



Review article

A Survey of Data Mining Methods for Automated Diagnosis of Cardiac Autonomic Neuropathy Progression

Herbert F. Jelinek * and Andrei V. Kelarev

Centre for Research in Complex Systems and School of Community Health, Charles Sturt University, Albury, NSW, Australia

* **Correspondence:** Email: hjelinek@csu.edu.au; Tel: +6-126-051-9219; Fax: +6-126-051-9238

Abstract: Cardiac autonomic neuropathy (CAN) is a disease that occurs as a result of nerve damage causing an abnormal control of heart rate. CAN is often associated with diabetes and is important, as it can lead to an increased morbidity and mortality of the patients. The detection and management of CAN is important since early intervention can prevent further complications that may lead to sudden death from myocardial infarction or rhythm disturbance. This paper is devoted to a review of work on developing data mining techniques for automated detection of CAN. A number of different categorizations of the CAN progression have been considered in the literature, which could make it more difficult to compare the results obtained in various papers. This is the first review proposing a comprehensive survey of all categorizations of the CAN progression considered in the literature, and grouping the results obtained according to the categorization being dealt with. This novel, thorough and systematic overview of all categorizations of CAN progression will facilitate comparison of previous results and will help to guide future work.

Keywords: automated detection; cardiac autonomic neuropathy; data mining; machine learning; diabetes complications

1. Introduction

Clinical applications of data mining techniques have been actively investigated. For preliminaries and background information on this broad area let us refer the readers to the monographs [1–3]. In particular, for successful treatment of various conditions, it is important to find attributes that may help in the early detection of signs and symptoms of disease that may facilitate the prevention, early diagnosis and treatment (cf. [4–9]). Likewise, automated computer-based diagnosis plays an important role in eHealth and mobile applications (cf. [10,11]). The present review article deals with recent contributions to this broad research area for the special case of cardiovascular autonomic neuropathy (CAN), which is a well-known complication associated with diabetes (cf. [12–14]).

Cardiovascular (CVD) complications associated with diabetes account for 65% of all diabetic deaths [15]. The large impact of CVD associated with diabetes mellitus Type 1 and Type 2 has brought about the recommendation that people with diabetes should be regularly screened for the presence of comorbidities including autonomic nervous system dysfunction with the aim to decrease the incidence of cardiovascular related morbidity and mortality [16–18]. People with diabetes and autonomic neuropathy have an increased mortality rate (29%) compared to people with diabetes but no autonomic neuropathy (6%) [19,20]. As many as 22% of people with type 2 diabetes suffer from CAN, with prevalence increasing as duration of diabetes increases [21,22]. CAN leads to impaired regulation of blood pressure, heart rate and heart rate variability (HRV). The increased risk of cardiac mortality due to arrhythmias makes screening of people with diabetes for autonomic neuropathy vital so that early detection, intervention and monitoring can occur [23]. Autonomic neuropathy is also associated with non-response hypoglycemia and a reduction in counter-regulation of the hypoglycemic events [24,25]. Silent ischemia is significantly more frequent in patients with CAN than in those without CAN [26,27] and significantly more people with diabetes die from cardiovascular disease such as heart attack and stroke, which can be attributed to CAN [28]. Early subclinical detection of CAN and intervention are of prime importance for risk stratification in preventing the potentially serious consequences of CAN [29].

Data mining methods are an important adjunct to medical research in identifying disease markers that allow early detection, prevention or treatment of disease. Electronic patient records and large healthcare databases combined with data mining provide a means to improve the level of health by identifying latent features not identified previously that are strong indicators of disease [30]. Data mining methods have been used extensively in health care research to build prediction models that provide additional information for improving health care outcomes [31–33].

2. Tests of the Ewing Battery

Autonomic neuropathy in diabetics has been traditionally identified by performing the Ewing

battery of tests, which was recommended by the American Diabetes Association and the American Academy of Neurology. These tests evaluate heart rate (HR) and blood pressure (BP) responses to various activities [34–36]. The five tests in the Ewing battery are shown in Table 1 following [35]. Results of these tests provide a good assessment of diabetic autonomic neuropathy and aid in objective diagnosis instead of relying on self-reported clinical signs such as gustatory sweating, reflux, and incontinence. For the current study, patient results were only included if participants in the study were free of medication and comorbidities affecting their heart rate. The response of subjects to each of the Ewing tests is defined as normal, borderline or abnormal, as shown in Table 1, where HR is measured beats per minute (BPM) and BP is measured in mmHg.

Table 1. Summary of Ewing tests.

Values	HR response to standing	BP response to standing	HR response to deep breathing	Valsalva maneuver and HR response	BP response to sustained handgrip
Abnormal	≤ 1.00	≥ 30	≤ 10	≤ 1.10	≤ 10
Borderline	1.01–1.03	11–29	11–14	1.11–1.20	11–15
Normal	≥ 1.04	≤ 10	≥ 15	≥ 1.21	≥ 16

Let us refer to [35,37,38] and [13] for more explanations and details on conducting and interpreting tests included in the Ewing battery. From this grading CAN risk assessment can be divided into a normal and no CAN evident category and four CAN categories comprising: early, definite, severe and atypical, as shown in Table 2.

Table 2. Categorization of CAN based on Ewing tests.

