Oxidative Stress Markers Identify Cardiac Autonomic Neuropathy Progression: Applying Machine Learning Methods

Alaa Alqaryuti[1], Nadeen Faraj [1], Mohamed Abdelmagid [1], Maher Maalouf [1], and Herbert F Jelinek[2]

[1] Department of Industrial and Systems Engineering, Khalifa University, Abu Dhabi, UAE
[2] Department of Biomedical Engineering, Khalifa University, Abu Dhabi, UAE

Abstract

This study aims to highlight the association between oxidative stress and cardiac autonomic neuropathy (CAN) using machine learning algorithms for risk prediction. Oxidative stress is a significant factor in chronic diseases. Data from 2,621 participants were provided by the DiabHealth diabetes complications screening clinic at Charles Sturt University (CSU) for analysis, spanning the years 2002 to 2015. The oxidative stress markers considered in this study were 8-isoprostane, 8-hydroxy-2'-deoxyguanosine (8-OHdG), reduced glutathione (GSH), oxidized glutathione (GSSG) and glutathione redox ratio (GSH/GSSG). Machine learning methods, including Random Forest and Logistic Regression, were employed to develop two multi-class and one binary model. For ROC-AUC, all models achieved relatively high values where "Definite" in model 1 is 0.82, "Normal" in model 2 is 0.81, and "Abnormal" in model 3 is 0.81. The findings underline the potential of integrating machine learning methods in CAN prediction, offering substantial improvements over traditional methods. By exploring novel multi-class models and unveiling the capabilities of the random forest classifier, this research establishes a robust foundation for future investigations.

1. Introduction

Cardiovascular disease (CVD), mortality, and morbidity with Diabetes Mellitus (DM) are significantly impacted by cardiac autonomic neuropathy (CAN), a prevalent and poorly understood diabetes-related condition. CAN is usually inadequately diagnosed, particularly in clinical practice as the majority of diabetic patients have subclinical or asymptomatic CAN [1]. The capacity to diagnose CAN at a subclinical level has advanced from the initial five cardiac autonomic reflex tests (CARTs) that are time-consuming including heart rate variability (HRV) analysis, baroreflex sensitivity (BRS) tests, and cardiac imaging [2]. Oxidative stress markers, mitochondrial dysfunction, and high cholesterol levels have all been shown to contribute to the development of CAN [3, 4].

The main focus of this study is on oxidative stress markers since they are easy to assess. Urinary 8-isoprostane and 8-hydroxy-2'-deoxyguanosine (8-OHdG) are commonly used for assessing oxidative stress in diabetes [2]. In diabetic patients with CAN, these markers are elevated, indicating the of oxidative stress in the development of the disease. Glutathione (GSH) and its oxidized form, GSSG, constitute another significant antioxidant system. GSH depletion and higher GSSG levels are associated with CAN and diabetes [5].

The traditional technique for CAN diagnosis contains five tests based on heart rate (HR) and blood pressure (BP) responses [1]. These tests consist of the HR response to (1) deep breathing, (2) standing from a supine position, (3) the Valsalva maneuver, and blood pressure response to (4) standing, and (5) sustained handgrip [2]. The methods used to conduct these tests and their normal, borderline, and abnormal values are presented in Table 1.

In this study several machine learning methods have been used to develop models for CAN diagnosis and prediction. These include Random Forest (RF), Logistics Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machines (SVMs), and Gradient Boosting (GB) [6]. Among the most used measures of machine learning performance are accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (ROC-AUC).

This study is motivated by the need to replace traditional CAN testing methods, which require physical movement that may not be feasible for some patients. The aim is to highlight the association between oxidative stress and CAN using machine learning algorithms for risk prediction.

2. Methodology

Figure 1 illustrates the methodology followed in this paper. The dataset consists of 2621 patient entries from the year 2002 to 2015. The target variable was created using

Table 1: Values of cardiovascular autonomic function tests

Test	Normal	Borderline	Abnormal
HR response to deep breathing (beats/min)	≥ 15	$11 - 14$	≤ 10
HR response to standing (30:15 ratio)	≥ 1.04	$1.01 - 1.03$	≤ 1.00
HR response to Valsalva manoeuvre	≥ 1.21	-	≤ 1.20
BP response to standing (mmHg)	≤ 10	$11 - 29$	≥ 30
BP response to sustained handgrip (mmHg)	≥ 16	$11 - 15$	≤ 10

