

# MULTIVALUE STRESS DETECTING SYSTEM USING ADT AND SVM ALGORITHM

M.Mailsamy<sup>#1</sup>, S.L.Sadhana<sup>\*2</sup>, M.Saliha Barveen<sup>\*3</sup>, and R.Swathi Lakshmi<sup>\*4</sup>

<sup>#</sup> Assistant Professor, Department of Computer Science and Engineering, Vivekanandha College of Engineering for Women, Tiruchengode, ,Tamilnadu.,India

<sup>\*</sup> Department of Computer Science and Engineering, Vivekanandha College of Engineering for Women, Tiruchengode, ,Tamilnadu.,India

**Abstract—** Physiological sensors examination is getting increasingly more significant as the accessibility of sensor-empowered versatile, wearable, and implantable gadgets gets omnipresent in the developing information investigation. Physiological multi-sensor contemplates have been led effectively to identify pressure. In this examination, we center around ECG checking that would now be able to be performed with insignificantly intrusive wearable patches and sensors, to build up a productive and hearty instrument for precise pressure recognizable proof. We track down that high pressure can be effectively recognized by AI calculations from the ECG flags alone with 100% precision. Stress is sorted as a state of mental strain or pressing factor approaches as a result of disturbing or mentioning conditions. Where different wellsprings of stress commencement scientists are think about human frontal cortex as the essential wellspring of stress.

To study how each individual encounters stress in different forms, researchers conduct surveys and monitor it. This project presents the adt-svm to enhance the accuracy for detection of mental stress using ECG signals. Whale Optimization Algorithm has been modified to select the optimal kernel in the SVM classifier for stress detection. An integrated set of algorithms (ADT SVM) has been used for pre-processing, feature extraction, and selection. The proposed algorithm has been tested on ECG signals collected from 14 subjects to identify the stress level. The proposed approach was validated using accuracy, sensitivity, specificity, and F1 score with values better accuracy and was found to be better than the existing ones. These results are obtained by looking at the variation in the ECG signals between states of rest and high stress in an individual.

**Index Terms—**About four key words or phrases in alphabetical order, separated by commas.

## I. INTRODUCTION

### A. MACHINE LEARNING

Machine learning (ML) is the study of computer algorithms that improve automatically through experience and by the use of data. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as email filtering and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks.

A subset of machine learning is closely related to computational statistics, which focuses on making predictions using computers; but not all machine learning is statistical learning. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a related field of study, focusing on exploratory data analysis through unsupervised learning. In its application across business problems, machine learning is also referred to as predictive analytics.

### B. MENTAL STRESS ANALYSIS

Mental stress has become a social problem of the 21st century. It affects the functionality of the routine work and economy of an individual human and a nation as well.

Stress can be classified as positive and negative. Positive stress alerts us and avoids the danger leading to performance enhancement while negative stress can cause mental and behavioural changes as one does not get relaxation in between various challenges.

To distinguish human feelings and stress discovery, electroencephalogram (EEG) or ECG (Electro cardio gram) is a significant technique used in stress detection and has shown

some significant results

### C. CLASSIFICATION

Classification is a process related to categorization, the process in which ideas and objects are recognized, differentiated and understood. See Classification (general theory).

### D. TIME SERIES

Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. Time series forecasting is the use of a model to predict future values based on previously observed values. While regression analysis is often employed in such a way as to test relationships between one more different time series, this type of analysis is not usually called "time series analysis," which refers in particular to relationships between different points in time within a single series. Interrupted time series analysis is used to detect changes in the evolution of a time series from before to after some intervention which may affect the underlying variable.

Time series data have a natural temporal ordering. This makes time series analysis distinct from cross-sectional studies, in which there is no natural ordering of the observations (e.g. explaining people's wages by reference to their respective education levels, where the individuals' data could be entered in any order). Time series analysis is also distinct from spatial data analysis where the observations typically relate to geographical locations (e.g. accounting for house prices by the location as well as the intrinsic characteristics of the houses). A stochastic model for a time series will generally reflect the fact that observations close together in time will be more closely related than observations further apart. In addition, time series models will often make use of the natural one-way ordering of time so that values for a given period will be expressed as deriving in some way from past values, rather than from future values (see time reversibility).