Category	Normal	Early	Definite	Severe	Atypical
Decision criterion	All tests normal or one borderline	One of the three heart rate tests abnormal or two borderline	Two or more of the heart rate tests abnormal	Two or more of the heart rate tests abnormal plus one or both of the blood pressure tests abnormal or both borderline	Any other combination of tests with abnormal results

The categorization given in Table 2 is for diagnosing CAN categories as shown in [37]. The paper [35] compared these rules for determining the categories of CAN with two alternative scoring systems. The first one gave 0 for a normal result, ½ for a borderline outcome, and 1 for an abnormal outcome, resulting in a combined total score ranging from 0 to 5 for each participant. The second set of rules counted the number of outcomes that were abnormal, which again produced a total score in the range from 0 to 5 for each person. The paper [35] demonstrated that these scoring systems give

roughly equivalent categorizations and neither seems to carry a real advantage over the other.

It is not always possible for patients to perform all of the Ewing tests. For instance, the hand grip test may be difficult due to arthritis. The lying to standing tests often cannot be included in the test battery results due to mobility challenges of patients. Likewise, some patients have conditions where forceful breathing required for the Valsalva maneuver is contra-indicated. These issues result in CAN risk assessments being made in practice on the basis of only a subset of the Ewing battery.

This review groups the results obtained in the literature into several sections according to the categorization being studied. If a paper considers several classifications simultaneously, we include a more detailed summary of the new methods proposed in this paper in the section devoted to the categorization with the smallest number of classes since usually it is easier to handle and the best outcomes are achieved in this case. In the sections devoted to categorizations with a larger number of classes we only include an indication of the best values obtained for the corresponding categorization in the relevant paper.

The following categorizations of CAN progression have been considered in the literature:

- CAN2: the presence or absence of CAN (2 classes);
- CAN3: absence of CAN, early CAN, and definite CAN (3 classes);
- CAN4: normal, early, definite and severe CAN (4 classes);
- CAN5: normal, early, definite, severe, and atypical CAN (5 classes);
- CAN-early: early CAN and the absence of early CAN (2 classes);
- CAN-severe: severe CAN and the absence of severe CAN (2 classes).

Here the absence of early CAN is the union of normal, definite CAN and severe CAN classes. Likewise, by the absence of severe CAN we mean the union of normal, early CAN, and definite CAN classes.

3. DiabHealth Database

Many articles have used the large database of health-related parameters and tests amalgamated in the Diabetes Screening Complications Research Initiative (DiabHealth) [39] organized by Charles Sturt University in Australia. The collection and analysis of data in the project was approved by the Ethics in Human Research Committee of the university. The participants were instructed not to smoke and refrain from consuming caffeine containing drinks and alcohol for 24 hours preceding the tests as well as to fast from midnight of the previous day until tests were complete. The measurements were conducted from 9:00am until 12midday and were recorded in the DiabHealth database along with various other clinical data including age, sex and diabetes status, blood pressure (BP), body-mass index (BMI), blood glucose level (BGL), and cholesterol profile. Reported incidents of a heart attack, atrial fibrillation and palpitations were also recorded. DiabHealth has made it possible to collect a large database with over 2500 entries and several hundred features.

4. Heart Rate Variability for the Automated Diagnostics of CAN

The Ewing battery is commonly used for detecting CAN but is often not conclusive and therefore more sensitive and accurate tests are required. This section deals with one of the most important special classes of attributes, which have been applied for the automated detection of CAN in previous publications.

Heart Rate Variability (HRV) as a clinical tool using ECG recordings has been shown to be a sensitive marker for risk of future arrhythmias or CAN and is easier to use clinically compared to the Ewing battery [40]. The ability to use only HRV for accurate identification of CAN and CAN progression provides alternative test results to the physician in addition to invasive testing such as cholesterol, BGL and HbA1c results.

HRV attributes can serve as a safeguard measure detecting CAN from short heart rate recordings during a patient health review. Several articles have applied HRV features to the task of automated detection of CAN. The motivation to use HRV data is that it is more often available and easier to obtain in clinical practice than the Ewing battery features. HRV measures also provide many more variables compared to the five attributes in the Ewing battery.

HRV analysis involves determining the interbeat intervals (RR intervals) between successive QRS complexes on an ECG or directly from heart rate recordings. HRV information can include as many as 20–30 measures sensitive to different characteristics of the ECG time series that can be divided into time, frequency, and nonlinear measures [41]. ECGs are routinely assessed in clinical practice and although they do not directly indicate CAN, HRV can be determined from the interbeat interval tachogram or from a continuous heart rate recording [42].

Nonlinear HRV measures have become popular in recent times as they are more robust against nonstationarity and nonlinearity characteristics of the RR tachogram and are able to detect how aging and pathological conditions affect interbeat variation [43,44]. Nonlinear HRV features such as detrended fluctuation analysis (DFA), estimate complexity inherent in the signal. The correlation dimension (D_2) can also be applied [45].

Table 3 summarizes various HRV analysis methods. In particular, it uses the notion of normal to normal beat intervals, also called NN intervals, see [44,46,47] for more explanations.

Several entropy measures have been proposed such as approximate entropy, sample entropy, tone-entropy [48,49]. These measures have subsequently led to the multiscale entropy measures including the multi-scale Rényi entropy, which is a generalization of the Shannon entropy [50]. The Rényi entropy H is defined as

$$H_{\alpha}(X) = \frac{1}{1 - \alpha} \log_2 \sum_{n=1}^{\infty} p_i^{\alpha}$$

where α is the order of the Rényi entropy and p_i stands for the probability of X being equal to a particular value. The value of Rényi entropy for given π and α is denoted by $H(\pi, \alpha)$.

