Machine Learning-based Signal Processing by Physiological Signals Detection of Stress

T. Lakshmi Prasanthi [1], K.Prasanthi [2]

[1]Dept. Of ECE Sri Padmavathi Mahila Viswavidyalayam, Tirupati, India
[2]Dept. Of ECE Sri Padmavathi Mahila Viswavidyalayam, Tirupati, India

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 10 May 2021

Abstract: Stress is a normal part of daily life that most people experience at different times. However, chronic stress or high levels of stress will jeopardize our safety and disrupt our normal lives. As a result, the capacity to operate and manage in critical situations is greatly reduced. Therefore, it is necessary to understand pressures and design processes with an understanding of pressure. In this paper, we are introduced to the process of processing signals according to machine learning algorithms:We have used natural data collected, such as Respiration, GSR Hand, GSR Foot, Heart Rate, and EMG, from various studies in a variety of situations and areas while driving. After that, the data division at different times, such as 100, 200 seconds, and 300 seconds, was done differently.We have removed the mathematical features from the separated data and fed these features into the available separator. We used KNN, K's closest neighbor, and a vector support machine, which is very different.We divided the pressure into three levels: low, medium, and high.Our results show that the pressure level can be reached with an accuracy of 98.41% in 100 seconds and 200 seconds simultaneously and 99% with a time interval of 300 seconds.

Keywords: Machine learning; Processing Symptoms; GSR; EMG; HR; Breathing; KNN; SVM.

1. Introduction

Depression occurs when a person is unable to meet a great need that is not set. The effects of stress are felt ph ysically, mentally, and emotionally. Existing studies have shown that a person's health knowledge can be affected by physical and mental stress. Physical information, which can be obtained from organic or sensory organs, usually includes Electrocardiogram (ECG), Galvanic Skin Response (GSR), Electromyogram (EMG), Respiration (RESP), Finger Temperature (FT), Skin Temperature (ST), and blood volume drive (BVP). Acquisition of work-related stress using biological signals can be divided into two categories:

1. Using EEG signals (Electroencephalographs)
2. Using GSR, ECG, EMG, ST, respiration, etc. Or a combination of them

Although pressure detection using an EEG sensor results in high accuracy, the use of an electroencephalograph electrode due to its weight is not always available.

The second mode uses GSR, ECG, EMG, ST, and RESPIRATION sensors or their combination. Using these sensors is easier than EEG. However, after receiving signals from these sensors, Signals can be processed using Wavelet transform] or calculated mathematical or offline features to extract the desired object [2 - 13].

Jennifer A. Healey and Rosalind W. Picard [2, 3] wrote and analyzed physical symptoms such as ECG, EMG, GSR on foot and hand, and actual respiratory function (RESP) to determine driver stress levels in three different areas. They extracted 22 elements with five signals using discriminatory function (LDF) at different levels of stress.

K. Soman et al. [8] based on driver data published in [2, 3], using ECGs and Respiration Signals and QRS power spectrum and respiratory rate are released. They have shown a positive relationship between QRS power and drivers' respiratory pressure.

Y. Deng et al. [9], based on the driver database published in [2, 3], use a variety of feature sets and apply the learning techniques to five machines such as naïve Bayes (NB), vector support (SVM) machine, decision solution C4. 5, racist activity (LDF), and close neighbors (KNN) to differentiate the level of stress.

Yong Deng et al. [10], based on the driver database published in [2, 3], select the appropriate features and reduce their value from 22 to 5 using the main object analysis (PCA). These factors resulted in an accuracy of 78.94%.

P. Karthikeyan et al. [11] used ECG, EMG, Heart Rate Variability (HRV), GSR, and ST signals obtained from 40 subjects using cognitive, mathematical performance to apply pressure. While in the HRV Higher-order spectra (HOS), the results were 93.75% accurate, and without the use of HOS, pressure detection was reduced by about 75%.

H. Kurniawan et al. [12] GSR expression and symptoms were used using the Stroop Color test, the Trier Social Stress test, and the Trier Mental Challenge test stimulus promoting stress was collected. Various features of GSR and Speech were used separately, using four separators. The best GSR accuracy was 70%, and the expression was 92%.

Zhai and Barreto [13] used four body symbols, GSR, Blood Volume Pulse (BVP), Pupil Diameter (PD), and ST, to locate computer users and to use the three machine learning methods NB, SVM and Decision Tree, to differentiate the level of stress.

