International journal of application or innovation in engineering and management

Ijaiem.com/June 2025/ Volume 14/Issue 1/Article No-1/887-892

ISSN: 2319-4847

Using Heart Rate Variability Features as a Machine Learning Approach to Investigate the Obesity Effect on the Autonomic Nervous System

[1] S.Gouthami, [2] N.Ankitha,

[1]Assistant Professor, Megha Institute of Engineering & Technology for Women, Ghatkesar.
[2] MCA Student, Megha Institute of Engineering & Technology for Women, Ghatkesar.

Article Info

Received: 30-04-2025 ***Revised: 06 -06-2025*** ***Accepted: 17-06-2025*** ***Published:28/06/2025***

Abstract

A change in autonomic nervous system (ANS) activity is one hypothesis that tries to explain why obese persons have an increased risk of cardiovascular disease (CVD). Monitoring heart rate variability (HRV) allows one to detect shifts in ANS activity. It is possible to quantify ANS activity non-invasively using linear and non-linear HRV features. To better understand the effects of obesity on the autonomic nervous system, this study sought to measure HRV. Synthetic minority oversampling (SMOTE) was used to increase the number of control and obese persons from sixteen to forty-eight. The initial sample for the research consisted of sixteen women and sixteen men, with ages ranging from twenty-five to fifty. When we used the Independent t test to compare the two datasets, we discovered a statistically significant difference. According to the study, a decrease in parasympathetic activity leads to an imbalance in sympathovagal tone. Using the machine learning approach allowed us to corroborate the study's statistical findings. To differentiate between healthy-weight and overweight individuals, we may use a Machine Learning (ML) approach to identify the most relevant predictor. Changes to the sympathovagal balance are caused by reduced parasympathetic activity, according to the results of the ML algorithm and statistical analysis.

Keywords

The synthetic minority oversampling approach, obesity, machine learning, the autonomic nervous system, and SMOTE are some of the topics covered.

I. INTRODUCTION

This is one of the main reasons why people who are overweight die. Diabetes, hypertension, heart disease, and stroke are all increased risks in those who are overweight, which is defined by a persistent buildup of fat around the midsection and lower extremities. Obesity is strongly associated with an increased risk of cardiovascular disease, according to many studies [1]. An imbalance of autonomic activity increases the risk of CVD in individuals who are obese, according to the research [2]. When everything is running well, the autonomic nervous system (ANS) keeps the body in a steady state known as homeostasis. The autonomic nerve system (ANS) regulates the function of several internal organs, including the glands and blood vessels. The SNS and the PNS are two separate systems; the PNS includes half of the autonomic nervous system and is distinct from the SNS. Parasympathetic neural system (PNS) regulation of the rest-and-digest response helps save energy, whereas sympathetic nervous system (SNS) activation during fight-or-flight state provides energy. As the vagal nerve mediates communication between the ANS and the heart, heart rate variability (HRV) quantifies the extent to which the ANS influences the heart rate. Alterations to the ANS have a direct impact on the heart rate. An electrocardiogram's (ECG) RR interval may show variations in heart rate variability, or HRV. Also, HRV may turn out to be the least invasive way to study how fat affects ANS. Obesity increases the risk of CVD due to the substantially reduced HRV [2,3]. Both linear and non-linear HRV characteristics may be used to represent the variability in the organs controlled by the autonomic nervous system (ANS). Here is the paper's outline: Methodology Methods for machine learning, subjects, obesity criteria, and statistical testing are covered in Section II. The outcomes of the statistical analysis and machine learning are described in Section III. Part IV concludes the piece.

II. METHODOLOGY

Part One. Topic Following all applicable ethical guidelines, this project received approval from the College of Engineering Pune's Dean of Research and Development. Institutional level research was the only focus of the study. All subjects, including the subjects themselves, have given their informed permission to participate in the study. The research included electrocardiogram (ECG) recordings from sixteen control volunteers ranging in age from twenty to fifty, sixteen normotensive obese persons, and sixteen individuals without hypertension. We cannot draw any firm inferences from the results of the sixteen-person control and obese groups. We used the Synthetic Minority Oversampling Technique (SMOTE)[4] to artificially increase the number of control and obese participants in our sample. This strategy has a lot of fans and has been rather fruitful. To ensure that underrepresented groups are fairly represented, it generates random samples. The samples of the minority group are combined with those of its closest neighbors to create new synthetic data sets. For further details on how the SMOTE approach is put into action, see to Algorithm 1.

