

MACHINE LEARNING BASED CARDIOVASCULAR DISEASE DETECTION

¹A.SANDHYA, ²Dr. MOHAMMED ABDUAL HAFEEZ

¹Assistant Professor, Dept. of CSE, Megha Institute of Engineering and Technology for Women,
sandhyaanaveni@gmail.com

²Assistant Professor, Dept. of CSE, Megha Institute of Engineering and Technology for Women,
mdabdl.hafeez@gmail.com

Abstract: *Cardiovascular disease (CVD) is the leading cause of loss of life and disability worldwide. Knowledge of cardiovascular risk factors can also provide important information about the prevention, etiology, management, and treatment of this important public health problem. Obese humans are more likely to have cardiovascular disease (CVD), which is thought to be caused by a change in the activity of the autonomic nervous system (ANS). Changes in ANS activity can be detected using heart rate variability (HRV). HRV is a non-invasive tool for classifying ANS interests using linear and non-linear HRV functions. This article aims to understand the effects of obesity on ANS using HRV parameters. Initially, 16 control subjects and 16 obese subjects of each sex between the ages of 20 and 50 were included when the synthetic minority oversampling technique (SMOTE) was used to increase the sample size of subjects control and overweight 16 to 48. Independent tests determined the statistical differences between the agencies. The study's statistical results suggest a sympathetic imbalance due to a decrease in parasympathetic activity Machine Learning (ML) algorithm determines the most significant visionary that can distinguish between control and obese subjects.*

Keywords: *Obesity, Cardiovascular disease (CVD), Heart rate variability (HRV), Autonomic nervous system (ANS).*

I. INTRODUCTION

Recent studies have strongly warned that poor emotional states, tendencies, and work stress are associated with illness and poor physical fitness [1]. Employment stress, in particular, has been linked to

major financial implications, including increased absenteeism, increased employee turnover, decreased employee activity satisfaction, and an associated decrease in worker productivity. Work stress is also a major public health risk related to heart disease. Cardiovascular disease (CVD) is

the leading cause of illness and death in men and women. It is especially true in developed countries [2]. A wide range of CVD risk factors have been identified, but no consensus model has been proposed that could lead to such risk factors. A recent Whitehall Study article showed that work stress is associated with decreased coronary heart rate variability (HRV). HRV deficiency is a neutral risk factor for disease and death. The important role of the vags nerve in health and disease has long been considered. However, the simplest thing these days is that researchers and therapists have begun researching how can integrate this data into more information about the etiology, symptoms, pathways, outcomes, and treatment of the disease. This exemplary review shows that autonomic imbalance, which lacks vaginal blocking effects, is associated with an increase in disease and mortality of all causes. We also look at evidence linking vaginal features to increasing and emerging risk factors, such as work pressure, heart disease, and death. Importantly, we are discussing the evidence that the factors that increase HRV are related to risk reduction and improved fitness profiles. Therefore, the autonomic imbalance version may also provide a unified technique to understand the risk of heart disease and the role of HRV in all causes of death.

Autonomic imbalance and disease

Evidence is emerging on the position of the autonomic nervous system (ANS) in various diseases. The main branches of the ANS are generally thought to be: the sympathetic machine, which is associated with electrical stimulation, and the parasympathetic machine, which is associated with vegetative and regenerative abilities. In general, these branches are interested in dynamic balance. However, it can rapidly modulate the activity of both branches in response to changing environmental requirements. The concepts of the characteristics of living things based on complexity maintain that the balance, adaptation, and health of organisms are maintained through changes in the dynamic relationship between the instrument factors [3]. Thus, given the ever-changing environmental requirements, patterns of systematic variability rather than static phases are preserved. Because the device "works at some distance from equilibrium," the device typically seeks to find the local force minima to reduce the organism's energy needs.

Another consequence of this approach is that autonomous imbalances, where an ANS department dominates, are related to a lack of dynamic flexibility and fitness. Experimentally, a large body of evidence indicates that autonomic imbalance, in

which the sympathetic apparatus is usually hyperactive and the parasympathetic apparatus is hypoactive, is related to various pathological conditions. In particular, when the sympathetic sector dominates for a long time, the electrical needs in the machine become excessive and ultimately cannot be met, resulting in death. On the road to death, however, premature aging and disease represent a management mechanism of independent imbalances. Thus, an imbalance of autonomy may be the ultimate common way to reduce illness and death for various conditions and diseases, including heart disease. Heart rate fluctuations can be used to independently estimate the risk of imbalance, disease, and death. Parasympathetic activity and HRV have been linked to immune disorders and inflammation, leading to heart disease, diabetes, osteoporosis, arthritis, Alzheimer's disease, periodontal disease, certain types of cancer, and muscle loss. Are attached. Strength and height Weakness and disability. Heart rate variable (HRV) measures in time and rate domains were effectively used to index vaginal activity. In the time zone, the same previous deviation of internet distances (IBI), total deviation of R to R duration (SDNN), base rectangle consecutive difference (RMSSD), and baroreflex sensitivity measures (an index) are

suggested. The cardiovascular system's response to changes in blood pressure is a beneficial indicator of vaginal attachment.