For most of these operations, high accuracy is achieved whenever a large number of sensors or multiple features are used, meaning that a large amount of processing and time is unavoidable. This paper has attempted to obtain the highest accuracy of various sensory numbers and various features. The process of finding pressure in this paper is divided into three stages; first, subtract 78 features from five body parts; secondly, depending on the number of sensors used, select the appropriate features and finally, the SVM and KNN machine learning techniques used for classification.

2. Method

A. Data Acquisition

PHYSIONET website's physiological signals is used. This data contains various signals from healthy people travelling in roads including city roads as more stress, normal highway as moderate stress and remaining as less stress. It contains data of seventeen 17 drivers and in this we have taken seven drivers (drivers six, seven, eight, ten, eleven, twelve and fifteen) that has total info. Five signs for each driver; Galvanic Foot Skin Response (FGSR), Galvanic Skin Response (HGSR), Electromyogram (EMG), Heart Rate (HR) Based Electrocardiogram (ECG), and Rpirpiration (RESP). For example, different 'drive06' signals in Healey and Picard's experiments are shown in Fig. 1.

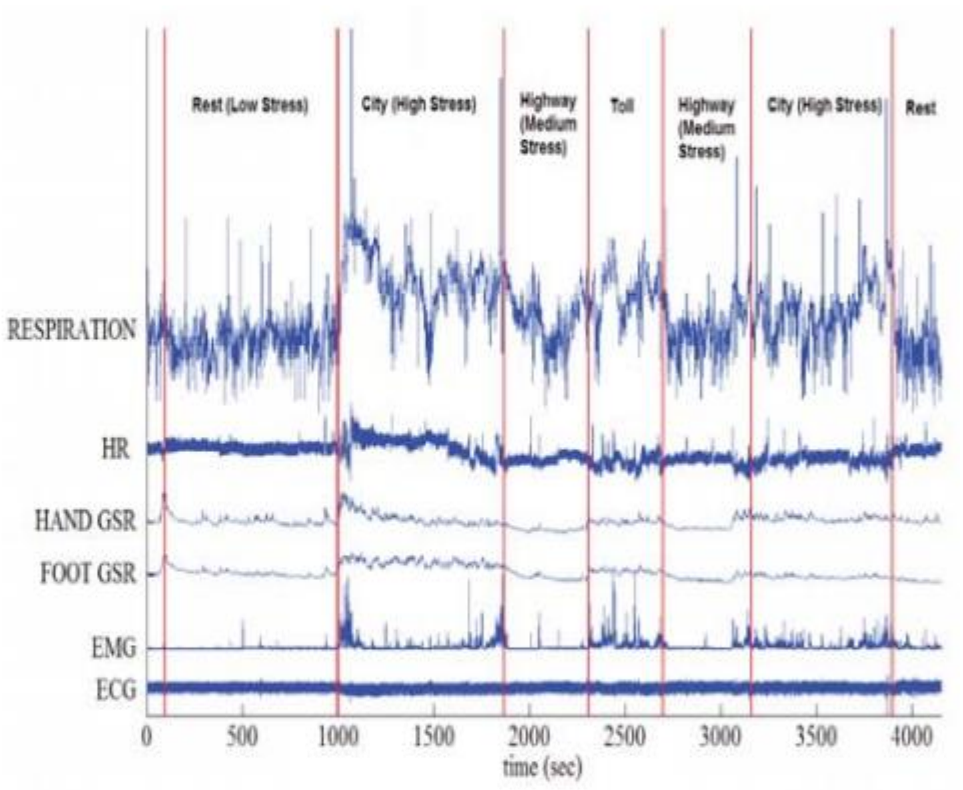

Fig. 1 Different signals for 'drive06'

For processing these data, the whole data is divided into an individual segment for different states. Here various time intervals segmentations such as 100, 200 and 300 seconds for three levels of low stress (relax), medium stress and high stress is applied to signals mentioned in fig. 1.

Each signal are divided into nine segments for 100 seconds intervals state; which three segments with 50% overlap belong to the first rest period as low stress (Fig.2), three segments with 50% overlap belong to the first city period as high stress (Fig.3), and three segments with 50% overlap belong to first highway period as medium stress (Fig.4).