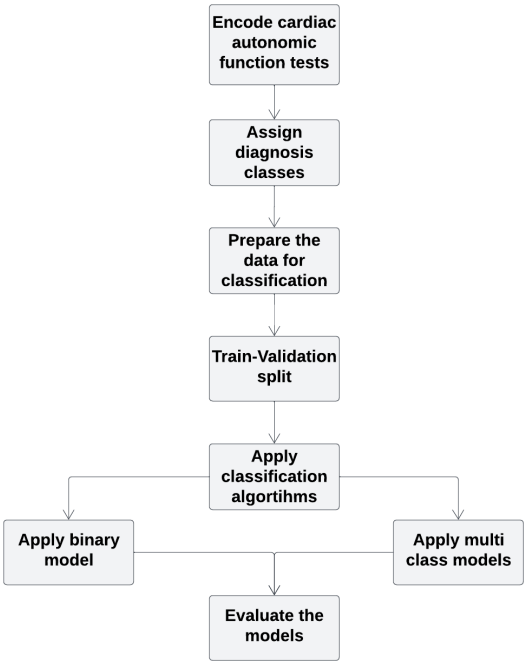


Figure 1: Methodology Flow

a rule-based method, resulting in a distribution of Normal: 862, atypical: 514, Early: 997, Severe: 108, and Definite: 140.

CAN severity is set as follows: Each test in table 1 is considered as normal, borderline, or abnormal based on the obtained value. Five classes of CAN severity are established; normal, early, definite, severe, and atypical. Table 2 describes the classifications based on test conditions. As described in Table 3, two multi-class models and a binary model are considered in this study. In our primary approach (Model 1), all distinct stages of CAN are modeled separately, capturing the progression from Normal to Severe involvement. We further explore two alternative modeling frameworks. In Model 2, we combine 'Early' and 'Definite' involvements into a single 'Moderate' category, providing a three-class representation. For Model 3, we adopt a binary classification perspective, merging 'Early', 'Definite', and 'Severe' involvements into a unified 'Abnormal' category, juxtaposes against the 'Normal' class. This enables us to evaluate the models' efficacy across varying granularity levels.

Table 2: Patient classification based on test conditions

CAN stage	**HR & BP Tests**
Normal	All tests normal or one borderline
Early	One of the three heart rate tests abnormal or two borderline
Definite	Two or more of the heart rate tests abnormal
Severe	Two or more of the heart rate tests abnormal, plus one or both of the blood pressure tests abnormal or both borderline
Atypical	Any other combination

Table 3: Models Definition

Models	**CAN stages**	**Definition**
Model 1	4	Normal, Early, Severe, and Definite
Model 2	3	Normal, Early, Moderate (Definite and Severe)
Model 3 (Binary)	2	Normal, Abnormal (Early, Definite, and Severe)

The classification methods applied to the data set in Python are the following; LR calculates the probability of a patient having CAN based on the values of the predictor variables, i.e., oxidative stress markers. SVM algorithm works by finding the hyperplane that maximizes the margin between classes in the feature space. KNN, RF and GB are utilized to classify the stage of CAN for a patient, based on the oxidative stress markers. All methods are used for testing models 1, 2, and 3. The performance of the classification models is evaluated using a suite of metrics, including accuracy, precision, recall, F1-score, and ROC-AUC. In addition to these metrics, 5-fold cross-validation is used to assess the robustness of the models. The model is trained on 80% of the data and tested on the remaining 20%.

3. Results

This study predicted CAN severity based on the results of the five CARTs using multiple classifiers. Table 4 shows the classifier accuracy of Model 1. The RF classifier achieved the highest accuracy of 0.6487 (± 0.0120). Given the superior performance of the RF classifier, it was selected for further modeling. According to the RF classifier analysis, all features played a significant role in determining CAN stages. "u-8-OHdG" scored approximately 22.78%, making it the most influential element. However, the differences in importance among the features were relatively small.

Table 4: Accuracies of different classifiers using 5-fold cross-validation

Classifier	Accuracy
LR	0.4944 (± 0.0219)
SVM	0.5323 (± 0.0131)
KNN	0.6241 (± 0.0104)
GB	0.6166 (± 0.0218)
RF	**0.6487 (± 0.0120)**

Evaluating the classifier performance on the testing dataset, it exhibited an accuracy of 65.64%. The precision metrics were: 68% for "Normal", 65% for "Early", 53% for "Definite", and 69% for "Severe". The recall for "Definite" stood at 28%, while "Early" and "Normal" stages exhibited recalls of 73% and 67%, respectively. The "Severe" stage achieved a recall of 46%. F1-Score trends mirrored these figures, highlighting potential limitations in the precise identification of the "Definite" stage.