Table 3. HRV measures.

HRV method	Measure	Description
Time	SDNN	The standard deviation of the normal to normal beat intervals
Domain	RMSSD	The square root of the mean squared difference of the NN intervals
Frequency	Total Power	Variance of N-N intervals over the temporal segment (freq < 0.4)
Domain	VLF	Power in very low frequency range (freq < 0.04)
	LF	Power in low frequency range (freq 0.04 to 0.15)
	HF	Power in high frequency range (freq 0.15 to 0.4)
Nonlinear	SD1, SD2	The standard deviations perpendicular to and along the line-of-identity of the Poincaré plot
	ApEn	Approximate entropy
	SampEn	Sample entropy
	D ₂	Correlation dimension
	DFA	Detrended fluctuation analysis: α_1 - Short-term fluctuation slope; α_2 - Long-term fluctuation slope
	ShanEn	Shannon entropy
	Rényi En	Rényi entropy

5. Data Mining Methodology

The following standard measures of the effectiveness or performance of classifiers have been considered in the literature devoted to the diagnostics of CAN: accuracy, precision, recall, F-measure, sensitivity, specificity and Area Under Curve (AUC) also known as the Receiver Operating Characteristic or ROC area. These measures are standard and well-known. The readers interested in detailed explanations of these standard measures can find them in the monograph [51].

Several articles devoted to CAN divide the data set into a training set and a validation set to assess the effectiveness of the classifiers being designed ([52]). On the other hand, 10-fold cross validation, which is a standard well-known technique organizing experiments to prevent overfitting machine learning models to data, have also been reported. It can be implemented in WEKA and is invoked by default as stratified 10-fold cross validation, see [51]. It divides data into ten stratified folds and creates training sets and hold out testing sets ten times for ten consecutive tests with hold out sets automatically. Another method designed in the literature to prevent over-fitting is the 5×2 cross-validation introduced and recommended in [53] for comparison of classifiers. This method carries out five iterations of twofold cross-validation. The results of cross-validation implemented in WEKA are included in the output of all classifiers automatically, which makes it easy to apply cross validation in experiments concerning classifiers implemented in WEKA [54,55].

6. Binary Classification CAN2

The paper [38] carried out a study demonstrating the usefulness of HRV and complexity analyses based on short term ECG recordings as a screening tool for CAN2 categorization. Binary classification CAN2 was investigated because so few patients realize that they have CAN since it can be asymptomatic for a substantial period of time during diabetes progression. The study utilized a total of 17 sets of ECG recordings from DiabHealth from diabetic subjects with definite CAN and without CAN. Poincaré plot indices were used and compared with the traditional time and frequency, and the sample entropy (SampEn) to analyze variability (short and long term) and complexity of HRV respectively. It was shown that reduced ($p > 0.05$) Poincaré plot patterns and lower ($p < 0.05$) SampEn values corresponded to patients with definite CAN, which can be a practical diagnostic and prognostic marker. Furthermore, a simple decision tree was trained to predict CAN2 categories using SampEn and Poincaré plot indices. It produced outcomes with 100% sensitivity and 75% specificity achieving an accuracy of 88.24%.

In order to reduce the cost of performing medical tests required to collect the attributes yet maintain diagnostic accuracy, it is essential to optimize the features used for classification and to keep the number of features as small as possible. Feature selection methods of this kind are outlined in this section. The binary classification CAN2 was also studied in [52]. Instead of concentrating on the role of attributes of a particular type, the article applied data mining feature selection methods to derive a set of features to be used for the automated detection of CAN. The experiments undertaken divided the data set into a training set and a validation set. A hybrid of Maximum Relevance filter (MR) and Artificial Neural Net Input Gain Measurement Approximation (ANNIGMA) wrapper approaches was defined and was used to reduce the number of features necessary for optimal classification. The combined heuristic MR-ANNIGMA exploits the complimentary advantages of both the filter and wrapper heuristics to find significant features [52].

The best accuracy obtained by applying this method for CAN2 was 80.66%. The feature selection approach applied in [52] however identified an effective set of ECG components associated with CAN2, which have clinical relevance. More information on the relation of ECG features, CAN and hypertension have been established in [56,57].

The binary classification CAN2 was also considered in [57–59]. The best classifiers reported in those publications achieved predictive accuracies of 94.61% [57], 94.84% [58], and 97.74% [59].

The paper [60] dealt with the binary classification CAN2 for diabetes patients only. It carried out a comprehensive study of the effectiveness of several decision trees including ADTree, J48, NBTree, RandomTree, REPTree, and SimpleCart and various ensembles of decision trees generated by applying AdaBoost, Bagging, Dagging, Decorate, Grading, MultiBoost, Stacking, and two multilevel combinations of AdaBoost and MultiBoost with Bagging. The best classifier designed in [60] achieved classification results for CAN2 with the ROC area equal to 0.947.

