International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 08, 2020
ISSN: 1475-7192

PREDICTION OF CARDIAC ARRHYTHMIA USING RECURRENT NEURAL NETWORKS GATED RECURRENT UNITS WITH CAUSAL DIAGRAM

[1]U.M. Prakash, [2]Anagha Chakravarthi, [3]Gargi Singh Shekhawat

ABSTRACT - *Cardiac Arrhythmia is a medical condition in which irregularity in the heart beat is observed. The aim of this paper is to detect said arrhythmia using a dataset containing various patient details. A popular technique is used namely Recurrent neural networks (RNN) in which a comparatively newer technique is used i.e. Gated recurrent units. This is again implemented with the observations from causality diagram to localize on the right attributes to work with.*

Keywords*— Cardiac Arrhythmia, Prediction, Recurrent neural networks (RNN), Long short-term memory, causality and Gated recurrent units****.***

I. INTRODUCTION

A group of conditions in which the irregularity in heartbeat is observed is termed as Cardiac arrhythmia. The irregularity is when the rate of heartbeat is either too slow (Bradycardia), below 60 beats/minute, or too fast (Tachycardia), above 100 beats/minute. Arrhythmia of many types have no symptoms and their direct consequences are cardiac arrest, hence the major scope for this project. At times, it is tough for even a medical professional to detect an arrhythmia as there are very minute irregularities. Therefore, using machine learning techniques to detect an arrhythmia would be very beneficial for the doctor and also the patient. The project aims at using Recurrent neural networks Gated recurrent technique with the observations from causality diagram.

II. DATA SET

For this project, the database is extracted from the UCI Machine Learning repository. It has (452) rows where each row represents a different patient's medical record. There are

279 attributes such as weight, age and other data. This data is related to patient's ECG. The data set has 16 types of classes. Class 1 corresponds to no arrhythmia i.e. the normal ECG. Different types of arrhythmia are

[1] *Department of CSE. SRM Institute of Science and Technology Chennai-603203, India, prakash.mr@ktr.srmuniv.ac.in*

[2]*UG student, Department of CSE. SRM Institute of Science and TechnologyChennai-603203, India, anaghachakravarthi_ravindraraju@srmuniv.edu.in*

[3]*UG student, Department of CSE. SRM Institute of Science and TechnologyChennai-603203, India, gargisingh_di@srmuniv.edu.in*

Received: 21 Dec 2019 | Revised: 18 Jan 2020 | Accepted: 05 Feb 2020

International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 08, 2020
ISSN: 1475-7192

referred by classes2 to 15 and class 16 represents unlabeled patients. Since, detecting at an early stage provides for a much higher chance of recovery and survival, this study will aim at predicting different into their specific categories using machine learning techniques such as RNN_GRU and diagram.

Sr. No.	Attribute	Type	No. of such Attributes	Remarks
1	Heart Rate	Linear	1	No. of heart beats per minute
2	Sex	Nominal	1	Sex of the patient
3	Age, Height, Weight	Linear	3	Physical Description of Patient
4	Duration of wave	Linear	5	Measured in milliseconds
5	Vector angles	Linear	5	Measured in degrees
6	Number of intrinsic deflections	Linear	12	
7	Areas of all segments divided by 10	Linear	12	Area=width*height/2
8	QRSTA	Linear	12	QRSA + 0.5 * width of T wave * 0.1 * height of T wave
9	Existence of ragged wave	Nominal	36	
10	Existence of diphasic derivation	Nominal	36	
11	Average width of wave	Linear	60	Measured in milliseconds
12	Amplitude of wave	Linear	96	Multiplied with 0.1*millivolt

Figure 1: Attributes of DataSet

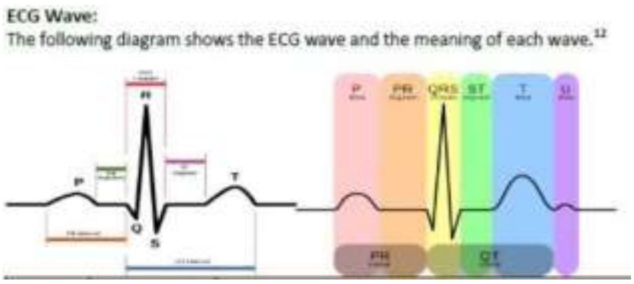

Figure 2

Figure 3: Attributes of the ECG

Sr. No.	Predictor	Description
1	P Wave	Represents atrial depolarization
2	PR interval	Measured from the beginning of the P wave to the beginning of the QRS complex
3	ORS	Represents ventricular depolarization
4	J-Point	The point at which the QRS complex finishes and the segment begins
5	ST segment	Represents the period when the ventricles are depol
6	T wave	Represents ventricular repolarization

Received: 21 Dec 2019 | Revised: 18 Jan 2020 | Accepted: 05 Feb 2020

III. DATA PREPROCESSING

The data we obtained have several columns with single values and also some columns with missing values. Therefore, from the data set these columns were deleted. The resulting data set contained 257 features and 452 instances.

