Research on anxiety detection based on personalized data markers

Yuyang Tan1*, Qinwei Zeng1, Hongchang Zhang1
1School of Computer Engineering and Science, Shanghai University, Shanghai 200444, China;
*Corresponding author’s e-mail: tanyuyang@shu.edu.cn

Abstract. The use of machine learning methods to detect human anxiety has become the mainstream. However, the label of current data usually relies on the subject's manual labeling, and there are strong subjective factors. In addition, traditional methods have the disadvantages of insufficient feature extraction, inaccurate classification results, and insufficient generalization ability. In response to the above problems, we propose an objective data labeling method, which extracts features on the basis of maintaining the characteristics of the original physiological signals, constructs a more complete feature set, and builds a long short-term memory network (LSTM) model on this basis perform an anxiety state detect. Experiments show that after processing the data by our method, the performance of all models has been greatly improved, among which LSTM shows the best performance compared with traditional algorithms.

1. Introduction
According to experimental data, about 7.4% of people will face their specific fears at least once in their lives [1]. Exposure therapy is a commonly used method of treatment, and it has been the best way to treat phobias for a long time in the past [2]. With the development of technology, the current virtual reality exposure therapy (VRET) and augmented reality exposure therapy (ARET) have gradually replaced the traditional exposure therapy and become the mainstream, and research has proven that they are effective for various anxiety disorders, such as Flying anxiety, social phobia and arachnophobia, and post-traumatic stress disorder [3-5]. In order to reduce the damage to the patient during the treatment process, the therapist needs to control the intensity and frequency of the fear object in real time according to the patient's anxiety state. However, human manipulation judgments usually fail to achieve accurate results. Therefore, it is of great significance to study an accurate and efficient anxiety recognition algorithm to replace manual operation.

Healey and Picard [6] concluded that using a linear discriminator to detect pressure levels (low, medium, and high) can obtain an accuracy of 97.4%. The physiological indicators used are electromyography, electrocardiogram, electrical skin activity, and respiration. The data comes from 24 drivers in Boston. Keshan [7] et al. and Chen [8] et al. use the same data, but use different classifiers and window lengths (5 minutes and 10 seconds). The results they obtained showed that using a support vector machine as a classifier, while using electrocardiogram (ECG), electrical skin activity (EDA) and respiration (RSP) signals as features as input, can reach 89% in the detection of two pressure levels The accuracy of using neural networks and decision trees is used to classify the three pressure levels, and the accuracy is about 70%. In addition, Baru [9] and others also used breathing and temperature sensors to obtain different physiological measurement data with a time window of 60 seconds. Their research explored three different classifiers (neural network, support vector machine and case-based Reasoning)
to distinguish between the two levels of pressure, and an accuracy rate of 85.6 was obtained in case-based reasoning. In a related VRET experiment, Handouzi [10] and others hope to distinguish between two different levels of fear in 7 patients with social phobia. They used 6 different virtual reality scenarios to induce anxiety and extracted features from blood volume pulse signals. The support vector machine algorithm with a 20-second time window can recognize the two anxiety levels with an accuracy of 76%.

On the basis of the above research, we use the LSTM model as the identification method, combined with the data labeling method proposed in this paper to eliminate the subjectivity of artificial labeling, shorten the time window size to extract more data, and make the features more suitable for the original physiological signal pattern Trend distribution, the introduction of six features to establish a new feature set, as far as possible to improve the accuracy of the algorithm model.

2. Data processing and feature extraction
ECG, RSP and EDA are widely used to detect the stress and anxiety state of patients. Most studies believe that ECG signal is one of the most suitable physiological signals for detecting the above symptoms. It is considered simple and can provide the most effective information [13]. The data used in our study comes from the physiological data of 57 volunteers in the Physionet [11] database [12]. The data set contains records from 57 subjects, all controlled between 18 and 40 years old. They divided the subjects into four groups with 20 subjects in each group. The experiment collected data on three physiological indicators of these subjects when watching the scary spider clips: EDA, ECG, and RSP. The collection time of each subject is approximately 35 minutes, of which the length of watching the horror movie is 16 minutes, which is composed of 16 video clips (the whole process is collected at a frequency of 100HZ), and the rest period is 5 minutes.