In [41], visualization methods for determining the categories of CAN2 were studied. The

authors concentrate on visualization using only data derived from HRV. A variety of measures were extracted from the sequence of interbeat time intervals (RR intervals). The multiscale Renyi entropy was calculated using $-5 < \alpha < +5$, where $\alpha = 1$ gives the Shannon entropy and $\alpha = 2$ produces the squared entropy. Sample Entropy was also calculated in order to provide a comparison. All features calculated from HRV were visualized using a Spider diagram. The results show that this visualization technique not only captures that binary classification CAN2, but provides additional insights by displaying a comprehensive picture of the complexity of the disease. In this relation, let us note that it would be also interesting to investigate the applications of conceptual graphs for the visualization of the diagnostics of CAN progression, since conceptual graphs are well known effective tools for formal visual reasoning in the medical domain [61,62].

Karmakar and colleagues [63] undertook a multi-lag Tone-Entropy (T-E) analysis of HRV data for CAN2. A total of 41 ECG recordings from DiabHealth were utilized with definite CAN and without CAN. T-E values of each patient were calculated for different beat sequence lengths (denoted by len and ranging from 50 to 900) and lags (denoted by m and ranging from 1 to 8). For all values of the len and m parameters, it was discovered that the group of normal patients has a lower mean tone value compared to that of definite CAN patients, whereas the mean entropy value was higher in normal patients than in patients with definite CAN. Leave-one-out cross-validation tests using a quadratic discriminant (QD) classifier were applied to investigate the performance of multi-lag T-E features. This produced 100% accuracy for T-E with len = 250 and m = {2, 3} settings, which is better than the performance of T-E technique based on m = 1. The results demonstrated the usefulness of multi-lag T-E analysis over single lag analysis for the diagnosis of CAN2 categorization.

Investigating data transfer, the paper [64] solved the problem of minimizing data transfer between different data centers of the cloud during the diagnosis of CAN2 by classifiers deployed in the cloud. A new model of clustering-based multi-layer distributed ensembles (CBMLDE) was introduced. It was designed to eliminate the need to transfer data between different data centers for training of the classifiers. Ten-fold cross validation and a dataset derived from DeabHealth were used in order to determine the best combinations of options for setting up CBMLDE classifiers. The results demonstrated that CBMLDE classifiers not only completely eliminate the need in patient data transfer, but also have significantly outperformed all base classifiers and simpler counterpart models in all cloud frameworks.

7. Ternary Classification CAN3 and Quaternary Classification CAN4

Ternary classification, CAN3, has been considered simultaneously with the quaternary classification, CAN4, in previous papers. These papers also included CAN2. The paper [65] considered CAN2, CAN3, and CAN4. The paper only used complete data without addressing the problem of missing values and applied feature selection methods incorporated in the implementation of Random Forest in R [66,67] to select relevant features. Multilevel classifiers were investigated in [65].

The best classifier produced AUC values of 0.997 for CAN2, 0.994 for CAN3, and 0.990 for CAN4. It is essential to note that the results of [65] cannot be applied to handle missing values, since all tests used a large set of features and a complete dataset.

The Ewing battery of tests is still commonly used, but the question of which if any of the single tests included in the Ewing battery may perform equally well as the 5-test battery and which one of the five tests this may be has not been investigated prior to Stranieri and colleagues [29] handled all three classifications CAN2, CAN3, and CAN4 to address this question. An optimal order of the Ewing tests was determined using the Optimal Decision Path Finder procedure proposed in [68]. In addition, visual aids were developed in [29] to simplify the selection of the next Ewing test during applications of this procedure in practice. Only simple basic decision trees were used and the best accuracy achieved was equal to 94.14% for CAN2 [29].

The paper [69] introduced a new parameter, the beat-to-beat TQ-RR ratio derived from ECG recordings and was investigated in conjunction with the systolic-diastolic interval interaction (SDI) parameter. Performance of both QT-TQ and TQ-RR based SDI measures was explored to diagnose the categories of CAN3. ECG recordings of 72 diabetic subjects without CAN, 55 subjects with early CAN and 15 subjects with definite CAN from the DiabHealth study were utilized. The outcomes obtained demonstrated that variability of the TQ-RR based SDI measure can distinguish all three categories of CAN3 with p-value $p < 0.001$. In contrast, the variability of the QT-TQ based SDI measures showed significant difference only between the normal subjects and definite CAN categories. This demonstrates TQ-RR based SDI parameter turned out more sensitive in the detection of CAN3 categories compared to the QT-TQ based measures.

The paper [13] used ten-fold cross validation to compare the effectiveness of applications of decision trees, ensemble classifiers and multi-level ensemble classifiers for neurological diagnostics of CAN. It investigated and compared the effectiveness of AdaBoost, Bagging, MultiBoost, Stacking, Decorate, Dagging, and Grading in their ability to enhance the performance of decision trees—ADTree, J48, NBTree, RandomTree, REPTree, SimpleCart, as well as several other base classifiers—Decision Table, FURIA, J48, NBTree, Random Forest and SMO. In addition, Jelinek et al. also investigated the meta-classifiers AdaBoost, Bagging, Dagging, Decorate, Grading, MultiBoost and Stacking to determine which function best in enhancing the performance of base classifiers further within the framework of a multi-level classification paradigm [13]. Several other studies handled other different multi-stage systems ([29,70–72]). The best outcomes were obtained by two options combining Bagging and Decorate into one multi-level ensemble classifier. The first option was Bagging in the second level after applications of Decorate based on Random Forest in the first level. The other optimal solution was using Decorate in the second level to combine the results of Bagging applied to Random Forest as a base classifier. These combinations achieved an AUC of 0.997 for CAN2, 0.994 for CAN3, and 0.990 for CAN4, respectively.