Algorithm 1 SMOTE Algorithm

Input : Dataset $D, \{y_i \epsilon T\}$ where $i = 1, 2, \dots T$
  Number of minority samples (T)
  SMOTE Percentage (P)
  Number of nearest neighbors (k)
Output : Synthetic Samples
  for $i = 1, 2, \dots T$ **do**
   Find the k-nearest neighbors of y_i
   $\hat{P} = \left[\frac{P}{100}\right]$
   while $\hat{P} \neq 0$ **do**
    Select randomly k-nearest neighbors $y_{n(i)}$
    Choose randomly $\delta \epsilon [0, 1]$
    $y_{new} = y_i + \delta\,(y_{n(i)} - y_i)$
    $T \leftarrow T + y_{new}$
    $\hat{P} = \hat{P} - 1$
   end while
  end for

B. Elements that impact According to the World Health Organization (WHO), a body mass index (BMI) of 30 or more is considered obese. The body mass index (BMI) is calculated by dividing the square of your height by the weight in kilograms [10]. Those whose BMI is between the healthy range of 18–25 (kg/m2) are not obese, while those whose BMI is more than 30 (kg/m2) are. Section C: Electrocardiogram Recording and Analysis of Heart Rate Variability (HRV) A total of sixteen patients, including both lean and obese individuals, were monitored with an electrocardiogram (ECG) using industry-standard 500 Hz sampling equipment for fifteen minutes while they slept on their backs. Using Biomedical Workbench LabVIEW's heart rate variability analyzer, the most recent five minutes were examined. There were two methods used to find the HRV parameters: linear and non-linear [7]. Two domains, time and frequency, are analyzed in the linear approach. Statistical metrics like RMSSD, SDNN, mean HR, and mean RR may be extracted from the time domain by using the RR interval signal. When we average the RR interval and the heart rate, we get the mean RR and mean HR, respectively. Standard deviation of the normal to normal (SDNN) interval is different from root-mean-square of the standard deviation of the normal to normal (NN) interval (RMSSD). The frequency domain characteristics of the HRV were found by using the Fast Fourier Transform (FFT) method. Total power (TP) (ms2), power at low frequencies (LF) and high frequencies (HF), LF and HF in normalized units, and the LF/HF ratio are among the frequency-domain metrics that may be retrieved from HRV. It is possible to extract two non-linear characteristics, SD1 and SD2, from the HRV signal. The SD1 and SD2 measures are the building blocks of the Poincare Plot. When looking at the NN interval, the short-term variability is represented by SD1, while the long-term variability is represented by SD2. Straight and non-linear HRV characteristics were used to assess the control and obese groups [8,9,10,11,12]. Part D: Data Analysis The results of the Independent t test showed that the control and obese groups were significantly different from one another. The findings are presented by taking the mean and either adding or subtracting the standard deviation. Statistical significance was defined as a p-value less than 0.05. Methods for Deep Learning (Part E) The most startling result was a statistically significant difference between the obese and control groups with respect to both linear and non-linear HRV characteristics. No statistically significant predictor that might have distinguished the lean from the obese groups was found. We identified important variables that adjust for obesity and risk factors by using a non-linear machine learning approach. Several non-linear ML methods, including Gradient Boosting Decision Tree (GBDT) [14] and Classes and Regression Tree (CART) [13], were used in this research.

It is possible to find valuable predictors using any of these approaches. The feature significance score let us identify the leading predictor. To quantify the significance of each feature, we use the metric of feature importance. A significant predictor is defined as a characteristic with a significance value greater than 90%. The main predictive attribute will be fed into the ML system, and its output will be assessed using performance measures. First Section: Key Success Criteria The accuracy, sensitivity, specificity, precision, F1 score, and square of the receiver operating characteristic curve (AUC) were the six quality assessment metrics used for segmentation. Using the following confusion matrix, all of these classification metrics were generated.