2.1. Data tagging
Most researchers label data by the intuitive feelings of patients, which will cause the labeling of the data to have strong subjective factors. So we defined three data labeling methods to eliminate manual intervention. The basic principle of our labeling data is based on: Heart rate (HR) and EDA will increase as the body is stimulated [13]; there are differences in the standard of physiological values for each person during relaxation. In this paper, we calculate the normalized mean value of the HR signal to reflect the magnitude of the numerical change of the human body at different moments compared to the stationary phase. The greater the magnitude of the change, the more intense the physiological response of the human body at this time, and vice versa the more stable.

Method one (TAG333): We calculate the normalized mean value of the HR signals of all subjects for each video segment (see formula 1) and rank them. Mark the three HR video clips with the highest value as ‘high’, how to mark the next three groups of video clips as ‘medium’, and mark the three groups of video clips with the lowest average value as ‘low’. The advantage of data labeling in this way is that the amount of data in each category is balanced, but this labeling of data will have the following problem: the same segment has the same impact on everyone.

\[
HRN_{\text{mean}} = \frac{1}{N} \sum_{n=1}^{N} \left( \frac{1}{M} \sum_{m=1}^{M} x_{mn} - \frac{1}{M} \sum_{m=1}^{M} \mu_{m} \right)
\]

In formula (1), N represents the total number of records of this video, and M represents the total number of participants in the experiment (57 people), \(x_{mn}\) represents the record value of the m-th subject on the segment. \(\mu_{m}\) represents the mean value of the HR signal of the m-th subject at rest.

Method two (TAG885): We normalized the HR of each subject on each record of each video clip with their own five-minute HR mean during the rest period, so each subject performed a calculation on each record. Videos have normalized HR values. For different testers, we marked the eight video clips with the highest value as ‘high’, the remaining eight video clips as ‘medium’, and the data value of the rest period as ‘low’. Therefore, the ‘high’ and ‘medium’ video clips marked for each person are different.
This effectively prevents the same video clip from causing the same impact on everyone, but it also causes uneven data distribution.

\[
HRNmean = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)
\] (2)

In formula (2), \(N\) represents the total number of records of this video, \(x_n\) represents the HR value of the nth record of the subject on the segment. \(\mu\) represents the average value of the HR signal during the rest period of the person.

Method three (TAG444): Similar to the second method, but we change the amount of data used, we mark the four video clips with the highest value as 'high', and mark the next four video clips as 'medium', Mark the last four minutes of the break as 'low'.

2.2. Time window selection
The advantages of choosing a shorter time window are: 1. Increase the total amount of data. 2. The smaller the length of the window, the more subtle changes in the psychological state can be reflected [8]. 3. Make the features we extracted more closely fit the trend distribution of the original physiological graphics. Although the smaller the length of the time window, the more it can reflect the changes in the details of the mind, but it should be noted that the EDA and ECG signals show meaningful changes in the shortest time [8]. Finally, we used three different time window sizes of 50 seconds, 10 seconds and 5 seconds to carry out the experiment

2.3. Feature extraction
Frank [13] et al. found through control experiments: the normalized mean of heart rate (HRNmean), the standard deviation of heart rate (HRstd), the normalized mean of electrical skin activity (EDANmean), the absolute value of the normalized first-order difference of electrical skin activity The average value (EDANFD), the number of directed responses to electrical skin activity (EDAOR), and the average range of directed responses to electrical skin activity (EDAmOR) six feature sets (for the convenience of memory, we named the feature set TYY) are for our experiment The most relevant feature set for the result. The calculation formula for each feature is shown below.

\[
HRNmean = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)
\] (3)

In formula (3), \(N\) represents the total number of records of this video, \(x_n\) represents the HR value of the subject on the nth record, \(\mu\) represents the average value of the HR signal of the person during the rest period.

\[
EDANmean = \frac{1}{N} \sum_{n=1}^{N} \frac{x_n - \min (x_i)}{\max (x_i) - \min (x_i)}
\] (4)

In formula (4), \(N\) represents the total number of records of this video, \(x_n\) represents the EDA value of the subject on the nth record, \(x_i\) represents the EDA signal value of the person in the resting phase.

\[
HRstd = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N} (x_n - u_x)^2}
\] (5)

In formula (5), \(N\) represents the total number of records of this video, \(x_n\) represents the HR value of the subject on the nth record, \(u_x\) represents the average value of the HR signal value in the current segment.

\[
EDANFD = \frac{1}{N-1} \sum_{n=1}^{N-1} |\hat{x}_{n+1} - \hat{x}_n|
\] (6)
In formula (6), $N$ represents the total number of records of this video, $\hat{x}_n$ which is the signal value of EDA after standardization.

$$nOR = \#OR$$

$$mmOR = \frac{1}{nOR} \sum_{n=1}^{nOR} mOR_n$$

In formula (7), $nOR$ represents the number of directional responses, in formula (8), $mOR_n$ represents the response range, and $mmOR$ represents the average response range.