Time series analysis can be applied to real-valued, continuous data, discrete numeric data, or discrete symbolic data (i.e. sequences of characters, such as letters and words in the English language).

### E. STRESS MEDICINE

- Stress (biology), an organism's response to a stressor such as an environmental condition
- Stress (linguistics), relative emphasis or prominence given to a syllable in a word, or to a word in a phrase or sentence

Stress (mechanics), the internal forces that neighbouring particles of a continuous material exert on each other

- Occupational stress, stress related to one's job

- Psychological stress, a feeling of strain and pressure
- Surgical stress, systemic response to surgical injury.

### F. DIMENSIONALITY ON REDUCTION DATAMINING

Machine Learning techniques are the result of a long process of research and product development. This development began when commerce data was initial amass on computers, continual with development in data admission, and more just, make knowledge that agree to users to find the way through their data in real time. Machine Learning obtains this evolutionary development further than traditional data admission and navigation to potential and practical information liberation. Machine Learning is prepared for claim in the business community because it is hold up by three technologies that are now adequately mature:

- Massive data collection
- Powerful multiprocessor computers
- Machine Learning algorithms

Commercial databases are rising at unparalleled rates. A current META Group study of data storehouse projects originate that 19% of respondents are further than the 50 gigabyte level, while 59% wait for to be there by second quarter of 1996.1 in some manufacturing such as retail, these numbers be able to be a lot well-built. The supplementary need for better computational engines can now be met in a cost-effective method with similar multiprocessor computer skill. Machine Learning algorithms exemplify system that have survive for at least 10 years, but have only just been apply as grown-up, dependable, comprehensible tools that time after time outperform older arithmetical process.

In the development from commerce data to trade in sequence, every new step has constructed upon the preceding one. For example, lively data admission is significant for drill-through in data steering applications, and the capability to amass huge databases is significant to data mining.

From the user's point of view, the four steps listed in Table 1 were revolutionary because they allowed new business questions to be answered accurately and quickly.

### G. THE SCOPE OF MACHINE LEARNING

Machine Learning obtain its name from the resemblance flanked by penetrating for precious business in order in a large database for instance, finding concurrent goods in gigabytes of store scanner data and mining a ton for a vein of expensive ore. Both growths need moreover separate through a huge quantity of fabric, or cleverly searching it to discover precisely where the worth live. Given databases of enough size and excellence, Machine Learning skill can generate new trade opportunities by given that these capabilities:

- **Automated prediction of trends and behaviours.** Machine Learning automates the procedure of finding

extrapolative in sequence in huge databases. Questions that traditionally necessary wide hands-on study can now be respond straight from the data quickly. A characteristic example of a prognostic problem is under attack advertising. Machine Learning uses data on precedent promotional mailings to recognize the aim mainly likely to exploit arrival on asset in prospect mailings. Other extrapolative trouble comprises forecasting insolvency and other appearance of default, and recognizes section of a population likely to react similarly to agreed events.

- Automated discovery of previously unknown patterns.

Machine Learning tools brush during databases and recognize beforehand hidden patterns in one step. An instance of pattern finding is the study of sell sales data to recognize apparently unconnected products that are often acquire jointly. Other pattern discovery trouble comprises detecting fraudulent credit card dealings and recognizes irregular data that could symbolize data admission keying errors.

- More columns. Analyst must often boundary the figure of variables they scrutinize when action hands-on study due to time restriction yet variables that are leftover because they seem insignificant may carry in sequence about unidentified patterns. High presentation Machine Learning permits users to discover the full deepness of a file, devoid of preselecting a separation of variables.

- More rows. Overweight example acquiesce inferior estimation errors and discrepancy, and permit users to build deduction about little but significant section of a populace.