Further insights can be drawn from the ROC-AUC for each CAN stage, where "Normal", "Early", "Definite", and "Severe" were found to be 0.81, 0.73, 0.82, and 0.76 respectively, as illustrated in Figure 2. These ROC-AUC values emphasize the classifier's strong capability to discriminate between positive and negative instances for each stage, especially for the "Definite" stage, despite its earlier noted limitations.

The classifier exhibited notable accuracy in identifying "Normal" and "Early ", with 114 and 143 correct predictions, respectively. Figure 3(a) shows the confusion matrix for the RF classifier. However, distinctions between adjacent stages proved challenging, especially discerning "Normal" from "Early ". The "Severe " stage saw 11 accurate identifications.

To investigate the RF classifier performance for different groupings, a multi-class model and a binary model were created. As mentioned previously, the four-class model acts as the baseline model but achieves low accuracy. For Model 2, performance significantly improved with an accuracy of 71%. The ROC-AUC for "Normal",

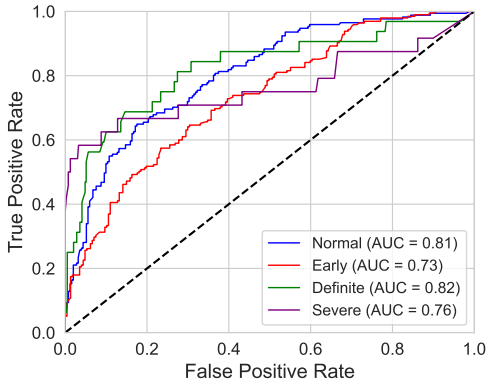
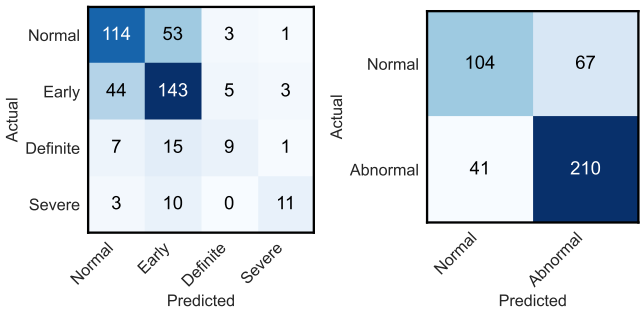


Figure 2: Model 1 - ROC-AUC



(a) Model 1 (b) Model 3

Figure 3: Confusion matrix for Random Forests classifier

"Moderate", and "Sever" were 0.81, 0.78 and 0.76 respectively as shown in Figure 4. This suggest that the fusion of 'Definite' and 'Severe' into 'Moderate' seems to aid in achieving a more discernible differentiation between stages. For Model 3, which comprises a binary classification of CAN stages into Normal and Abnormal, the RF classifier achieved an accuracy of 74.40%. The precision, recall, and F1-score for the Normal class stand at 0.72, 0.62, and 0.66 respectively. In contrast, the Abnormal class has values of 0.76, 0.84, and 0.80 for the same metrics, respectively. As shown in figure 3(b) the confusion matrix indicates that 104 out of 171 Normal instances were correctly identified, while the Abnormal class saw a higher true positive rate with 210 out of 251 instances being accurately classified. Comparatively, the binary model outperforms Model 1 and Model 2 in terms of overall accuracy. The distinction between the Normal and Abnormal classes in Model 3 seems to enable better generalization. Notably, the Abnormal class in Model 3 has a higher recall than both the merged Moderate category in Model 2 and the separate stages in Model 1. Figure 5 assess the classifier's performance in distinguishing between 'Normal' and 'Abnormal' CAN stages. A notable

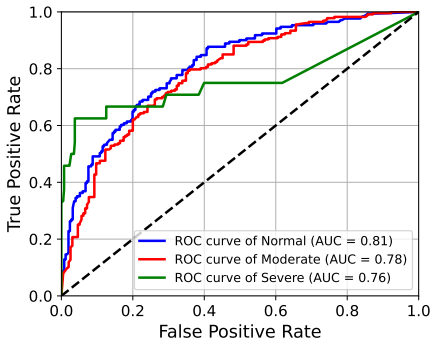


Figure 4: Model 2 -ROC-AUC

ROC-AUC of 0.81 is achieved, which suggests that our classifier possesses a commendable ability to discern between these two stages.