In further experiments, we continue to explore the construction of feature sets with higher performance. Previous studies have shown that when a person is under stress, the P wave and R wave interval (PR interval), the Q wave and T wave interval (QT interval) and QRS duration (QRS duration) in the ECG image will change. Produce obvious changes, these changes are considered to be useful indicators to detect heartbeat pressure [7]. Based on the above theory, we conducted a tenth set of experiments. In the feature set (TYY), we added six features extracted from the original ECG signal that can reflect the pattern change law, including QRS interval average (Average QRS interval), RR interval average (Average RR interval), QQ interval average (Average QQ interval), SS interval average (Average SS interval), QR interval average (Average QR interval), RS interval average (Average RS interval), create a new The feature set of (for the convenience of memory, we named the new feature set CYH) to explore its impact on the accuracy of anxiety classification results.

3. Classifier selection

Traditional machine learning methods have large deficiencies in generalization capabilities, and are limited by the amount of data and feature dimensions, and cannot use deep learning models. However, now the data processed by our method has been improved in both scale and dimension, and considering that we are using time series data, the advantage of the LSTM model is that it can process time series information. The storage unit and gate effectively learn the dependence between long-term and short-term data, so we use LSTM as our preferred classifier model.

4. Analysis of results

We use ten-fold cross for verification, and use the calculated results to provide a measure of the performance of different classifiers. We set four indicators to measure performance: accuracy (AUC), true high rate (THR), true middle rate (TMR) and true low rate (TLR). THR, TMR, and TLR respectively represent the accuracy rates of the model's correct classification of the three anxiety states of 'high', 'medium' and 'low'. Table 1, Table 2, and Table 3 show that we use the window time of 50 seconds, five algorithms, three different marking feature methods, and the results of two classifications and three classifications. The feature set used is TYY. Under this condition, using TAG885 to extract the feature long and short-term memory network can achieve the highest accuracy rate of 89.9% in the second classification, and the support vector function can achieve the highest classification accuracy of 68.5% in the third classification. At this time, due to the small amount of data and feature set, the LSTM model failed to show excellent performance.

| Classifier | levels | AUC | THR | TMR | TLR |
|------------|--------|-----|-----|-----|-----|
| SVM        | 2 levels | 89.3 | 91.4 | -   | 87.2 |
| KNN        | 2 levels | 84.2 | 84.9 | -   | 83.5 |
| DT         | 2 levels | 80.8 | 81.3 | -   | 80.3 |
| NB         | 2 levels | 88.5 | 88.1 | -   | 88.9 |
| LSTM       | 2 levels | 87.4 | 86.1 | -   | 88.7 |
| SVM        | 3 levels | 67.1 | 60.3 | 59.9 | 81.4 |
Table 2 Method 2 and experimental results with a window size of 50 seconds

| Classifier | Method | 2 levels | 3 levels | 4 levels | 5 levels |
|------------|--------|----------|----------|----------|----------|
| SVM        | 2 levels | 88.3     | 89.3     | -        | 88       |
| KNN        | 2 levels | 89.6     | 87       | -        | 90.2     |
| DT         | 2 levels | 85.1     | 84.3     | -        | 85.9     |
| NB         | 2 levels | 86.6     | 87.8     | -        | 86.4     |
| LSTM       | 2 levels | 89.9     | 89.7     | -        | 90.1     |
| SVM        | 3 levels | 68.5     | 56.6     | 60.1     | 87.4     |
| KNN        | 3 levels | 67.1     | 59.9     | 61.1     | 81.5     |
| DT         | 3 levels | 65.3     | 61.1     | 52.9     | 80.6     |
| NB         | 3 levels | 64.7     | 55.9     | 60.2     | 79.4     |
| LSTM       | 3 levels | 66.9     | 55.1     | 57.4     | 87.9     |