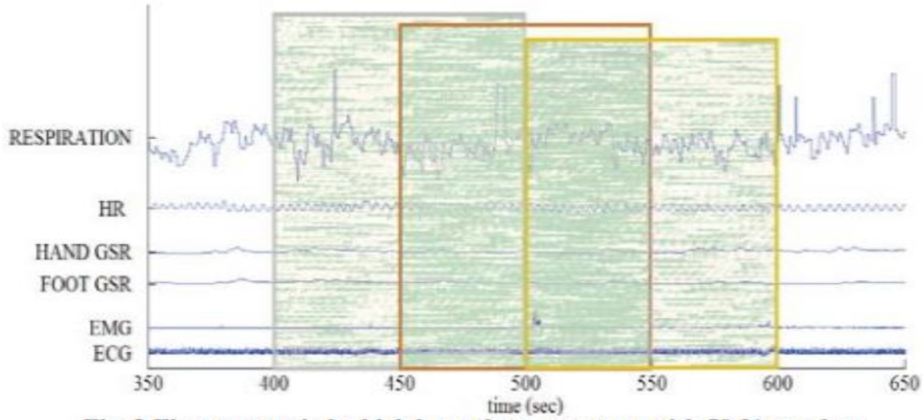

Fig. 2 First rest period which have three segments with 50 % overlap

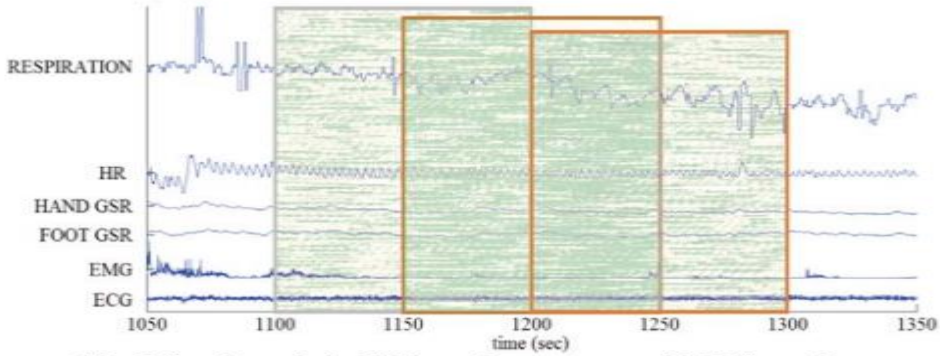

Fig. 3 First city period which have three segments with 50 % overlap

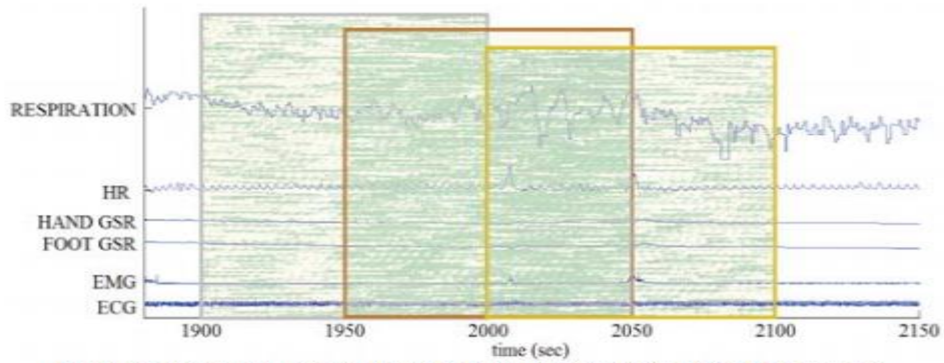

Fig. 4 First highway period which have three segments with 50 % overlap

The same job is done at 200 seconds intervals. For 300 seconds intervals, each signal is divided into three segments, the first segment belong to the first rest period as low stress (Fig.5), the second segment belongs to the first city period as high stress (Fig.6), and the third segment belong to first highway period as medium stress (Fig.7).