A new machine learning algorithm for the diagnosis of CAN progression based on HRV attributes was proposed in [14]. The Multi-Layer Attribute Selection and Classification (MLASC)

algorithm was applied to the DiabHealth data and accuracy was determined by tenfold cross-validation to investigate optimal options to be employed at the different layers of MLASC and compared the effectiveness of MLASC with other counterpart versions and simplified methods. In addition, [14] introduced a new feature selection method, Double Wrapper Subset Evaluator with Particle Swarm Optimization (DWSE-PSO), and incorporated it in one of the steps of MLASC. The outcomes of experiments presented in [14] compared several automated feature selection and optimization methods in their effectiveness as a subpart of MLASC. These feature selection techniques had not been considered before for the task of classification of CAN progression using HRV attributes. The results obtained in [14] deal with three classifications: CAN2, CAN3, and CAN4 and demonstrated that the MLASC algorithm produced the best results when applying DWSA with PSO attribute selection, and performed better in comparison with other approaches. Diagnosis by this method achieved an AUC level of 0.96 for CAN2, 0.95 for CAN3, and 0.94 for CAN4, respectively.

Decision making and diagnosis in medical practice is most often based on incomplete data due to either unavailability of diagnostic laboratory services, technical issues or lack of patient cooperation as well as counter-indications for undertaking certain diagnostic tests. The first article specifically addressed to the problem of missing values in this clinical paradigm by using the Ewing formula and regression to classify CAN [71]. Previous articles circumvented the problem of missing values in CAN data by deleting all records and fields with missing values and applying classifiers trained on different sets of features that were complete. Most of these also added alternative features to compensate for the deleted values. The paper [71] introduced and investigated a new method for classifying CAN data with missing values. In contrast to all previous papers, the new method proposed by Abawajy et al. did not delete attributes with missing values, did not use classifiers, and did not add features. Instead it applied regression and meta-regression combined with the Ewing formula to identify CAN2 and CAN3. The best outcomes were obtained by the additive regression meta-learner based on M5Rules and combined with the Ewing formula. It achieved the best accuracy of 99.78% for CAN2, and 98.98% for CAN3.

8. Quinary Classification CAN5

The quinary classification CAN5 was investigated in [73] concentrating on the impact of blood biochemistry features for improving the predictive accuracy of the diagnosis of CAN in situations where one or more of the Ewing tests cannot be completed because of the individual difficulties faced by each patient in performing the tests. A novel class of iterative multilevel ensembles, called automated iterative multitier ensembles (AIME) was introduced. AIME employs varied ensembles allocated to several layers by combining them into one integrated system being generated automatically so that one ensemble acts as an integral part of another one. Experiments presented in this paper relied on the standard ten-fold cross validation to investigate the effectiveness of AIME ensembles carried out for the quinary classification, CAN5. The results obtained demonstrated that

several blood biochemistry attributes can be used to supplement the Ewing battery for the detection of CAN in the absence of a complete set of Ewing tests. The results show that AIME provided higher accuracy as a multitier CAN5 classification system. For CAN5 categorization, the best accuracy of 99.57% was obtained by the AIME that combined Decorate as the top layer with Bagging on middle layer applied to Random Forest as a base classifier.

9. Binary Classification of Early CAN

Binary classification of early CAN was considered simultaneously with CAN2 and the ternary classification, CAN3 in [74]. It investigated the problem of determining of these categorizations based only on HRV. A variety of measures may be extracted from HRV, including time domain, frequency domain, and more complex non-linear measures. Among the latter, Renyi entropy has been proposed as a suitable measure that can be used to discriminate CAN from normal healthy patients. There are several different ways that can be used to calculate various variants of the Renyi entropy, which depend on a number of parameters. The paper [74] compares nine different methods to calculate Renyi entropy by applying several variations of the histogram method and a density method based on sequences of RR intervals. The effectiveness on nine methods in achieving the best separation of the different categories of CAN3 is then compared. The results obtained showed that that the histogram method using single RR intervals yields an entropy measure that is either incapable of discriminating CAN from controls, or it provides little information that could not be gained from the standard deviation (SD) of the RR intervals. In contrast, probabilities calculated using a density method based on sequences of RR intervals yielded an entropy measure that provided good separation between groups of participants and provided information not available from the SD. This showed that different approaches to calculating probability for determining the Renyi entropy may affect the success of detecting CAN3 categories. Thus, the results of [74] bring clarity to the question of how best to calculate the Renyi entropy for the successful detection of CAN3 categories.

10. Binary Classification Severe CAN

The paper [75] applied multiscale Allen factor to determine a marker for cardiac neuropathy from ECG recordings as features to be used for the machine learning methods and automated detection. It introduced the Graph-Based Machine Learning System (GBMLS). This method is intended to enhance the effectiveness of the diagnosis of severe diabetic neuropathy. We applied it to the multiscale Allen factor (MAF) features as a collection of attributes determined from the recorded ECG biosignals. These attributes can be collected as a result of routine ECG investigation of patients regardless of the presenting medical problems. The experiments compare sensitivity and specificity of the automated detection produced by GBMLS with analogous outcomes achieved by various other machine learning approaches. To this end the authors used a comprehensive collection of important

classifiers and clusterers available in the open source machine learning software package Scikit-learn. The experiments have demonstrated that the best outcomes were obtained by GBMLS in combination with MAF, which improved sensitivity to 0.89 and specificity to 0.98 and outperformed several other classifiers and clusterers including Random Forest with sensitivity of 0.83 and specificity of 0.92.