Table 3 Method 3 and experimental results with a window size of 50 seconds

| Classifier | Method | 2 levels | 3 levels | 4 levels | 5 levels |
|------------|--------|----------|----------|----------|----------|
| SVM        | 2 levels | 87.7     | 88       | -        | 87.4     |
| KNN        | 2 levels | 86.9     | 87.1     | -        | 86.7     |
| DT         | 2 levels | 84.8     | 85.6     | -        | 82.4     |
| NB         | 2 levels | 83.4     | 84.1     | -        | 81.7     |
| LSTM       | 2 levels | 86.4     | 86.5     | -        | 86.3     |
| SVM        | 3 levels | 66.1     | 58.6     | 56.1     | 83.6     |
| KNN        | 3 levels | 62.2     | 54.3     | 53.4     | 78.9     |
| DT         | 3 levels | 65.8     | 55.2     | 58.4     | 83.8     |
| NB         | 3 levels | 62.8     | 59.1     | 53.5     | 75.8     |
| LSTM       | 3 levels | 64.3     | 54.2     | 58.7     | 80       |

Table 4, Table 5, and Table 6 show that we use a time window of 10 seconds, five algorithms, three different marking feature methods, and the results of two classifications and three classifications. The feature set used is contractual. We can find that whether it is two- or three-category, under the premise of using TAG885 tag features, the long- and short-term memory network can achieve the highest accuracy. The accuracy of the two-category and the three-category are 91.3% and 76.9%, respectively. Compared with the time when the window size is 50 seconds, as the amount of data increases, the data features are more and more able to reflect the characteristics of the original physiological signal. The accuracy of the second and third classifications of all models has been improved to varying degrees, and this At that time, the classification results we obtained have exceeded Frank et al. (second classification: 89.8%, third classification: 74.4%).

Table 4 Method 1 and experimental results with a window size of 10 seconds

| Classifier | Method | 2 levels | 3 levels | 4 levels | 5 levels |
|------------|--------|----------|----------|----------|----------|
| SVM        | 2 levels | 88.3     | 88.8     | -        | 87.8     |
| KNN        | 2 levels | 86.1     | 87       | -        | 85.2     |
| DT         | 2 levels | 87.5     | 87.1     | -        | 87.9     |
| NB         | 2 levels | 86.2     | 86.9     | -        | 85.5     |
| LSTM       | 2 levels | 90.1     | 89.1     | -        | 91.1     |
| SVM        | 3 levels | 70.1     | 61.9     | 65.3     | 83.1     |
| KNN        | 3 levels | 67.3     | 56.9     | 63.1     | 81.9     |
| DT         | 3 levels | 68.8     | 61.2     | 55.6     | 89.6     |
| NB         | 3 levels | 66.8     | 59.1     | 59.5     | 81.8     |
| LSTM       | 3 levels | 69.9     | 61.8     | 62.4     | 85.5     |
Table 5  Method 2 and experimental results with a window size of 10 seconds

| Classifier | levels | AUC  | THR  | TMR  | TLR  |
|------------|--------|------|------|------|------|
| SVM        | 2 levels | 87.9 | 86.9 | -    | 89.1 |
| KNN        | 2 levels | 87.3 | 88.4 | -    | 86.5 |
| DT         | 2 levels | 84.6 | 83.2 | -    | 85.7 |
| NB         | 2 levels | 89.1 | 89.9 | -    | 88.4 |
| LSTM       | 2 levels | 91.3 | 91.9 | -    | 90.4 |
| SVM        | 3 levels | 72.3 | 65.1 | 61.7 | 89.8 |
| KNN        | 3 levels | 69.4 | 61.3 | 57.9 | 89.4 |
| DT         | 3 levels | 70.7 | 61.2 | 62.5 | 87.6 |
| NB         | 3 levels | 68.8 | 59.3 | 62.5 | 84.6 |
| LSTM       | 3 levels | 76.9 | 71.8 | 68.3 | 91.3 |

Table 6  Method 3 and experimental results with a window size of 10 seconds

| Classifier | levels | AUC  | THR  | TMR  | TLR  |
|------------|--------|------|------|------|------|
| SVM        | 2 levels | 84.4 | 84.1 | -    | 84.7 |
| KNN        | 2 levels | 82.3 | 81.1 | -    | 83.5 |
| DT         | 2 levels | 81.9 | 83.3 | -    | 80.5 |
| NB         | 2 levels | 83.3 | 81.9 | -    | 84.7 |
| LSTM       | 2 levels | 86.1 | 85.9 | -    | 86.3 |
| SVM        | 3 levels | 71.2 | 68.1 | 62.3 | 83.2 |
| KNN        | 3 levels | 70.1 | 58.4 | 66.1 | 85.6 |
| DT         | 3 levels | 66.7 | 56.9 | 59.1 | 84.1 |
| NB         | 3 levels | 65.4 | 61.7 | 52.9 | 81.6 |
| LSTM       | 3 levels | 71.9 | 66.3 | 64.8 | 84.6 |