II. REVIEW OF LITERATURE

This illustrative review shows that autonomic imbalance, which lacks vaginal blocking effects, is associated with rapid disease and mortality. We also look at the evidence that links vaginal function to possible installed and emerging factors, including blood pressure, heart disease, and death. Importantly, we are discussing the evidence that the factors that increase HRV are related to risk reduction and a better health profile. Therefore, the model of autonomic imbalance may also offer a coherent technique for understanding the role of HRV in the risk of heart disease and all causes of death.

Heart rate variability and mortality

In one of the first studies to investigate the relationship between indices of HRV and mortality,

Kleiger et al. [4] Approximately 900 patients with myocardial infection (MI) were confirmed to be a significant neutral predictor of HRV death in this high-risk group. Numerous studies have supported the idea that reduced vaginal time, as configured by HRV, predicts death in a more likely, low-risk population. In an

older sample of the Framingham Heart Study (FHS), frequency domain measures have been linked to aggregate mortality after overcoming dramatically different risk factors. Seven hundred thirty-six women and men with an average age of 72 years provided ambulatory HRV data in the time and frequency domain [22]. Eight HRV measurements have been tested, including five frequency area measurements. All five frequency domain measures have been associated with commendable mortality due to commendable causes other than the LF / HF ratio (a significant degree of sympathovagal stability where a higher number suggests additional relative sympathetic dominance). Risk factors remain after control. A standard deviation (SD) difference in log-transformed LF power was associated with an additional 1.7 times relative difficulty in each cause of death in this pattern.

Similarly, in Hoorn's study, a possible test of glucose tolerance in the modern population, the multiple time-frequency domain index of HRV was calculated, and five were associated with all causes of death during 9-12 months. Age, sex, and glucose tolerance after a period of at least pb0.10 phase increase after control. The discovery became more powerful for

people at high risk for diabetes, high blood pressure, or heart disease.

Liao et al. [5] The ARIC study examined the relationship between two-minute sleep apnea HRV and hypertension in a randomized, randomized pattern of 2,061 black and white men and women. Only 64 people had high blood pressure during the three years of follow-up. However, baseline HF power reversal was associated with an improvement in those with high blood pressure. In cross-sectional analysis, HF power adjusted for age, race, sex, smoking, diabetes, and training decreased significantly compared to the hypertension group (both untreated and untreated). Furthermore, those living within the lowest quarter of HRV had a 2.44 times higher risk of developing high blood pressure than those living within the highest quarter.

Singh et al. [6] Cross-sectional and prospective analyzes tested the association between hours of recording ambulatory heart rate and hypertension in women and men. Cross-sectional analysis indicates that, when age, BMI, smoking, and alcohol use are adjusted, multiple measures of the HRV time-area-frequency index are dramatically higher in hypertension than in normal women and men. During the 4-year follow-up period, 119 boys and 125 girls developed hypertension. These analyzes

showed that low-frequency low electricity was associated with the development of hypertension in men but not in women anymore.

Singh et al. [7] examined the relationship between HRV and blood glucose levels in 1,919 women and men. The first two hours of the ambulatory heart rate recording were used to calculate some of the HRV's time-frequency domain index. Fasting glucose levels have been used to classify people with normal or poor fasting glucose and select people with diabetes (similar to those diagnosed with diabetes). Numerous indicators of HRV, including the strength of LF and HF within the FHS, are inversely related to fasting glucose levels and are more pronounced in diabetic patients and those who are weak in fasting than in those with fasting glucose limits.

Hyano et al. [8] noted that both acute and chronic smoking was associated with decreased vaginal tone. Similarly, smoking was associated with significantly higher LF / HF ratios in taxi drivers within five minutes of exposure to daily driving conditions. Smoking at night, in particular, is thought to have a more severe effect on cardiac modulation than smoking during daylight hours. The authors also suggest that the sympathetic and parasympathetic withdrawal response from smoking may also contribute to heart disease risk.