IV. DATA CLEANING

Th data set consists of 279 attributes of 452 patients. The attributes are as shown in the section above. The first step was to replace the missing values with null values. After that the columns were identified with same value in all the 452 rows. Out of all the attributes, 17 of them were found to have a single value. According to our observations, these columns have not shown any change, hence have been deleted from the dataset.

The missing values in the data set were identified in the next step. Missing values were observed in five columns. (Figure 4)

Attribute	No. of missing values
J_Angle	375
P_Angle	22
T_Angle	8
QRST_Angle	1
Heart	1

Figure 4 : Columns with missing values

Class	No. of instances
7	3
8	1
14	4
15	5
16	22

Figure 5 : Instances in chosen classes

The columns J Angle and P Angle were removed as they had a large number of values that were sparse. The sparse values in the other columns T Angle, QRST Angle were replaced with the average value of the same column. It was observed that the classes (Figure 3) had minimal examples of data. For our model, there is no issue of multiple instances as in the end, we will be combining all such instances into a single class named "arrhythmia".

Merging of classes – As seen earlier, classes 2-16 have been combined as a single one.

Sr. No.	Class	Diagnosis	Instances
1	1	Arrhythmia	207
2	2	Normal	245

Figure 6 : Instances

V. REASONING FOR APPROACH

According to the model, we have created for the detection of arrhythmia there is a binary bucket system where the result may either be arrhythmic or not. The class under which arrhythmia falls is considered the positive one. The following describes a step-by-step approach for implementing the model. The approach uses GRU (Gated Recurrent Unit) to predict the arrhythmia from the ECG dataset. To make the Neural Networks more efficient and to have better accuracy, causality is studied of different attributes in the dataset. The attributes with major causalities are chosen and a final causality plot is drawn for the visualization purposes. Thenon the causality plots markings are done of both arrhythmatic and normal condition of the patient. The studied causalities are then fed to the GRU neural networks.

VI. MODELS AND OBSERVATIONS

The following sections give insight into the observations received through the diagram and Gated recurrent units.

A. Causal Learning

To establish a causality between variables a controlled study is used. It is the most proven way. In a controlled study, the sample is divided in two, in which either groups are compared with each other in several aspects. These two groups end up receiving treatment according to the results given by them. The only way we can make predictions about how our distribution changes as a consequence of an interaction is if we know how the variables are causally related. This information about causal relationships is shown through casual diagram.

Here Causal Relations were used to find the causality between different attributes of the data set and following were the observations :

QRS and QRST Angles were found to have the maximum causality constant of .646 (64.6 %)

trueOther significant causalities which were observed in
the data set are (please refer to the Figure 7)true