Table 7, Table 8, and Table 9 show that we use a time window of 5 seconds, five algorithms, three different marking feature methods, and the results of two classifications and three classifications. The feature set used is contractual. Similarly, we found that when using TAG885 for labeling, the long-short-term memory network can get the highest accuracy in the second and third classifications, which are 92.7% and 82.9%, respectively. Compared with other labeling methods, all classifiers are Can get the highest accuracy rate compared to its own on TAG885. Although the true high rate and true hit rate of the three classifications have also been improved a bit compared to the previous ones, there is still a large room for improvement. We guess that this result is because the features we currently use cannot be clearly distinguished.' Medium' and 'high' anxiety state.

Table 7  Method 1 and experimental results with a window size of 5 seconds

| Classifier | levels | AUC  | THR  | TMR  | TLR  |
|------------|--------|------|------|------|------|
| SVM        | 2 levels | 91.7 | 90.1 | -    | 93.3 |
| KNN        | 2 levels | 88.2 | 89.7 | -    | 86.7 |
| DT         | 2 levels | 89.5 | 90.2 | -    | 88.8 |
| NB         | 2 levels | 87.9 | 86.2 | -    | 89.6 |
| LSTM       | 2 levels | 92.4 | 92.1 | -    | 92.7 |
| SVM        | 3 levels | 77.1 | 68.8 | 70.3 | 92.2 |
| KNN        | 3 levels | 74.4 | 71.3 | 69.1 | 82.8 |
| DT         | 3 levels | 74.1 | 69.9 | 73.1 | 79.3 |
| NB         | 3 levels | 72.8 | 62.3 | 68.4 | 87.7 |
| LSTM       | 3 levels | 78.3 | 72.4 | 76.8 | 85.7 |
Table 8 Method 2 and experimental results with a window size of 5 seconds

| Classifier | levels | AUC  | THR | TMR | TLR |
|------------|--------|------|-----|-----|-----|
| SVM        | 2 levels | 91.4 | 93.1 | -   | 90.1|
| KNN        | 2 levels | 88.9 | 88.1 | -   | 89.5|
| DT         | 2 levels | 88.5 | 87.9 | -   | 89.9|
| NB         | 2 levels | 89.1 | 90.7 | -   | 88.4|
| LSTM       | 2 levels | 92.7 | 91.2 | -   | 93.9|
| SVM        | 3 levels | 75.6 | 69.4 | 71.1| 88.2|
| KNN        | 3 levels | 75.1 | 70.1 | 67.4| 89.6|
| DT         | 3 levels | 76.3 | 71.1 | 66.5| 90.9|
| NB         | 3 levels | 78.9 | 74.9 | 73.7| 89.3|
| LSTM       | 3 levels | 82.9 | 77.7 | 76.9| 93.5|

Table 9 Method 3 and experimental results with a window size of 5 seconds

| Classifier | levels | AUC  | THR | TMR | TLR |
|------------|--------|------|-----|-----|-----|
| SVM        | 2 levels | 89.9 | 90.3 | -   | 89.5|
| KNN        | 2 levels | 85.3 | 83.9 | -   | 86.7|
| DT         | 2 levels | 89.1 | 87.2 | -   | 91  |
| NB         | 2 levels | 83.6 | 84.1 | -   | 83.1|
| LSTM       | 2 levels | 88.2 | 87.5 | -   | 88.9|
| SVM        | 3 levels | 75.3 | 71.1 | 68.8| 86  |
| KNN        | 3 levels | 74.9 | 69.4 | 65.7| 89.6|
| DT         | 3 levels | 73.8 | 61.8 | 69.5| 90.1|
| NB         | 3 levels | 73.1 | 66.5 | 60.9| 91.9|
| LSTM       | 3 levels | 76.9 | 73.3 | 71.8| 85.6|

In order to explore the correlation of the six features related to the peak extracted from the ECG signal to the prediction of anxiety state, we have a time window size of 5 seconds and TAG885 extraction features under the premise, Six features extracted from ECG were added to the TY collection, a new feature collection CYH was established and a control experiment was carried out. The results of the experiment are shown in Table 10 below. We found that after adding the six features, the accuracy of all classifiers has been improved to varying degrees. Among them, the long and short-term memory network can get the highest accuracy in the two classification and the three classification, which are 95.6% and 84.3%.

