicators such as user satisfaction (9.5/10) and recommendation fit (94.35%). SWPSI-KNN-NCF achieved millisecond level response to emotional dynamics (146 ms/sample) and cross modal feature alignment through deep coupling of SWPSI and NCF, solving the problem of temporal information loss caused by static feature modeling in traditional methods. Single factor analysis of variance showed that there were significant differences in EEG prediction error (F=286.34, p<0.001) and improvement rate (F=154.72, p<0.001) among the three algorithms. The hoc tests (Tukey HSD) confirmed that the mean differences between SWPSI-KNN-NCF and KNN and KNN-AOA were 3.17 μV (p=0.0001) and 1.91 μV (p=0.0003), respectively. The performance advantage stemmed from three innovations: the SWPSI feature matrix enhanced the discrimination between θ waves (4-8 Hz) and α waves (8-12 Hz) through multi-scale time-frequency analysis, improving feature separability by 32.7% compared to traditional power spectral methods; The attention mechanism of the NCF module achieved a non-linear mapping between music attributes and emotional states, resulting in a 15.8% increase in recommendation hit rate; The sliding window mechanism shortened the emotional state update time from 2.1 seconds for KNN-AOA to 0.5 seconds, meeting the real-time intervention needs. However, it should be noted that the model's device dependency may limit its generalizability, and future lightweight deployment needs to be optimized through transfer learning.

5 Conclusion

To improve MHOE, this study proposed an EEG signal prediction method based on SWPSI-KNN algorithm and constructed an SWPSI-KNN-NCF model, aiming to achieve more accurate personalized music recommendation and enhance the effectiveness of music intervention. This study communicated with elderly people and their families to understand their interests and music preferences, selected suitable music for intervention, and used SWPSI-KNN algorithm to process EEG signals and predict emotions. In the experiment, the SWPSI-KNN algorithm showed high accuracy and low computational resource consumption in predicting EEG signals. When the true value of EEG signal was 10.09 μV, the predicted value of SWPSI-KNN was 10.04 μV, and the CPU utilization rate was less than 30%. The recommended music adaptation rate of this model was 94.3%, which was better than the 78.56% of traditional KNN. In summary, the proposed SWPSI-KNN-NCF model can recommend the most suitable music for different users, maximizing the effectiveness of psychological health music intervention and effectively treating the mental health problems of the elderly. However, the proposed method relies on professional EEG acquisition equipment, which limits its application scenarios. The short-term intervention effect evaluation has not yet verified the sustainability of long-term mental health improvement, and the adaptability of music recommendation logic to cultural background differences needs to be strengthened. In the future, low-cost portable biological signal acquisition terminals will be developed to achieve convenient deployment in community scenarios through embedded system optimization. In the future, a dynamic dose-response model can be established through longitudinal follow-up studies lasting 6-12 months to optimize music intervention strategies. At the same time, a cross-cultural knowledge graph of music intervention for the elderly can be constructed, integrating the mapping relationship between regional music elements and neural response characteristics.

References

[1] Rodwin A H, Shimizu R, Travis Jr R, James K J, Banya M, Munson M R. A systematic review of music-based interventions to improve treatment engagement and mental health outcomes for adolescents and young adults. Child and Adolescent Social Work Journal, 2023, 40(4): 537-566. DOI: 10.1007/s10560-022-00893-x.

[2] McCrary J M, Altenmüller E, Kretschmer C, Scholz D. Association of music interventions with health-related quality of life: a systematic review and meta-analysis. JAMA Network Open, 2022, 5(3): e223236-e223236. DOI: 10.1001/jamanetworkopen.2022.3236.

[3] Musgrave G. Music and wellbeing vs. musicians' wellbeing: examining the paradox of music-making positively impacting wellbeing, but musicians suffering from poor mental health. Cultural Trends, 2023, 32(3): 280-295. DOI: 10.1080/09548963.2022.2058354.

[4] Moll-Bertó A, López-Rodrigo N, Montoro-Pérez N, Mármol-López M I, Montejano-Lozoya R. A Systematic review of the effectiveness of non-pharmacological therapies used by nurses in children undergoing surgery. Pain Management Nursing, 2024, 25(2): 195-203. DOI: 10.1016/j.pmn.2023.12.006.