A fresh Gartner collection superior knowledge study letter listed Machine Learning and fake cleverness at the top of the five key knowledge areas that "will obviously have a major impact across a wide range of industries within the next 3 to 5 years."2 Gartner also listed corresponding architectures and Machine Learning as two of the top 10 new knowledge in which corporation will spend throughout the after that 5 years. According to a fresh Gartner HPC investigate Note, "With the fast go forward in data imprison, programme and storage, large-systems consumer will increasingly need to apply new and innovative ways to excavation the after-market worth of their vast food of feature data, utilize MPP [massively parallel processing] systems to make new source of business benefit (0.9 probability)."3

The mainly usually used processes in Machine Learning are:

- Artificial neural networks: Non-linear prognostic replica that study through education and resemble organic neural networks in association.
- Decision trees: Tree-shaped arrangement that symbolize sets of conclusion. These terminations make rules for the organization of a dataset. Explicit result tree process contain

Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID) .

- Genetic algorithms: Optimization method that uses procedure such as hereditary grouping, mutation, and natural assortment in a plan support on the impression of evolution.

Nearest neighbor method: A method that organize every evidence in a dataset support on a combination of the classes of the k record(s) most similar to it in a historical dataset (where  $k \geq 1$ ). Sometimes called the k-nearest neighbor technique.

- Rule induction: The extraction of useful if-then rules from data based on numerical meaning.

## II. PROBLEM STATEMENT :

Generally, the actual refinement is divided into two phases, called the match and the refine phases. In NMS, such a natural "evolution" process of pieces can also be controlled by users.

NMS allows users to cease, quicken, or slow down the "evolution" by setting different parameters, such as initial "vitality", fading rate, and increasing rate. Besides such macro control, users can also directly manipulate any individual piece. For example, users can mark a piece as crucial, indicating that it should never be expired from the system. They could also directly delete some useless pieces.

## III. LITERATURE REVIEW

### DETECTING STRESS DURING REAL-WORLD DRIVING TASKS USING PHYSIOLOGICAL SENSORS

Jennifer A. Healey et al., has proposed This paper presents methods for collecting and analyzing physiological data during real world driving tasks to determine a driver's relative stress level. Electrocardiogram, electromyogram, skin conductance and respiration were recorded continuously while drivers followed a set route through open roads in the greater Boston area. Data from twenty four drives of at least fifty minute duration were collected for analysis

### EVALUATING FEATURE SELECTION FOR STRESS IDENTIFICATION

Yong Deng, Zhonghai Wuet al., has proposed in modern society, more and more people are suffering from stress. The accumulation of stress will result in poor health condition to people.

### AN INVESTIGATION OF DECISION ANALYTIC METHODOLOGIES FOR STRESS IDENTIFICATION

Yong Deng, Chao-Hsien Chuet al., has proposed in modern society, more and more people are suffering from some type of stress. Monitoring and timely detecting of stress level will be very valuable for the person to take counter measures. In this paper, we investigate the use of decision analytics

methodologies to detect stress.

#### SENSOR FEATURE SELECTION AND COMBINATION FOR STRESS IDENTIFICATION USING COMBINATORIAL FUSION

Yong Deng, Zhonghai Wu et al. has proposed the identification of stressfulness under certain driving condition is an important issue for safety, security and health. Sensors and systems have been placed or implemented as wearable devices for drivers.

#### A NOVEL METHOD OF STRESS DETECTION USING PHYSIOLOGICAL MEASUREMENTS OF AUTOMOBILE DRIVERS

Mandeep Singh, Abdullah Bin Queyamet et al. has proposed Stress while driving is an important factor in many number of fatal road accidents worldwide. There has been much work done in driver stress detection. In this research, we present a method based on a correlation analysis and developed a mathematical function for the estimation of automobile driver stress level. The proposed methodology monitors driver's stress level using features extracted from selected physiological parameters. The results obtained indicate a strong correlation between the stress level of driver and the stress function formed. Threshold approach is used to perform a classification of affective states as "Low Stress", "Moderate Stress" and "High Stress" based on different traffic conditions.

#### IV. SYSTEM ARCHITECTURE:

The architecture of the proposed stress detection system is depicted in Fig 1. A stress detection system was designed using time series analysis of physiological signals such as ECG, EMG, hand GSR and foot GSR. The signals are reduced and classified

using a Support Vector machine classifier. The performance of the stress detection system is assessed in terms of accuracy, sensitivity and specificity.