TABLE I. CONFUSION MATRIX

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

Positive, True, False Negative, and True Negative are the four possible outcomes: TP, FP, FN, and TN. Here are the specifics of the performance measure: 1. The proportion of accurate predictions to total occurrences is called accuracy. Here is how the accuracy was calculated:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \dots (1)$$

2. The sensitivity of a classifier is defined as the percentage of real positive occurrences that it correctly categorized. Method for calculating sensitivity:

$$Sensitivity = \frac{TP}{TP + FN} \quad \dots (2)$$

Third, classifiers' specificity tells us how well they can spot false negatives.

$$Specificity = \frac{TN}{TN + FP} \quad \dots (3)$$

4. Accuracy 4. The true fraction of the occurrences is defined by this indication when it is projected to be true. Here is how precision is calculated:

$$Precision = \frac{TP}{TP + FP} \quad \dots (4)$$

Fifth, the F1 Score is the sonic average of the accuracy and recall. If the classification method is doing well, the result should be 1, and if it is not, the result should be 0. We compute the F1 score as follows:

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad \dots (5)$$

Another key statistic for evaluating the effectiveness of machine learning algorithms is the area under the receiver operating characteristic (ROC) curve. If the area under the curve (AUC) is close to 1, then the machine learning model is doing perfectly; if it is close to 0.5, then it is performing poorly [15].

III. RESULTS AND DISCUSSION

A. HRV Variables for Time-Domain Data Analysis Being overweight was associated with a significant decrease in time-domain HRV metrics including RR, SDNN, and RMSSD. The variance and variability of the RR interval time series signals are reduced when both the mean RR and SDNN go down. Reducing the RMSSD value (Table II) indicates a reduction in parasympathetic activity.

TABLE II. TIME DOMAIN HRV PARAMETERS

Time Domain HRV Features	Healthy Group (n=48)	Obese Group (n= 48)	P-Value
mean HR	67.33 ± 9.82	72.10 ± 10.51	0.0237
mean RR	855.17 ± 124.75	798.69 ± 116.47	0.0241
SDNN	57.42 ± 8.49	49.90 ± 7.50	0.0001
RMSSD	47.44 ± 7.25	40.23 ± 6.11	0.0001

frequency spectrum What HRV Is Worth The frequency domain variables TP(ms2), HF(ms2), and the LF: HF ratio were slightly lower in the obese group compared to the control group, whereas LF(ms2) and LF(nu) were similar. As TP (ms2) values decrease, volatility decreases as well. The parasympathetic tone is lower when the HF (ms2) value is low. An autonomic imbalance is indicated by a lower value of the LF: HF ratio in Table III.

TABLE III. FREQUENCY DOMAIN HRV PARAMETERS

Frequency Domain HRV Features	Healthy Group (n=48)	Obese Group (n= 48)	P-Value
TP (ms^2)	3265.70 ± 537.58	2543.04 ± 451.68	0.0001
LF (ms^2)	948.12 ± 156.55	926.26 ± 171.12	0.5154
HF(ms^2)	1011.06 ± 171.96	637.48 ± 111.17	0.0001
LF (nu)	48.07 ± 7.16	57.11 ± 8.36	0.0001
HF (nu)	49.94 ± 7.42	40.88 ± 6.07	0.0001
LF:HF	1.59 ± 0.26	1.08 ± 0.16	0.0001

A Look at the Non-linear HRV Settings The control group had higher non-linear HRV values than the obese group, which had lower SD1 and SD2. Table IV shows that the short-term variability of HRV signals was lower in obese individuals, as shown by their considerably lower SD1 feature values.

TABLE IV. NON LINEAR HRV PARAMETERS

Non-linear HRV Features	Healthy Group (n=48)	Obese Group (n= 48)	P-Value
SD1	33.72 ± 5.07	28.11 ± 4.30	0.0001
SD2	66.00 ± 9.62	65.83 ± 9.70	0.9303