11. Future work

Future work will further propose new hierarchical data mining methods in classification of CAN progression in combination with new HRV parameters that are robust against short heart rate recordings and allow patients to self-monitor disease progression with data integrated at a central site on the cloud for population studies.

Since conceptual graphs are effective tools for formal visual reasoning in the medical domain [61,62], it would be interesting to investigate the applications of conceptual graphs for the visualization of the diagnostics of CAN progression.

Acknowledgments

The authors wish to thank Bev de Jong for technical assistance in collecting the data for the DiabHealth screening, Roche Australia for providing the glucose sticks and measuring instruments. The authors are grateful to two anonymous reviewers for comments and corrections that have helped to improve the text of this paper.

Conflict of Interest

All authors declare no conflicts of interest in this paper.

References

1. Fernandez-Llatas C, García-Gómez JM (2014) *Data mining in clinical medicine*. New York: Springer.
2. Cerrito P (2010) *Cases on health outcomes and clinical data mining: studies and frameworks*. Hershey, Pennsylvania, USA: IGI Global.
3. Epstein I (2010) *Clinical data-mining: integrating practice and research*. Oxford, UK: Oxford University Press.
4. Petillo D, Orey S, Tan AC, et al. (2014) Parkinson's disease-related circulating microRNA biomarkers—a validation study. *AIMS Med Sci* 2: 7-14.

5. DeMarshall CA, Sarkar A, Nagele RG (2015) Serum autoantibodies as biomarkers for Parkinson's disease: background and utility. *AIMS Med Sci* 2: 316-327.
6. Ervin K, Pallant J, Terry DR, et al. (2015) A descriptive study of health, lifestyle and sociodemographic characteristics and their relationship to known dementia risk factors in rural Victorian communities. *AIMS Med Sci* 2: 246-260.
7. Shinde S, Mukhopadhyay S, Mohsen G, et al. (2015) Biofluid-based microRNA biomarkers for Parkinson's disease: an overview and update. *AIMS Med Sci* 2: 15-25.
8. White VJ, Nayak RC (2015) Re-circulating phagocytes loaded with CNS debris: a potential marker of neurodegeneration in Parkinson's disease? *AIMS Med Sci* 2: 26-34.
9. Khalid KE, Nsairat HN, Zhang JZ (2016) The presence of interleukin 18 binding protein isoforms in Chinese patients with rheumatoid arthritis. *AIMS Med Sci* 3: 103-113.
10. Fitzmaurice MJ, Adams K, Eisenberg JM (2002) Three decades of research on computer applications in health care: medical informatics support at the agency for healthcare research and quality. *JAMIA* 9:144-160.
11. Hage I, Hamade R (2015) Automatic detection of cortical bone's Haversian osteonal boundaries. *AIMS Med Sci* 2: 328-346.
12. Vinik AI, Erbas T, Casellini CM (2013) Diabetic cardiac autonomic neuropathy, inflammation and cardiovascular disease. *J Diabetes Investig* 4: 4-18.
13. Jelinek HF, Abawajy JH, Kelarev AV, et al. (2014) Decision trees and multi-level ensemble classifiers for neurological diagnostics. *AIMS Med Sci* 1: 1-12.
14. Jelinek HF, Abawajy JH, Cornforth D, et al. (2015) Multi-layer attribute selection and classification algorithm for the diagnosis of cardiac autonomic neuropathy based on HRV attributes. *AIMS Med Sci* 2: 396-409.
15. Deshpande AD, Harris-Hayes M, Schootman M (2008) Epidemiology of diabetes and diabetes-related complications. *Phys Ther* 88: 1254-1264.
16. Colagiuri S, Colagiuri R, Ward J (1998) *National diabetes strategy and implementation plan*. Canberra: Paragon Printers.
17. Pop-Busui R, Evans GW, Gerstein HC, et al. (2010) The ACCORD Study Group. Effects of cardiac autonomic dysfunction on mortality risk in the Action to Control Cardiovascular Risk in Diabetes (ACCORD) Trial. *Diab Care* 33: 1578-1584.
18. Spallone V, Ziegler D, Freeman R, et al. (2011) Cardiovascular autonomic neuropathy in diabetes: clinical impact, assessment, diagnosis, and management. *Diab Metab Res Rev* 27: 639-653.
19. Ziegler D. (1994) Diabetic cardiovascular autonomic neuropathy: prognosis, diagnosis and treatment. *Diab Metabol Rev* 10: 339-383.
20. Gerritsen J, Dekker JM, TenVoorde BJ, et al. (2001) Impaired autonomic function is associated with increased mortality, especially in subjects with diabetes, hypertension or a history of cardiovascular disease. *Diab Care* 24: 1793-1798.