Table 10 Experimental results after adding six features

| Classifier | levels | AUC  | THR | TMR | TLR |
|------------|--------|------|-----|-----|-----|
| SVM        | 2 levels | 92.1 | 93.4 | -   | 90.5|
| KNN        | 2 levels | 89.7 | 88.4 | -   | 91.3|
| DT         | 2 levels | 89.4 | 88.4 | -   | 90.2|
| NB         | 2 levels | 90.9 | 91.4 | -   | 90.2|
| LSTM       | 2 levels | 95.6 | 94.2 | -   | 96.9|
| SVM        | 3 levels | 77.3 | 70.9 | 70.2| 90.6|
| KNN        | 3 levels | 76.9 | 73.5 | 68.9| 89.2|
| DT         | 3 levels | 78.4 | 72.8 | 67.7| 92.5|
| NB         | 3 levels | 79.3 | 72.1 | 76.8| 88.4|
| LSTM       | 3 levels | 84.3 | 80.1 | 79.5| 93.3|
5. Conclusion
This article proposes a new method of data tagging (TAG885) combined with a smaller time window (5 seconds). Based on this method, the anxiety state of the human body is detected through the deep learning method, and a series of biological signals obtained by online monitoring technology are used as input for model training. The results of the experiment show that the TAG885 tag feature is used, and the time window size is 5 seconds. Under the premise of extracting features, regardless of two-level (low, high) or three-level (low, medium, and high) anxiety classification, the accuracy of all classifier models has been greatly improved, and the LSTM classifier shows the best performance. We propose a new feature set (CYH) for human anxiety detection. The results show that all models have also been improved to varying degrees when the new feature set is selected as input. The results show that all the models are improved to different degrees when the new feature set is selected as the input, and the LSMT classifier shows the best performance, which is significantly higher than other classifiers. Compared with the results of Frank et al., the accuracy of the second classification has increased by 7.1 percentage points to 95.6%, and the accuracy of the three classification has increased by 9.9 percentage points to 84.3%.

References
[1] Wardenaar K J , Lim C C W , Al-Hamzawi A O , et al. The cross-national epidemiology of specific phobia in the World Mental Health Surveys[J]. Psychological Medicine, 2017, 48(5):1.
[2] Wolitzky-Taylor K B , Horowitz J D , Powers M B , et al. Psychological approaches in the treatment of specific phobias: A meta-analysis[J]. Clinical Psychology Review, 2008, 28(6):1021-1037.
[3] Nexhmedin, Morina, Hiske, et al. Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments[J]. Behaviour Research & Therapy, 2015.
[4] Fodor L A , Cote? C D , Pim C , et al. The effectiveness of virtual reality based interventions for symptoms of anxiety and depression: A meta-analysis[J]. Scientific Reports, 2018, 8(1):10323-.
[5] Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials.[J]. Journal of Anxiety Disorders, 2018.
[6] Healey J A , Picard R W . Detecting stress during real-world driving tasks using physiological sensors[J]. IEEE Transactions on Intelligent Transportation Systems, 2005, 6(2):156-166.
[7] Keshan N , Parimi P V , Bichindaritz I . Machine learning for stress detection from ECG signals in automobile drivers[C]// IEEE International Conference on Big Data. IEEE, 2015.
[8] Chen L L , Zhao Y , Ye P F , et al. Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers[J]. Expert Systems with Applications An International Journal, 2017:279-291.
[9] Barua S , Begum S , Ahmed M U . Supervised Machine Learning Algorithms to Diagnose Stress for Vehicle Drivers Based on Physiological Sensor Signals[J]. Stud Health Technol Inform, 2015, 211:241-248.
[10] W. Handouzi, C. Maaoui, A. Pruski, A. Moussaoui and Y. Bendjouid, "Short-Term Anxiety Recognition Induced by Virtual Reality Exposure for Phobic People," 2013 IEEE International Conference on Systems, Man, and Cybernetics, Manchester, UK, 2013, pp. 3145-3150, doi: 10.1109/SMC.2013.536.
[11] Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. e215–e220.
[12] Ihmig, F. R., Gogasascoechea, A., Schäfer, S., Lass-Hennemann, J., & Michael, T. (2020). Electrocardiogram, skin conductance and respiration from spider-fearful individuals
watching spider video clips (version 1.0.0). PhysioNet. https://doi.org/10.13026/sq6q-z04.

[13] Ihmig F R, Antonio G H, Neurohr-Parakenings F, et al. On-line anxiety level detection from biosignals: Machine learning based on a randomized controlled trial with spider-fearful individuals[J]. PLOS ONE, 2020, 15.