Discoveries Made via Machine Learning (B) In order to distinguish the obese from the control group, we use machine learning approaches to identify the most relevant predictor. Statistical analysis, however, revealed a dramatic reduction in non-linear HRV properties, as well as features in the time and frequency domains, and rendered them useless as predictors. By giving each characteristic a score, the feature significance approach facilitates the selection of the most relevant predictor. To be considered the most important predictor, a trait must have a significance value higher than 0.90, which is 90%. Important factors in this research were HF (ms2), mean RR, and LF to HF ratio. The CART method produced mean RR and LF: HF, although HF (ms2) was the leading predictor according to the GBDT. The GBDT and CART ML algorithms are fed just these predictions. By feeding the CART technique the mean RR and LF: HF ratio, we were able to get a 96.55% accuracy, 100% sensitivity, 92.86% specificity, 93.75% precision, F1 score of 0.96, and AUC of 0.96. The following metrics were measured when HF (ms2) was input into the GBDT method: area under the curve (AUC) 0.92, F1 score 0.93, specificity 92.86%, precision 93.33%, and accuracy 93.10%. According to Table V, the CART ML method has a 96.55% accuracy rate for detecting obese persons, whereas the GBDT ML technique gets a 93.10% accuracy rate.

TABLE V. EVALUATION OF IMPORTANT HRV PREDICTOR

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

Low average RR, HF (ms2), and the LF: HF ratio in obese people is indicative of impaired ANS function. Accordingly, the heart's output is affected by the autonomic nervous system.

IV. CONCLUSION

This study contrasted control subjects with overweight individuals using short-term HRV analysis to examine the effects of obesity on ANS. This study made use of both real and synthetic HRV data. The findings of the inquiry were shown via the use of statistical tests and a machine learning technology. Using a machine learning method, a critical predictor of HRV was identified. Results showed that compared to the control group, patients who were overweight had much lower HRV values. As a result of reduced parasympathetic activity, the results show that the sympathovagal balance has changed. Also, the GBDT algorithm (which had a 93.10% classification accuracy) and the CART algorithm (which had a 96.55% classification accuracy) both verified this.

REFERENCES

[1] Hurt RT, Kulisek C, Buchanan LA, McClave SA. The obesity epidemic: challenges, health initiatives, and implications for gastroenterologists. Gastroenterol Hepatol (N Y). 2010;6(12):780792.

[2] Thayer JF, Yamamoto SS, Brosschot JF. The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors. Int J Cardiol. 2009; 141(2):122131.

[3] Rossi RC, Vanderlei LC, Gonalves AC, Vanderlei FM, Bernardo AF, Yamada KM, da Silva NT, 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] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling

technique. Journal of artificial intelligence research. 2002 ;16:321-57.

[5] World Health Organization: obesity and overweight. Fact sheet No. 311,[updated June 2016]. Available from: http://www.who.int/mediacentre/factsheets/fs311/en/. Accessed December 13, 2016.

[6] Romero-Corral A, Somers VK, Sierra-Johnson J, et al. Accuracy of bodymass index to diagnose obesity in the US adult population. Int J Obes(Lond). 2008;3(6):959966.

[7] Heart rate variability: standards of measurement, physiological interpretationand clinical use. task force of the European society of cardiology and the North American Society of Pacing and Electrophysiology. Circulation. 1996;93(5):10431065.

[8] Vanderlei LCM, Pastre CM, Hoshi RA, Carvalho TD, Godoy MF. Basic notions of heart rate variability and its clinical applicability. Rev Bras Cir Cardiovasc. 2009;24(2):205217.

[9] Laederach-Hofmann K, Mussgay L, Ruddel H. Autonomic cardiovascular regulation in obesity. J Endocrinol. 2000;164(1):5966.

[10] Billman GE, Huikuri HV, Sacha J, Trimmel K. An introduction to heart rate variability: methodological considerations and clinical applications. Front Physiol. 2015;6:55.

[11] Billman GE. The effect of heart rate on the heart rate variability response to autonomic interventions. Frontiers in physiology. 2013;4:222.

[12] Shenoy AR, Doreswamy V, Shenoy JP, Prakash VS. Impact of obesity on cardiac autonomic functions in middle aged males. National Journal of Physiology, Pharmacy and Pharmacology. 2014;4(3):236-9.

[13] Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. CRC press; 1984.

[14] Natekin A, Knoll A. Gradient boosting machines, a tutorial. Frontiers in neurorobotics. 2013;7:21.

[15] Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian Journal of internal medicine. 2013;4(2):627-635.