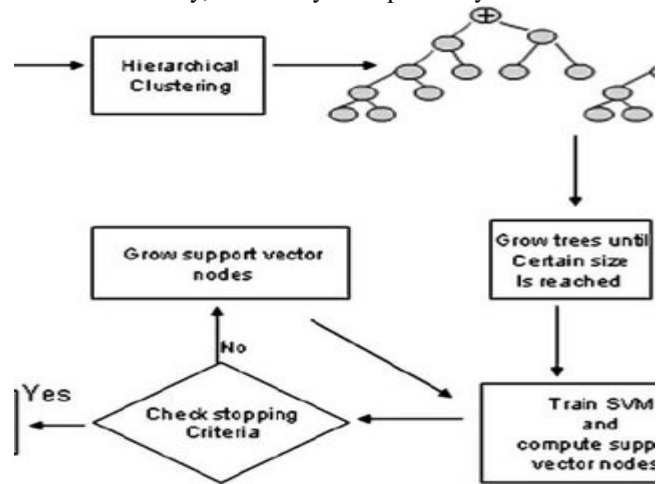


Fig. 1: System architecture of proposed stress detection

## V. EXISTING SYSTEM

In existing system, this is impractical in our scenario for two reasons. First, it may have the undesirable effect of changing a stress patient's existing test data groups, potentially undoing the Patient's own manual efforts in organizing her history. Second, it involves a high computational cost, since we would have to repeat a large number of attribute test data group similarity computations for every new test data. As existing approaches to extract stress prediction suffer from scalability, it is imperative to address the scalability issue. Connections in stress prediction are not homogeneous. This relation-type information, however, is often not readily available in stress prediction. A direct application of collective inference or label propagation would treat connections in a stress network as if they were homogeneous.

### A. DRAWBACKS

- Less security.
- Poor performance.
- Complex data processing to find stress prediction.
- The data retrieval based on user requirement is not done.

## VI. PROPOSED METHODOLOGY

The suggested method has four main phase algorithms, namely, (i) preprocessing, (ii) feature extraction, (iii) feature selection, and (iv) classification (stress level identification), besides data acquisition stage. In the proposed system, Naive Bayes algorithm and support vector machine algorithm are used to classify stressed and nonstressed user. Ranking is based on the number of stressed posts tweeted by each user. Then the user with highest stress is found. Data Preprocessing In data preprocessing, the noisy and incomplete data are removed. The twitter dataset has the user ID, stress category and user's tweet. For example: Incomplete tweet data like "u guys know y" and "very much so" are removed from the dataset.

### ADVANTAGES

- We show through comprehensive experimental evaluation the effectiveness and the robustness of our proposed search log-based method, especially when combined with approaches using other signals such as text similarity.
- We will focus on evaluating the effectiveness of the proposed algorithms in capturing test data relevance.
- Relevance Measure.

### MODULES DESCRIPTION

#### DATA ACQUISITION

The dataset used in the work was taken from. Around 14 human subjects with normal signal value of an average age of 26 years (ranging between 22 and 46) volunteered for the study.

The subjects sat in a dimly lit room, at a distance of 110 cm

from a signal having a block of 100 numerical data. Subjects were also holding a touch-sensitive button. (e 32 electrodes mounted on the ECG data attributes.) The Load method provides a technique for filling a single DataTable with data, retrieved from an IDataReader instance. This method provides the same functionality, but allows you to load multiple result sets from an IDataReader into multiple tables within a DataSet .

### FEATURE SELECTION AND SIGNAL ANALYSIS

The generic problem of supervised feature selection SVM can be outlined as follows. We aim to find a feature subset of size  $m$  which contains the most informative features. The two well-performing feature selection algorithms on the ECG dataset are briefly outlined below. This module is used to filter that ranks the input attributes according to their relevance. A cutting rule enables the selection of a subset of these attributes. It is required to define the target attribute which in this domain of research applies to the nature of the stress (Normal/Abnormal) and the predictor attributes. After computing the Fisher score for each feature, it selects the top- $m$  ranked features with large scores. The next subsection directs focus on another technique of feature selection using ADT Process.