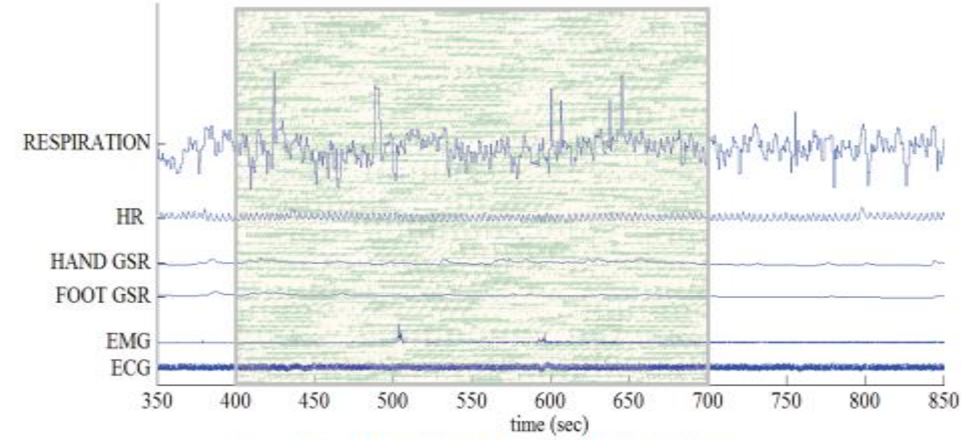

Fig. 5 First rest period for 300 seconds state

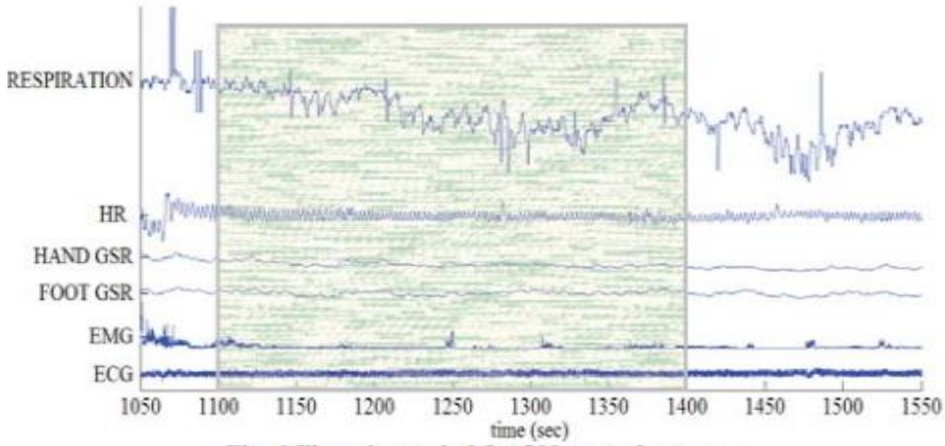

Fig. 6 First city period for 300 seconds state

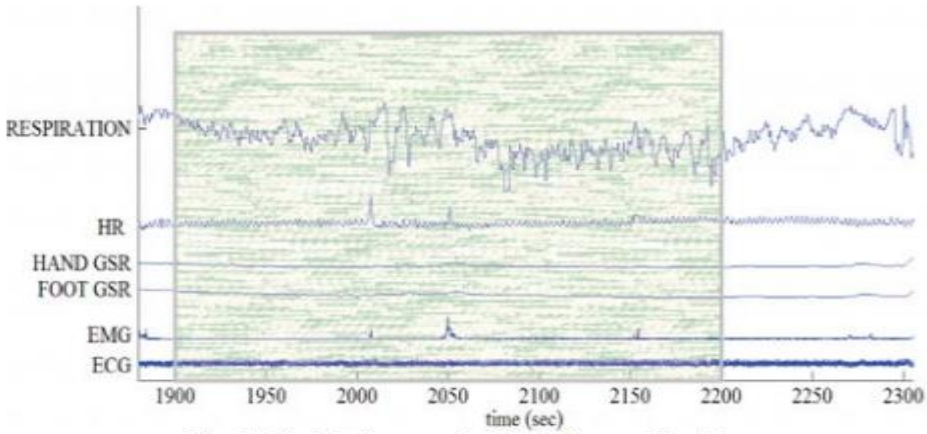

Fig. 7 First highway period for 300 seconds state

B. Feature Extraction

78 features are extracted for each segment. All features are chosen based on the most important and most used for physiological signals according to [2-13]. These features are summarized in table 1.