21. Pappachan JM, Sebastian J, Bino BC, et al. (2008) Cardiac autonomic neuropathy in diabetes mellitus: prevalence, risk factors and utility of corrected QT interval in the ECG for its diagnosis. *Postgrad Med J* 84: 205-210.
22. Zoppini G, Cacciatori V, Raimondo D, et al. (2015) Prevalence of cardiovascular autonomic neuropathy in a cohort of patients with newly diagnosed type 2 diabetes: the Verona newly diagnosed type 2 diabetes study (VNDS). *Diabetes Care* 38: 1487-1493.
23. Jelinek HF, Imam HM, Al-Aubaidy H, et al. (2013) Association of cardiovascular risk using nonlinear heart rate variability measures with the Framingham risk score in a rural population. *Frontiers Physiol* 4: 1-8.
24. Tesfaye N, Seaquist ER (2010) Neuroendocrine responses to hypoglycemia. *Ann NY Acad Sci* 1212: 12-28.
25. Hoeldtke RD, Boden G (1994) Epinephrine secretion, hypoglycemia unawareness, and diabetic autonomic neuropathy. *Ann Intern Med* 120: 512-517.
26. Balcioğlu AS, Müderrisoğlu H (2015) Diabetes and cardiac autonomic neuropathy: Clinical manifestations, cardiovascular consequences, diagnosis and treatment. *World J Diab* 6: 80-91.
27. John SC, Easton JD (2003) Are patients with acutely recovered cerebral ischemia more unstable? *Stroke* 4: 24-46.
28. Ko SH, Kwon HS, Lee JM, et al. (2006) Cardiovascular autonomic neuropathy in patients with type 2 diabetes mellitus. *J Korean Diab Assoc* 30: 226-235.
29. Stranieri A, Abawajy J, Kelarev A, et al. (2013) An approach for Ewing test selection to support the clinical assessment of cardiac autonomic neuropathy. *Artif Intell Med* 58: 185-193.
30. Devoe JE, Gold R, McIntire P, et al., (2011) Electronic health records vs Medicaid claims: completeness of diabetes preventive care data in community health centers. *Ann Fam Med* 9: 351-358.
31. Bellazzi R, Ferrazzi F, Sacchi L (2011) Predictive data mining in clinical medicine: a focus on selected methods and applications. *Wiley Int Rev Data Mining Know Discover* 1: 416-430.
32. Habibi S, Ahmadi M, Alizadeh S, (2015) Type 2 diabetes mellitus screening and risk factors using decision tree: results of data mining. *Glob J Health Sci* 7: 304-310.
33. Ali R, Hussain J, Siddiqi MH, et al. (2015) H2RM: a hybrid rough set reasoning model for prediction and management of diabetes mellitus. *Sensors (Basel)* 15: 15921-15951.
34. Agelink MW, Malessa R, Baumann B, et al. (2001) Standardized tests of heart rate variability: normal ranges obtained from 309 healthy humans, and effects of age, gender and heart rate. *Clin Auton Res* 11: 99-108.
35. Ewing DJ, Martyn CN, Young RJ, et al. (1985) The value of cardiovascular autonomic functions tests: 10 years experience in diabetes. *Diab Care* 8: 491-498.
36. Pumprla J, Howorka K, Groves D, et al. (2002) Functional assessment of HRV variability: physiological basis and practical applications. *Int J Cardiol* 84: 1-14.

37. Ewing DJ, Campbell JW, Clarke BF (1980) The natural history of diabetic autonomic neuropathy, *Q J Med* 49: 95-100.
38. Khandoker AH, Jelinek HF, Palaniswami M (2009) Identifying diabetic patients with cardiac autonomic neuropathy by heart rate complexity analysis, *Biomed Eng OnLine* 8: 1-12.
39. Cornforth D, Jelinek HF (2007) Automated classification reveals morphological factors associated with dementia. *Appl Soft Comput* 8: 182-190.
40. Malliani A (2005) Heart rate variability: from bench to bedside. *Eur J Intern Med* 16: 12-20.
41. Cornforth DJ, Tarvainen MP, Jelinek HF (2014) Visualization methods for assisting detection of cardiovascular neuropathy. Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE; 26-30 Aug. 2014, pp. 6675-6678.
42. Dietrich DF, Schindler C, Schwartz J, et al. (2006) Heart rate variability in an ageing population and its association with lifestyle and cardiovascular risk factors: results of the SAPALDIA study. *Europace* 8: 521-529.
43. Ho YL, Lin C, Lin YH, et al. (2011). The prognostic value of non-linear analysis of heart rate variability in patients with congestive heart failure – a pilot study of multiscale entropy. *PLoS One* 6: 1-6.
44. Goldberger AL, Amaral LAN, Hausdorff JM, et al. (2002) Fractal dynamics in physiology: Alterations with disease and aging. *PNAS* 99: 2466-2472.
45. Sturmberg JP, Bennett JM, Picard M, et al. (2015) The trajectory of life. Decreasing physiological network complexity through changing fractal patterns. *Front Physiol* 6: 1-11.
46. TFESC/NASPE (1996) Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Eur Heart J* 17: 354-381.
47. Peng CK, Havlin S, Stanley HE, et al. (1995) Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos* 5: 82-87.
48. Oida ET, Moritani KT, Yamori Y (1999) Diabetic alteration of cardiac vago-sympathetic modulation assessed with tone–entropy analysis. *Acta Physiol. Scand* 165: 129-135.
49. Lake DE, Moorman JR (2011) Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. *Am J Physiol Heart Circ Physiol* 300: H319-325.
50. Tarvainen MP, Cornforth DJ, Jelinek HF (2014) Principal component analysis of heart rate variability data in assessing cardiac autonomic neuropathy. Engineering in Medicine and Biology Society (EMBC2014), 36th Annual International Conference of the IEEE, 26-30 Aug. 2014, 6667-6670.
51. Witten H, Frank E, Hall MA (2011) *Data mining: practical machine learning tools and techniques with java implementations*. 3ed, New York, Sydney: Morgan Kaufmann.
52. Huda S, Jelinek HF, Ray B, et al. (2010) Exploring novel features and decision rules to identify cardiovascular autonomic neuropathy using a Hybrid of Wrapper-Filter based feature selection.

