Middlebury

A Deep Learning Algorithm for Detection of Cardiac Arrhythmia Using Synthetic Electrocardiography

Luke Lorenz and Caroline Cutter
Supervised by: Professor Shrutarshi Basu

December 12, 2022

1 Introduction

1.1 Overview

Cardiovascular disease is currently the leading cause of death in the United States and continues to increase due to an aging population[5]. Cardiac arrhythmia, where the heart beats too slowly, too fast, or in an irregular way, is a serious manifestation of cardiovascular disease[5] Current diagnostic methods for atrial fibrillation, Aortic stenosis, and arrhythmia include electrocardiography (ECG), echocardiography, and chest x- rays. ECG's are the most commonly used diagnostic test among patients evaluated for heart-related issues, largely due to its non-invasive approach and relatively low cost. Early detection is vital for treatment and recovery of cardiovascular disease, however, one of the main problems doctors currently face is that diagnosis is done by naked eye, so early signs of arrhythmia are often missed or misdiagnosed [6]. Patients with these abnormalities are frequently asymptomatic for months before rapidly developing severe symptoms, such as dizziness, palpitations, fast heart beat, feelings of weakness, and even sudden heart failure.

1.2 Problem

Due to the high incidence rate of cardiovascular disease, the goal of this project is to explore current deep-learning techniques in an effort to create a model that successfully detects cardiac arrhythmia using synthetic ECG data to aid doctors in diagnosing different cardiovascular diseases. This work is important to study due to the high incident/death rates of cardiovascular disease and that early detection of arrhythmia is key to better outcomes.

2 Related Literature

2.1 Preface

Cardiac arrhythmia is a broad spectrum of disorders of abnormal heart rhythms, including tachycardia (beats too fast), bradycardia (beats too slow), or irregular heartbeat. Arrhythmia affects approximately 1.5 and 5 percent of the general population, however, there is a wide spectrum of ways in which arrhythmia can manifest, including different symptoms, rates, and severity.

2.2 Biological Background

2.2.1 Physiology of the Heart

The heart is a pump responsible for controlling the circulatory system which delivers blood and oxygen throughout the body. The heart consists of four chambers, the left/right atrium, and the left/right ventricle. The heart receives deoxygenated blood into the right atrium, passes blood the right ventricle to re-oxegenate the blood, then passes the re-oxegenated blood to the left atrium then the left ventricle. It is then passed through a large artery called the aorta, then through smaller coronary to the rest of the body. The location of these coronary arteries is significant because it helps doctors understand affected areas with ECG data [6].

The heart is built out of specialized muscles that are organized through connecting fibers that contract in wave-like fashion. The synchronized pumping of the heart is controlled by electrical activity created by pacemaker and non-pacemaker cells. A typical heartbeat begins with pacemaker cells in the sinoatrial node (SA) firing to contract and fill the ventricles with blood. Then, the electrical signal activates the pacemaker cells in the atrioventricular node (AV) which pumps the blood out of the heart [6]. Although very technical, these structures of the heart are important to understand for our project because these structures, muscles and electrical activation gives rise to the characteristics in ECG recordings and allows us to understand areas of defect for different arrhythmia and how to create useful data/ML models based on these characteristics.

2.2.2 Electrocardiogram (ECG)

An electrocardiogram is a recording of the heart's electrical activity during a heartbeat, which graphs voltage over time. This is a non-invasive technique, where electrodes (conductive pads) are attached to the surface of the body to measure the electrical potential difference between leads, or two electrode attachments (Figure 2). There are two categories of leads, limb leads which use the limb electrodes to view the heart in a vertical plane (I, II, III, aVL, aVF, aVR), and precordial leads which view the heart in a horizontal plane (V1 – V6) (Figure 3). The ECG records the electrical cardiac activity between these leads, where depolarization (increase in electrical activity) is currently traveling towards the electrode and repolarization (decrease in electrical activity). Our ECG trace consists of a 12-lead trace for 10 seconds at 250 Hz, as it lines up with average ECG data in the real world.

The ECG trace allows doctors to systematically understand the heart rhythm because typical conductance propagates in a predictable pattern, so deviations can