### CLASSIFICATION SCORING

When the number of descriptors is very large for a given problem domain, a learning algorithm is faced with the problem of selecting a relevant subset of features backward regression includes regression models in which the choice of predictor variables is carried out by an automatic procedure.

The iterations of the algorithm for logistic regression are given in steps as stated as follows .

Step 1: The feature set with all 'ALL' predictors.

Step 2: Eliminate predictors one by one.

Step 3: 'ALL' models are learnt containing 'ALL-1' descriptor each.

These iterations are further continued till either a pre-specified target size is reached or the desired performance statistics (classification accuracy) is obtained. After feature relevance.

We classify the nature of the ECG cases in the ECG dataset using ADT SVM classification algorithm.

### ADT\_SVM VALIDATION

The following steps illustrates ADT-SVM as follows,

- input:  $G$  = set of antigens to be recognized,  $n$  the number of worst elements to select for removal
- output:  $M$  = set of memory attributes capable of classifying unseen patterns
- begin
- Create an initial random set of attributes,  $A$  for all antigens in  $G$  do
- Generate clones of a subset of the attribute in  $A$  with the highest predicted.
- The number of clones for an attribute is proportional to its predicted.
- Mutate attributes of these clones to the set  $A$ , and place a copy of the highest predicted
- Replace the  $n$  lowest predicted attributes in  $A$  with new randomly generated attributes.

• end

## VII.CONCLUSION

In conclusion, we have obtained a perfect 100% accuracy in detecting high stress levels when the difference in average heart rate feature is used for analysis, and 98% with all the eight features. Thus, the accuracy of detecting high stress based on individual variations in ECG extracted features is higher than that previously published results detecting stress based on stress state alone.

False negative rate with ADT-SVM method was perfect in all experiments. Future works include adding the T-wave related features in our analysis since ECG studies have shown that QT is an important biomarker of cardiac abnormality and adding other signals.

## REFERENCES

- [1] J. Healey, R.W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *Intelligent Transportation Systems, IEEE Transactions on*, 2005, vol. 6, no 2, pp. 156-166.
- [2] DENG, Yong, WU, Zhonghai, CHU, Chao-Hsien, et al. Evaluating feature selection for stress identification. In : *Information Reuse and Integration (IRI), 2012 IEEE 13th International Conference on. IEEE, 2012. p. 584-591.*
- [3] DENG, Yong, CHU, Chao-Hsien, WU, Z., et al. An Investigation of Decision Analytic Methodologies for Stress Identification. *The International Journal on Smart Sensing and Intelligent Systems (ISSN: 1178-5608)*, 2012.
- [4] DENG, Yong, WU, Zhonghai, CHU, Chao-Hsien, et al. Sensor feature selection and combination for stress identification using combinatorial fusion. *International Journal of Advanced Robotic Systems*, 2013, vol. 10, p. 306-313.
- [5] SINGH, Mandeep et QUEYAM, Abdullah Bin. A novel method of stress detection using physiological measurements of automobile drivers. *International Journal of Electronics Engineering*, 2013, vol. 5, no 2, pp. 13-20
- [6] SINGH, Mandeep et QUEYAM, Abdullah Bin. Correlation between Physiological Parameters of Automobile Drivers and Traffic Condition. *International Journal of Electronics Engineering*, 2013, vol. 5, no 2, pp. 6-12
- [7] AVCI, Cafer, AKBAŞ, Ahmet, et YÜKSEL, Yusuf. Evaluation of Statistical Metrics by Using Physiological Data to Identify the Stress Level of Drivers.
- [8] BEGUM, Shahina. Intelligent driver monitoring systems based on physiological sensor signals: A review. In : *Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on. IEEE, 2013. p. 282-289.*
- [9] ZHOU, Yingchun et SEDRANSK, Nell. Functional data analytic approach of modeling ECG T-wave shape to measure cardiovascular behavior. *The Annals of Applied Statistics*, 2009, p. 1382-1402.
- [10] Mandeep Singh and Abdullah Bin Queyam, "A Novel Method of Stress Detection using Physiological Measurements of Automobile Drivers," *International Journal of Electronics Engineering*, vol. 5, no. 2, pp13-20, Dec 2013.