III. PROPOSED METHODOLOGY

A. Subject

This examination has been completed with the permission of the Dean of Research and Development, College of Engineering, Pune, after following all ethical recommendations for the work of Stage Studies of the Institute. Investigators and subjects have volunteered. The test consisted of obtaining 16 normal obese and 16 controlled subjects' electrocardiograms (ECG) between the ages of 20 and 50 of both sexes. However, sixteen controls and coarse samples are insufficient to test statistical results. Therefore, we artificially elevated the length of control and coarse subject patterns using the Synthetic Minority Oversampling Technique (SMOTE) [4]. This is a strong and widely used method. Create a random set to balance the minority beauty. Samples of the new artificial records are randomly drawn between selected minority beauty patrons and their close friends. Details of the implementation of SMOTE technique are given in Algorithm 1.

B. Criteria to decide

According to the World Fitness Organization (WHO) guidelines, the use of BMI determines obesity. BMI is calculated as $BMI = \text{weight (kg)} / \text{height (m}^2)$ [10].

Subjects with a BMI value of 18 to 25 (kg / m²) are considered normal or not overweight, and subjects with a BMI value above 30 (kg / m²) are considered obese.

C. ECG Recording and HRV Analysis

ECG of 16 obese and controlled subjects was recorded using a state-of-the-art ECG device with a sampling rate of 500 Hz at rest for 15 minutes. The last five minutes of HRV analysis are used inside the Biomedical Workbench LabVIEW Cardiac Load Variability Analyzer. HRV parameters are determined using a linear and non-linear approach [7]. The linear form involves the assessment of the area of time and the area of frequency. In the time domain, the RR c program language's period sign extracts statistical parameters including implicit HR, implicit RR, SDNN, and RMSSD. Implicit HR and suggested RR are common values for coronary heart rate and RR interval. The SDNN represents the standard deviation of the normal to normal (NN) interval, and the RMSSD represents the average rectangular base of the corresponding previous deviation of the NN interval. The Fast Fourier Transform (FFT) technique calculated the HRV frequency area parameters. Frequency domain parameters that can be extracted from HRV are general power (TP) (ms²), low frequency (LF) and high frequency (HF) power, low

frequency (LF) and high frequency (HF) in normalized units. , And LF / HF ratio. Two nonlinear HRV properties are derived from the HRV signal, SD1 and SD2. SD1 and SD2 are Poincaré chart measures. SD1 represents a short-term change in the NN c programming language, and SD2 NN c is a long-term change in the programming language. The control and coarse group were analyzed based on linear and nonlinear HRV parameters.

D. Machine Learning Algorithm

While analyzing the control group and the overweight group using statistical tests, we found that the linear and nonlinear parameters of HRV show a big difference. But the key prediction that could tell the coolest company became the use of statistical control. Therefore, we have used a nonlinear machine study algorithm to detect significant predictions between weight and control issues. In this test, we have used nonlinear systems control algorithms, namely, the Classification and Regression Tree (CART) and the Gradient Boosting Discussion Tree (GBDT).

Both algorithms are used to discover important predictions. The key predictor became the use of feature significance ratings. The importance of features underscores the importance of each ability. A feature score of more than 90% is

considered an important predictor. The key predictive function can be provided as an input to the ML algorithm, and its overall performance is evaluated using the performance matrix.

F. Performance Metrics

Six classification quality evaluation measures were used, including accuracy, sensitivity, specificity, accuracy, F1-score, and the recipient's working characteristic curves (AUC). These class measures are calculated using the following confusion matrix.

Table.1 Confusion Matrix

		Actual Value	
		Positive	Negative
Predicted Value	Positive	TP	FP
	Negative	FN	TN

Where, TP- True Positive, FP- False Positive, FN- False Negative, TN-True negative

Accuracy: Accuracy is the ratio of the total number of instances of the correct prediction. Accuracy calculated as follows

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Sensitivity: Sensitivity is used to determine the portion of the actual positive instances case classified adequately by the classifier. Sensitivity calculated as follows

$$Sensitivity = \frac{TP}{TP + FN}$$

Specificity: Specificity is used to know the ability of classifiers to identify incorrectly classified negative cases

$$Specificity = \frac{TN}{TN + FP}$$

Precision: Precision is an indicator that defines the true portion of the instances when predicted to be true. Precision calculated as follows-

$$Precision = \frac{TP}{TP + FP}$$

F1 Score: F1 Score is a harmonic mean of recall and precision. It must be one for good performance and zero for the bad performance of the classification algorithm. F1 score calculated as follows

IV. RESULTS AND DISCUSSION

ML algorithms are used to find the most important predictor that separates obese subjects from their control. However, within the statistical analysis, it was found that most of the time, domain, frequency domain, and nonlinear HRV parameters have decreased significantly but no longer

provided any significant predictions. The key predictor can be found using the feature importance perspective, which provides each feature's significance score. Feature Significance Rating indicates 0.9 or more than 90% of key predictions.