- In: Marusic S, Palaniswami M, Gubbi J, et al, editors. Intelligent sensors, sensor networks and information processing, ISSNIP 2010. Sydney: IEEE Press, 297-302.
53. Dietterich TG (1998) Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Comput* 10: 1895-1924.
54. Bouckaert, RR, Frank E, Hall M, et al., WEKA manual for version 3-7-13, <http://www.cs.waikato.ac.nz/ml/weka/>, viewed 15 May 2016.
55. Hall M, Frank E, Holmes G, et al. (2009) The WEKA data mining software: an update. *SIGKDD Explor* 11: 10-18.
56. Baumert M, Schlaich MP, Nalivaiko E, et al. (2011) Relation between QT interval variability and cardiac sympathetic activity in hypertension. *Am J Physiol Heart Circ Physiol* 300: H1412-1417.
57. Kelarev AV, Dazeley R, Stranieri A, et al. (2012) Detection of CAN by ensemble classifiers based on ripple down rules. *Lect Notes Artif Int* 7457: 147-159.
58. Kelarev AV, Stranieri A, Yearwood JL, Jelinek HF (2012) Empirical study of decision trees and ensemble classifiers for monitoring of diabetes patients in pervasive healthcare. Proceedings of Network-Based Information Systems, NBIS2012, 26-28 September 2012, Melbourne, Australia, pp 441-446.
59. Abawajy J, Kelarev AV, Stranieri A, Jelinek HF (2012) Empirical investigation of multi-tier ensembles for the detection of cardiac autonomic neuropathy using subsets of the Ewing features. Workshop on New Trends of Computational Intelligence in Health Applications, CI-Health 2012. *CEUR Workshop Proceed* 944: 1-11.
60. Kelarev AV, Abawajy J, Stranieri A, Jelinek HF (2013) Empirical investigation of decision tree ensembles for monitoring cardiac complications of diabetes. *Int J Data Warehousing Mining* 9: 1-18.
61. Kamsu-Foguem B, Tchuenté-Foguem G, Foguem C (2014) Conceptual graph operations for formal visual reasoning in the medical domain. *IRBM—Innovation Res BioMed Eng* 35: 262-270.
62. Kamsu-Foguem B, Tchuenté-Foguem G, Foguem C (2014) Using conceptual graphs for clinical guidelines representation and knowledge visualization. *Inform System Front* 16: 571-589.
63. Karmakar CK, Khandoker AH, Jelinek HF, et al. (2013) Risk stratification of cardiac autonomic neuropathy based on multi-lag tone–entropy. *Med Biol Eng Comp* 51:537-546.
64. Chowdhury M, Abawajy J, Kelarev A, Jelinek HF (2016) A clustering-based multi-layer distributed ensemble for neurological diagnostics in cloud services. *IEEE Trans Cloud Comp*. DOI: 10.1109/TCC.2016.2567389.
65. Kelarev A, Stranieri A, Yearwood J, et al. (2012) Improving classifications for cardiac autonomic neuropathy using multi-level ensemble classifiers and feature selection based on random forest. Data mining and analytics, 11th Australasian Data Mining Conference, AusDMm 2012. *Conferences Res Practice Inform Techno* 134: 93-102.
66. Williams GJ (2009) Rattle: a data mining GUI for R. *R J* 1: 45-55.

67. Williams G (2011) *Data mining with Rattle and R: the art of excavating data for knowledge discovery (use R!)*. New York, Dordrecht, Heidelberg, London: Springer.
68. Chi CL, Street NK, Katz C (2010). A decision support system for cost-effective diagnosis. *Artif Intell Med* 50:149-161.
69. Imam MH, Karmakar C, Khandoker AH, et al. (2014) Analysing cardiac autonomic neuropathy in diabetes using electrocardiogram derived systolic-diastolic interval interactions. *Comp Cardiol* 41: 85-88.
70. Abawajy J, Kelarev A, Chowdhury M, et al. (2013) Multistage approach for clustering and classification of ECG data. *Comput Meth Prog Bio* 112: 720-730.
71. Abawajy J, Kelarev A, Chowdhury M, et al. (2013) Predicting cardiac autonomic neuropathy category for diabetic data with missing values. *Comput Biol Med* 43: 1328-1333.
72. Jelinek HF, Yatsko A, Stranieri A, et al. (2015) Diagnostic with incomplete nominal/discrete data. *Artif Intell Res* 4: 22-35.
73. Abawajy J, Kelarev A, Chowdhury MU, et al. (2016) Enhancing predictive accuracy of cardiac autonomic neuropathy using blood biochemistry features and iterative multi-tier ensembles. *IEEE J Biomed Health Informatics* 20: 408-415.
74. Cornforth D, Tarvainen M, Jelinek HF (2014) How to calculate Rényi entropy from Heart Rate Variability, and why it matters for detecting cardiac autonomic neuropathy. *Front Bioeng Biotechnol* 2: 1-7.
75. Jelinek HF, Cornforth DJ, Kelarev AV (2017) Machine learning methods for automated detection of severe diabetic neuropathy. *J Diab Compl Meds* (in print).



AIMS Press

© 2016 Herbert F. Jelinek et al., licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)