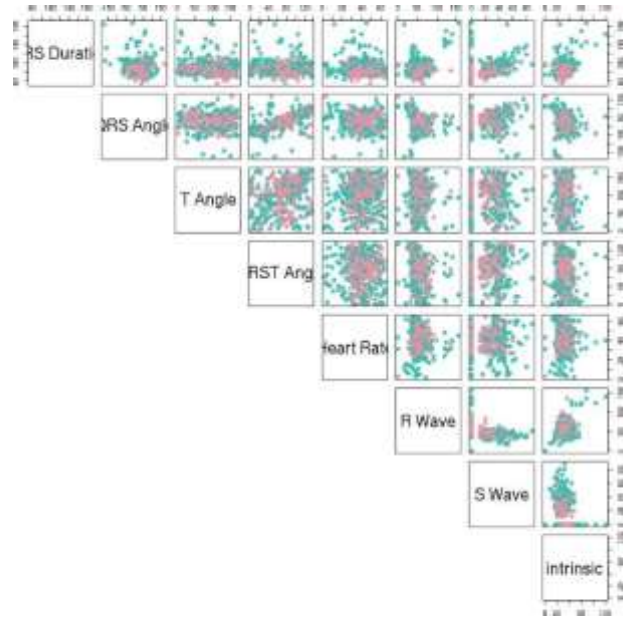

Figure 7 : Casual diagram

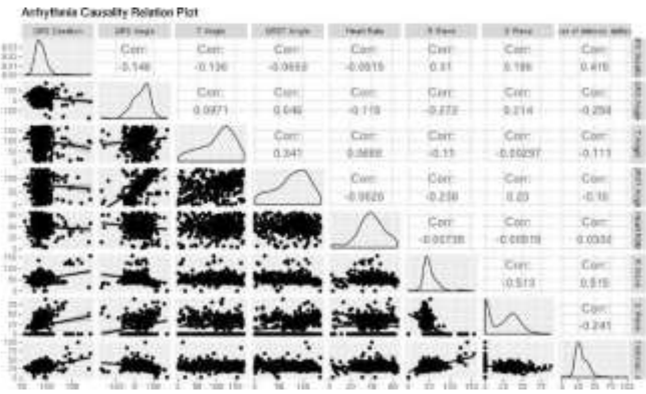

Figure 8 : Arrhythmia Causality Relation plot

- "R wave" and "Number of intrinsic deflections" -> 51.5%
- "QRS Duration" and "Number of intrinsic deflections" -> 41.5%
- "T Angle" and "QRST Angle" -> 34.1%

B. Recurrent *neural network*

In a recurrent neural network (RNN), connections between nodes are portrayed like a graph that is temporal. It is a type of artificial neural network. These graphs display a temporal dynamic behavior. RNNs use their inherent internal memory to save data.

1) Long short term memory (LSTM)

Long short-term memory (LSTM), a type of RNN, is used in the field of deep learning. It can process single data points (such as images), as well as complete sequences of data.

2) Gated recurrent units (GRU)

Gated recurrent units () are a gating mechanism in RNNs. GRUs show better performance attributes on smaller datasets.

Here, the studied causalities were to the GRU neural networks. As observed, main six attributes were taken in consideration –

R wave, Number of intrinsic deflections, QRS Duration, Number of intrinsic deflections, T Angle and QRST Angle

Observation:

The RNN-GRU(Recurrent Neural Network-Gated RecurrentUnit) Model :average prediction accuracy was observed as 84.9375. This prediction accuracy is a marginal improvement when compared to the prediction accuracy reported by the papers surveyed by us. The results are summarized in Table 1.

Minimum Recorded Accuracy	77.5%
Average Accuracy	84.9375%
Maximum Recorded Accuracy	90.625

Table 1 : Accuracy

VII. FUTURE WORK

As we mainly focused on prediction and improving its accuracy, we will then can shift our focus towards

2019-02-19T16:48:05.835087:0.835938step 701, loss 0.400085, classification of the arrhythmia as well. Various machinelearning algorithms can tested. We believe hybrid algori

2019-02-19T16:48:05.959743: step 702, loss 0.275614,0.90625will yield better results for classificatio

2019-02-19T16:48:06.080708:0.867188step 703, loss 0.321383,

VII. CONCLUSION

2019-02-19T16:48:06.204138: step 704, loss 0.332307,0.820312

2019-02-19T16:48:06.328098: step 705, loss 0.315942,0.851562

2019-02-19T16:48:06.452439: step 706, loss 0.27562,0.890625

2019-02-19T16:48:06.576073: step 707, loss 0.318855,0.84375

2019-02-19T16:48:06.701481: step 708, loss 0.30665,0.859375

2019-02-19T16:48:06.826317: step 709, loss 0.346533,0.8437

2019-02-19T16:48:06.890256: step 710, loss 0.5002, 0.775

The RNN-GRU model runs with an average accuracy of

84.9375 with maximum recorded accuracy of 90.625 and minimum of 77.5. The data was recorded on 700 different iterations of the model.

REFERENCES

1. CLoPAR: Classification based on Predictive AssociationThe paper presents the implementation of RNN-GRU with the help of causality diagram with a few techniques. The Rules M. Naderi Dehkordi, M. H. Shenassa
2. Prediction and Classification of Cardiac Arrhythmia
3. Vasu Gupta, Sharan Srinivasan, Sneha S Kudli
4. A kind of hybrid algorithm using rough set and SVM Lai- Sheng Wang, Yi-Tian Xu, Li-San Zhao
5. Machine Learning in Cardiac Health Monitoring and Decision Support System Shurouq Hijazi and Alex Page, Burak Kantarci, Tolga Soyata,
6. Combining Evidence in Hybrid Medical DecisionSupport Models Maurice E. Cohen, Donna L. Hudson
7. Comparative Study of Various Machine Learning Classifiers on Medical Nilima Karankar, Pragya Shukla, Niyati Agrawal
8. Automatic ECGBeat Arrhythmia Detection M.Bazarghan_1, Y. Jaberi 2, R. Amandi 3, M. Abedi 4
9. A Neural Network Architecture Combining Gated Recurrent Unit (GRU) and Support Vector Machine (SVM) for Intrusion Detection in Network Traffic Data
10. Identification of causal pattern using opinion analysis inIndonesian medical texts Susetyo BagasBhaskoro ; Saiful Akbar ; Suhono Harso Supangka bienFred M. Agarap

Received: 21 Dec 2019 | Revised: 18 Jan 2020 | Accepted: 05 Feb 2020