[5] Yıldırım D, Yıldız C Ç. The effect of mindfulness-based breathing and music therapy practice on nurses' stress, work-related strain, and psychological well-being during the COVID-19 pandemic: a randomized controlled trial. Holistic Nursing Practice, 2022, 36(3): 156-165. DOI: 10.1097/hnp.0000000000000511.

[6] Wang F, Huang X, Zeb S, et al. Impact of music education on mental health of higher education students: moderating role of emotional intelligence. Frontiers in psychology, 2022, 13(1): 938090-938091. DOI: 10.3389/fpsyg.2022.938090.

[7] De Witte M, Pinho A S, Stams G J, Moonen X, Bos A E, Van Hooren S. Music therapy for stress reduction: a systematic review and meta-analysis. Health psychology review, 2022, 16(1): 134-159. DOI: 10.1080/17437199.2020.1846580.

[8] Vajpeyee M, Tiwari S, Jain K, Mod, P, Bhandari P, Monga G, Vajpeyee A. Yoga and music intervention to reduce depression, anxiety, and stress during COVID-19 outbreak on healthcare workers. International Journal of Social Psychiatry, 2022, 68(4): 798-807.v DOI: 10.1177/00207640211006742.

[9] Pamungkas A, Isnanto R R, Nugraheni D M K. Implementation of K-Nearest Neighbor in Case-Based Reasoning for Mental Health Diagnosis Systems. Scientific Journal of Informatics, 2024, 11(4): 1109-1120. DOI: 10.15294/sji. v11i4.19912.

[10] Wibowo A P, Taruk M, Tarigan T E, Habibi M Improving Mental Health Diagnostics through Advanced Algorithmic Models: A Case Study of Bipolar and Depressive Disorders. Indonesian Journal of Data and Science, 2024, 5(1): 8-14. DOI: 10.56705/ijodas. v5i1.122.

[11] Cheng J P, Haw S C. Mental health problems prediction using machine learning techniques. International Journal on Robotics, Automation and Sciences, 2023, 5(2): 59-72. DOI: 10.33093/ijoras.2023.5.2.7.

[12] Kolenik T, Schiepek G, Gams M. Computational psychotherapy system for mental health prediction and behavior change with a conversational agent. Neuropsychiatric Disease and Treatment, 2024, 20(1): 2465-2498. DOI: 10.2147/NDT.S417695#d1e194.

[13] Kolenik T. Methods in digital mental health: smartphone-based assessment and intervention for stress, anxiety, and depression//Integrating Artificial Intelligence and IoT for Advanced Health Informatics: AI in the Healthcare Sector. Cham: Springer International Publishing, 2022, 1(1): 105-128. DOI: 10.1007/978-3-030-91181-2_7.

[14] Chinthamu N, Karukuri M. Data Science and Applications. Journal of Data Science and Intelligent Systems, 2023, 1(1): 83-91. DOI: 10.47852/bonviewJDSIS3202837.

[15] Ogunseye E O, Adenusi C A, Nwanakwaugwu A C, Ajagbe S A, Akinola S O. Predictive analysis of mental health conditions using AdaBoost algorithm. ParadigmPlus, 2022, 3(2): 11-26. DOI: 10.55969/paradigmplus. v3n2a2.

[16] Kumar P, Chandra S. Prediction and comparison of psychological health during COVID-19 among Indian population and Rajyoga meditators using machine learning algorithms. Procedia computer science, 2023, 218(1): 697-705. DOI: 10.1016/j.procs.2023.01.050.

[17] Ohannesian G S, Harfash E J. Epileptic seizures detection from EEG recordings based on a hybrid system of Gaussian mixture model and random forest classifier. Informatica, 2022, 46(6): 4203-4204. DOI: 10.31449/inf.v46i6.4203.

[18] Sahu R, Dash S R, Baral A. Identification of Students' Confusion in Classes from EEG Signals using Convolution Neural Network. Informatica, 2024, 48(1): 4604-4605. DOI: 10.31449/inf.v48i1.4604.

[19] Wan J, Zhang S. Hybrid Deep Learning Approach for Ship Navigation in Curved River Sections Using PPO and CNN. Informatica, 2024, 48(22): 15-29. DOI: 10.31449/inf.v48122.6909.

[20] Liu G, Zhao H, Fan F, Liu G, Xu Q, Nazir S. An enhanced intrusion detection model based on improved kNN in WSNs. Sensors, 2022, 22(4): 1407-1425. DOI: doi.org/10.3390/s22041407.