This test found that the proposed RR, LF: HF, and HF (ms²) ratios became the most important predictors. We have taken the advice of RR and LF: HF using the CART algorithm became the main predictor of HF, while HF (ms²) became the main predictor obtained using GBDT. We've used these predictions to input the CART and GBDT ML algorithms. When we applied the average RR and LF: HF ratio as input to the CART algorithm, we found 96.55% accuracy, 100% sensitivity, 92.86% specificity, 93.75% accuracy, F1 score of 0.96%. AUC of 0.96. When using HF (ms²) as an input in the GBDT principle set, we set the F1 rating to 0.93 with 93.10% accuracy, 93.33% sensitivity, 92.86% specificity, 93.33% accuracy, and AUC of 0.92 . The key predictor indicates that the CART and GBDT ML algorithms can classify overweight and control subjects with an accuracy of 96.55% and 93.10%, respectively (Table I).

Algorithm	Important HRV Predictor	Feature Importance Score	AC (%)	SE (%)	SP (%)	PR (%)	F1 Score	AUC
CART	mean RR	0.93	96.55	100	92.86	93.75	0.96	0.96
	LF/HF	0.94						
GBDT	HF(ms ²)	0.95	93.10	93.33	92.86	93.09	0.93	0.92

This proposed that obesity moderates the purpose of ANS, as mean RR, HF (ms²), and LF: HF ratio is condensed in obese subjects. Thus variations in ANS modulate cardiac activity.

V. CONCLUSION

In this paper we have tried to provide an overview of some of the evidence for the role of HRV in cardiovascular disease risk and mortality. In the present study, we have used short-term HRV. Analysis of obesity and control subjects to study the effects of Obesity in ANS We has used real and artificial HRV data. The statistical results suggest shows a significant decrease in HRV. Control and parameters of coarse articles compared to machine learning algorithms were used to find the main HRV. The predictor suggests a change in the statistical results. Sympathetic balance due to low parasympathetic activity. Furthermore, this was confirmed by the use of CART and GBDT. The algorithm which showed a rating accuracy of 96.55% And 93.10, respectively.

REFERENCES

- [1] Shubham Debnath, Stavros.Z, 2021, "A method to quantify autonomic nervous system function in healthy, able-bodied individuals",
- [2] Ana Paula, Fabio Santos L, 2021, "Role of Body Mass and Physical Activity in Autonomic Function Modulation on Post-COVID-19 Condition", 19(4), 2457
- [3] Rossi RC, de Abreu LC, "Impact of obesity on autonomic modulation, heart rate and blood pressure in obese young people". *Autonomic neuroscience*. 2015; 193:138-41
- [4] Billman GE, Trimmel K, "An introduction to heart rate variability: methodological considerations and clinical applications". *Front Physiol*. 2015;6:55.
- [5] Friedman BH, 1998, "Autonomic balance revisited: panic anxiety and heart rate variability". *J Psychosom Res* ;44:133–51.
- [6] Kang MG, Chang SJ, "Association between job stress on heart rate variability and metabolic syndrome in shipyard male workers". *Yonsei Med J* 2004;45: 838–46.
- [7] Kiecolt-Glaser JK, Glaser R, "Emotions, morbidity, and mortality: new perspectives from psychoneuroimmunology". *Annu Rev Psychol* 2002;53: 83–107.
- [8] Krantz DS, McCeney MK, "Effects of psychological and social factors on organic disease: a critical assessment of research on coronary heart disease". *Annu Rev Psychol* 2002;53:341–69.
- [9] Shenoy AR, Prakash VS, "Impact of obesity on cardiac autonomic functions in middle aged males". *National Journal of Physiology, Pharmacy and Pharmacology*. 2014
- [10] Prasadu Peddi (2019), *Data Pull out and facts unearthing in biological Databases, International Journal of Techno-Engineering, Vol. 11, issue 1, pp: 25-32.*
- [11] Prasadu Peddi (2019), "AN EFFICIENT ANALYSIS OF STOCKS DATA USING MapReduce", ISSN: 1320-0682, Vol 6, issue 1, pp:22-34.