Table 1 Feature symbol and description

	Feature Symbol				
Feature Description	EMG	HR	Foot GSR	Hand GSR	RESP
Mean Normalization	EMG1	HR1	FGSR1	HGSR1	RESP1
Root Mean Square (RMS)	EMG2	HR2	FGSR2	HGSR2	RESP2
The average power 0.01 to 0.1 Hz	EMG3	HR5	FGSR3	HGSR3	RESP3
The average power 0.1 to 0.2 Hz	EMG4	HR6	FGSR4	HGSR4	RESP4
The average power 0.2 to 0.3 Hz	EMG5	HR7	FGSR5	HGSR5	RESP5
The average power 0.3 to 0.4 Hz	EMG6	HR8	FGSR6	HGSR6	RESP6
The average power F1 to F2 Hz	EMG7	HR3	FGSR7	HGSR7	------
The average power F3 to F4 Hz	EMG8	HR4	FGSR8	HGSR8	------
Ratio low band / high band	EMG9	HR9	FGSR9	HGSR9	RESP7
Means of differences between adjacent elements	EMG10	HR10	FGSR10	HGSR10	RESP8
Means of differences between adjacent elements 2nd times	EMG11	HR11	FGSR11	HGSR11	RESP9
Interquartile Range (IQR)	EMG12	HR12	FGSR12	HGSR12	RESP10
Sum of Rise time from the 10% to 90% of reference levels	EMG13	HR13	FGSR13	HGSR13	RESP11
Peak 2 Peak	EMG14	HR14	FGSR14	HGSR14	RESP12
Sum of local peak	EMG15	HR15	FGSR15	HGSR15	RESP13
Number of local peak	EMG16	HR16	FGSR16	HGSR16	RESP14

In table 1, some frequency is unknown that we introduced them in table 2.

Table 2. Unknown Frequency in table 1.

Signal/Freq.(Hz)	F1	F2	F3	F4	Low Band	High Band
EMG	0.4	0.5	0.01	7	0.01 to 0.1	0.01 to 7
HR	0.01	0.15	0.15	0.5	0.01 to 0.15	0.15 to 0.5
Foot GSR	0.02	0.5	0.5	1.5	0.02 to 0.5	0.5 to 1.5
Hand GSR	0.02	0.5	0.5	1.5	0.02 to 0.5	0.5 to 1.5
Respiration	-	-	-	-	0.01 to 0.1	0.3 to 0.4

C. Feature Selection

Feature vectors have 78 characteristics of any part of all signals; This creates longer training times and a sophisticated computer, hence choosing the best features to improve the computation speed of the partition. The feature selection algorithm can be displayed in conjunction with the search process to select new feature subsets and a test rating that finds different feature subsets. The simplest algorithm to test each set of available features that detects the error rate. In the Faka software, all features are listed by Cfs Subset Eval and Info Gain Attribute Eval. Another algorithm is based on machine learning methods, which selects the best features. In informal software, the best features are found in the categories by selecting the EVA and SVM classifiers from the classifier.

In this paper, both algorithms are used to determine the best features. First, by combining both algorithms, all features are calculated. The mains were selected and then combined with the features selected in the second algorithm, and classification was done in different countries. In the second method, a second algorithm and division element are created for only one country. After the active features are selected, SVM and KNN with cross-validation are used for isolation. First, Matlab 2012a is used for signal processing and feature removal. SVM and KNN were later Subdivided using WEKA3.6.

Table 3 Analysis features for 100 seconds intervals by using SVM

Features used for 100 seconds state	Number of Features	SVM Classification	Number of Sensor
ALL	78	90.47 %	5
RESP 2,3,4,5,9,10,14 - HR 1 - EMG 8,14,16 -HGSR 8,10,13,15 - FGSR 4,9,10,15,16	20	98.41 %	5
RESP 2,5,9,10 - HR 1 - EMG 16 - HGSR 8 - FGSR 4	8	96.82 %	5
RESP 2,5,9,10 - HR 1 - HGSR 8	6	93.65 %	3
RESP 2,5,9,10 - HGSR 8	5	87.30 %	2
RESP 1,2,5,9,14	5	84.12 %	1
ClassifierSubsetEval feature selection : EMG 3,14,16 - HGSR 7,15 - FGSR 9 - RESP 4,9,10,14	10	82.54%	4

Table 4 Analysis features for 100 seconds intervals by using KNN

3. Result

For 3 states, 100, 200 and 300 seconds intervals, Results separately for SVM and KNN for a different sensors and features are shown in table from 3-8.