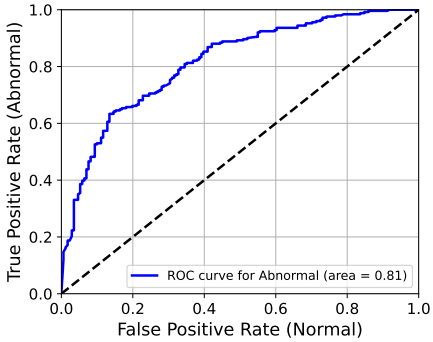


Figure 5: Model 3 - ROC-AUC

4. Discussion and Conclusion

In this study, the performance of various machine learning classifiers were explored, including LR, SVM, KNN, RF, and GB. The results of all experiments showed that the best accuracy was obtained by the RF classifier on the binary model (Model 3). For ROC-AUC, all models achieved relatively high values where "Definite" in model 1 is 0.82, "Normal" in model 2 is 0.81, and "Abnormal" in model 3 is 0.81. These values indicate that the progression of CAN is highly associated with oxidative stress markers. Some factors, such as patient age, gender, comorbidities, and medications, may increase the ROC-AUC percentage obtained in our study if they were considered [5].

The findings of this study support the results of previous studies that used machine learning methods for CAN detection using CARTs. As the RF classifier had the best performance compared to other models. Abdalrada et al., [7] obtained an accuracy of 94.1% and ROC 0.980 when adding HRV attributes to CAN classification. Moreover, Rashid et al., [8] obtained an accuracy of 98.67% by including the patient demographic (gender), clinical, and laboratory profiles for CAN testing. These studied included different attributes than our study, as we only used oxidative stress markers. Future studies should consider other markers and attributes such as HRV attributes and demographic profiles to obtain better results. Moreover, testing each oxidative stress marker separately should enhance the accuracy.

References

[1] Kuehl M, Stevens MJ. Cardiovascular Autonomic Neuropathies as Complications of Diabetes Mellitus. Nature Reviews Endocrinology 2012;8(7):405–416.

[2] Ewing D, Campbell I, Clarke B. Assessment of Cardiovascular Effects in Diabetic Autonomic Neuropathy and Prognostic Implications. Annals of Internal Medicine 1980;92(2_Part_2):308–311.

[3] Maschirow L, Khalaf K, Al-Aubaidy H, Jelinek H. Inflammation, Coagulation, Endothelial Dysfunction and Oxidative Stress in Prediabetes—Biomarkers as a Possible Tool for Early Disease Detection for Rural Screening. Clinical Biochemistry 2015;48(9):581–585.

[4] Voigt A, Jelinek HF. Humanin: A Mitochondrial Signaling Peptide as a Biomarker for Impaired Fasting Glucose-Related Oxidative Stress. Physiological Reports 2016;4(9):e12796.

[5] Pouvreau C, Dayre A, Butkowski EG, De Jong B, Jelinek HF. Inflammation and Oxidative Stress Markers in Diabetes and Hypertension. Journal of Inflammation Research 2018; 61–68.

[6] Jelinek H, Kelarev A. A Survey of Data Mining Methods for Automated Diagnosis of Cardiac Autonomic Neuropathy Progression. AIMS Medical Science 2016;3(2):217–233.

[7] Abdalrada AS, Abawajy J, Chowdhury M, Rajasegarar S, Al-Quraishi T, Jelinek HF. Meta Learning Ensemble Technique for Diagnosis of Cardiac Autonomic Neuropathy Based on Heart Rate Variability Features. In 30th International Conference on Computer Applications in Industry and Engineering, CAINE. 2017; 169–175.

[8] Rashid M, Alkhodari M, Mukit A, Ahmed KIU, Mostafa R, Parveen S, Khandoker AH. Machine Learning for Screening Microvascular Complications in Type 2 Diabetic Patients Using Demographic, Clinical, and Laboratory Profiles. Journal of Clinical Medicine 2022;11(4):903.

Address for correspondence:

Maher Maalouf
maher.maalouf@ku.ac.ae
P O Box 127788, Abu Dhabi, UAE
Khalifa University of Science and Technology