Table 4 Analysis features for 100 seconds intervals by using KNN

Features used for 100 seconds state	Number of Features	KNN Classification	Number of Sensor
ALL	78	85.71 %	5
RESP 2,3,4,9,10,14 - HR 1 - EMG 8,14 -HGSR 8,13,15 - FGSR 4,10,14,15,16	17	93.65 %	5
RESP 2,3,10,14 - HR 1 - EMG 14 - HGSR 8 - FGSR 14	8	95.23 %	5
RESP 2,10,14 - HR 1 - EMG 14 -- HGSR 8	6	93.65 %	4
RESP 2,10,14 - HGSR 8	4	87.30 %	2
RESP 2,5,9	3	92.06 %	1
ClassifierSubsetEval feature selection : EMG 3,14,16 - HGSR 7,15 - FGSR 9 - RESP 4,9,10,14	10	90.48%	4

For 100 seconds, high exactness of 98.41% was obtained utilizing SVM seclusion, each of the five sensors, and 20 highlights. Additionally, a phenomenal 96.82% precision is accomplished with a couple of highlights utilizing SVM split, 5 sensors, and 8 highlights. With the procurement of a couple of sensors and highlights, the best precision of 92.06% is acquired utilizing the KNN split, single sensor, and 3 highlights.

Table 5 Analysis features for 200 seconds intervals by using SVM

Features used for 200 seconds state	Number of Features	SVM Classification	Number of Sensor
ALL	78	93.65 %	5
RESP 2,4,5,6,9,14 - HR 1 - EMG 1,4 - HGSR 1, 8,10 - FGSR 4,6,15,16	16	98.41 %	5
RESP 2,4,14 - HR 1 - EMG 1 - HGSR 8,10 - FGSR 15	8	92.06 %	5
RESP 2,4,14 - HR 1 - HGSR 8,10	6	90.47 %	3
RESP 2,4,14 - HGSR 8,10	5	88.89 %	2
RESP 2,3,8,9,14	5	85.71 %	1
ClassifierSubsetEval feature selection : EMG 1,4 - HGSR 1,10 - HR 15 - RESP 3,14	7	92.06%	4

Table 6 Analysis features for 200 seconds intervals by using KNN

Features used for 200 seconds state	Number of Features	KNN Classification	Number of Sensor
ALL	78	88.89 %	5
RESP 3,4,5,9,14 - HR 1 - EMG 1,16 -HGSR 1,10 - FGSR 9,10,14,15	14	92.06 %	5
RESP 4,9,14 - HR 1 - EMG 1 - HGSR 1 - FGSR 10, 14	8	93.65 %	5
RESP 4,9,14 - HR 1 - EMG 1 - HGSR 1	6	90.47 %	4
RESP 4,9,14 - EMG 1 - HGSR 1	5	87.30 %	3
RESP 2,3,9	3	92.06 %	1
ClassifierSubsetEval feature selection : EMG 1,4 - HGSR 1,10 - HR 15 - RESP 3,14	7	95.23%	4

For 200 seconds interval state, the most extreme precision of 98.41% is accomplished utilizing an SVM classifier, every one of the 5 sensors, and 16 highlights. The best precision of 95.23% is accomplished for few highlights utilizing the KNN classifier, 4 sensors, and 7 highlights. For a couple of sensors and highlights, the best exactness of 92.06% is accomplished utilizing the KNN classifier, one sensor, and 3 highlights.

Table 7 Analysis features for 300 seconds intervals by using SVM

Features used for 300 seconds state	Number of Features	SVM Classification	Number of Sensor
ALL	78	80.95 %	5
RESP 3,4,5,6,9 - HR 1 - EMG 3 - HGSR 3 - FGSR 3	9	99 %	5
RESP 3,6, 9 -- HGSR 3	4	85.71 %	2
RESP 3,4,9	3	80.95 %	1
ClassifierSubsetEval feature selection : EMG 3 - HGSR 3 - RESP 9	3	95.24%	3

Table 8 Analysis features for 300 seconds intervals by using KNN

Features used for 300 seconds state	Number of Features	KNN Classification	Number of Sensor
ALL	78	76.19 %	5
RESP 3 ,9,14 - HR 1 - EMG 3 -HGSR 3 - FGSR 3	7	90.47 %	5
RESP 3, 9,14 - EMG 3 - HGSR 3	5	99 %	3
RESP 4,9,14 -HGSR 3	4	90.47 %	2
RESP 3,4,6,8,9	5	85.71 %	1
ClassifierSubsetEval feature selection : EMG 3 - HGSR 3 - RESP 9	3	95.24%	3

For 300 seconds, the greatest precision of 99% is acquired utilizing the KNN split, three sensors, and five highlights.

As demonstrated in Tables 2 to 7, better affectability with a couple of sensors and highlights is accomplished with long stretches. It should be said that the respiratory system is the main source of stress. At last, you can see the consequences of this paper and the other three papers in the table. As referenced, the aftereffects of this paper are precise, and a couple of highlights have been utilized..

Table 9 Compare results

Time int. (s)	Acc. (%)	Classifier	Sensor numbers	Features numbers	Ref.
300	97	Fisher	5	22	[3]
300	78.94	SVM	2	5	[9, 10]
300	99	SVM	5	9	paper
200	98.41	SVM	5	16	
100	98.41	SVM	5	20	

4. Conclusion

It has been shown that biological signs can distinguish feelings of anxiety by an alternate number of biological sensors, a few distinct elements, and various time intervals. Leading features are chosen for every one of the 78 features to be parted, with the greatest precision of 98.42% for 100-second intervals and 200-second spans, 99% for 300 seconds are accessible. It is shown that the main sensor for getting strain to breathe.

Utilizing more data about the human condition in various settings, we can make an example to recognize depression in various conditions and track down the specific measure of pressure that is useful to assist specialists with prescribing medication.

References

1. L. Salahuddin, "Heart rate variability analysis for mental stress measurement in mobile settings." MSc. Thesis; Korea Advanced Institute of Science and Technology (KAIST), Korea, 2007.
2. J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," Intelligent Transportation Systems, IEEE Transactions on, vol. 6, pp. 156-166, 2005.
3. J. A. Healey, "Wearable and automotive systems for affect recognition from physiology," Massachusetts Institute of Technology, 2000.
4. S. Norizam, M. Nasir, L. Sahrim, H. M. Zunairah, M. A. Siti Armiza, M. Mahfuza, et al., "Development of EEG-based stress index," 2012.
5. A. Saidatul, M. P. Paulraj, S. Yaacob, and M. A. Yusnita, "Analysis of EEG signals during relaxation and mental stress condition using AR modeling techniques," in Control System, Computing and Engineering (ICCSCE), 2011 IEEE International Conference on, 2011, pp. 477-481.
6. R. Khosrowabadi, C. Quek, K. K. Ang, S. W. Tung, and M. Heijnen, "A Brain-Computer Interface for classifying EEG correlates of chronic mental stress," in Neural Networks (IJCNN), The 2011 International Joint Conference on, 2011, pp. 757-762.

7. P. Karthikeyan, M. Murugappan, and S. Yaacob, "EMG signal based human stress level classification using wavelet packet transform," in Trends in Intelligent Robotics, Automation, and Manufacturing, ed: Springer, 2012, pp. 236-243.
8. K. Soman, V. Alex, and C. Srinivas, "Analysis of physiological signals in response to stress using ECG and respiratory signals of automobile drivers," in Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013 International Multi-Conference on, 2013, pp. 574-579.
9. Y. Deng, Z. Wu, C.-H. Chu, Q. Zhang, and D. F. Hsu, "Sensor Feature Selection and Combination for Stress Identification Using Combinatorial Fusion," International Journal of Advanced Robotic Systems, vol. 10, 2013.
10. Y. Deng, C.-H. Chu, H. Si, Q. Zhang, and Z. Wu, "An Investigation of Decision Analytic Methodologies for Stress Identification," The International Journal on Smart Sensing and Intelligent Systems (ISSN: 1178-5608), 2012.
11. P. Karthikeyan, M. Murugappan, and S. Yaacob, "Multiple Physiological Signal-Based Human Stress Identification Using NonLinear Classifiers," Electronics & Electrical Engineering, vol. 19, 2013.
12. H. Kurniawan, A. V. Maslov, and M. Pechenizkiy, "Stress detection from speech and Galvanic Skin Response signals," in Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on, 2013, pp. 209-214.
13. J. Zhai and A. Barreto, "Stress detection in computer users through noninvasive monitoring of physiological signals," Blood, vol. 5, p. 0, 2008.
14. C. Cortes and V. Vapnik, "Support-vector networks," Machine learning, vol. 20, pp. 273-297, 1995.
15. N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," The American Statistician, vol. 46, pp. 175- 185, 1992.