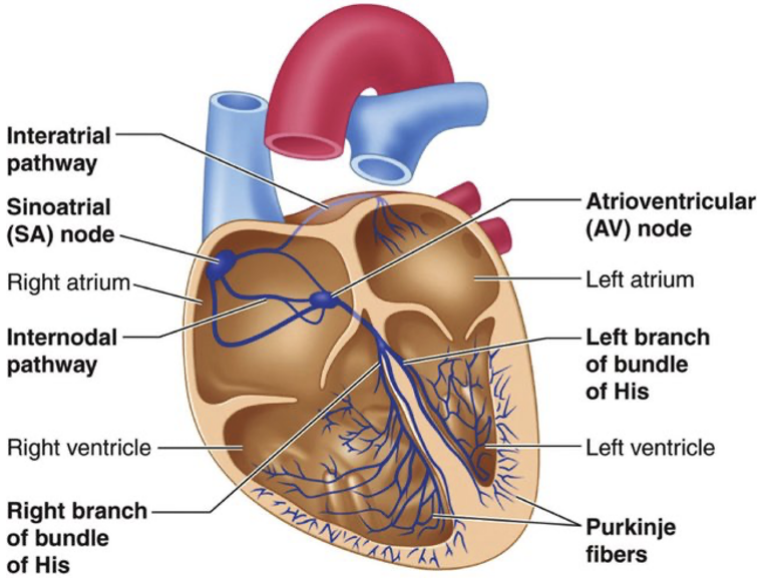

Figure 1: A diagram of the heart which shows the four chambers and the conductance flow.

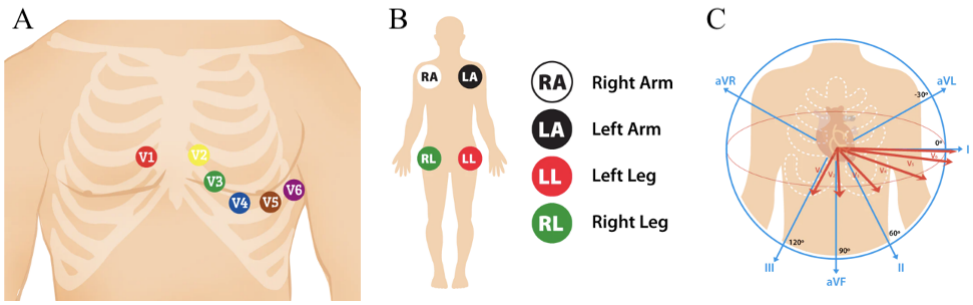

Figure 2: A) Precordial Electrode Placement B) Limb Electrode Placement C) 12 Lead groups (Red = Horizontal plane, Blue = vertical plane)

reveal the location and cause of abnormalities. For our project purposes, we decided to focus on ST interval elevation (Figure 3).

2.2.3 Heart Rate

A heart rate, also known as a pulse, is the number of heart beats per minute (BPM). On an ECG, this is automatically calculated as the time between two successive R waves. Most irregular heart rhythms are triggered by the detection of the variance in the R-R interval, which it is why it is important to look at heart rate while interpreting an ECG. In typical ECG traces, there are grids of 0.04 seconds in each box, so the heart rate would be 1,500 divided by the number of boxes between R waves. An average adult heart rate is 60 – 99 BMP. Bradycardia is less than 60 BPM and tachycardia is greater than 100 BPM, which both could signify abnormal sinus rhythm and therefore an arrhythmia. [6]

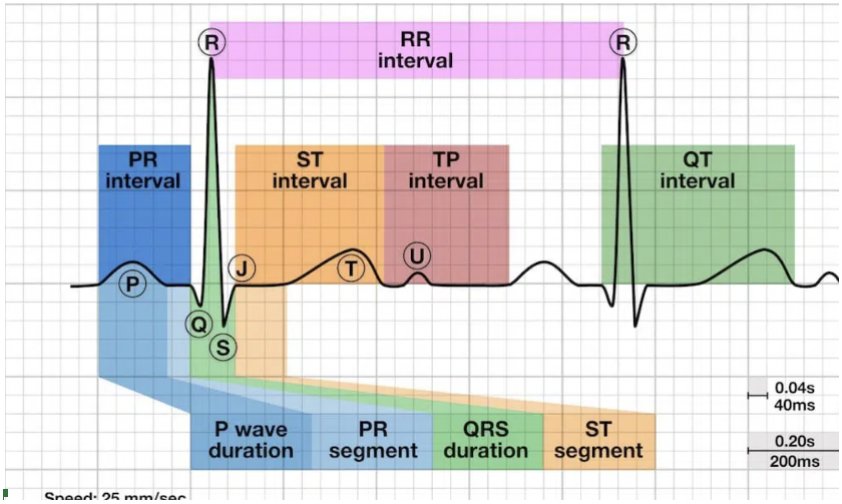

Figure 3: An ECG with segment labels.

2.2.4 Heart Rhythm

There are three main components to an ECG trace – the P wave (depolarization of atria), the QRS complex (depolarization of ventricles), and ST Segment (repolarization of ventricles) (Figure 4).[6]

2.2.5 The ST-Segment and Common Associated Arrhythmias

Our project focused on the ST segment elevation and depression. As stated above, the ST segment is the final ventricular depolarization and the beginning of ventricular re-polarization. Common cardiac problems associated with ST-segment abnormalities include cardiac rhythm disturbances, inadequate coronary artery blood flow, and electrolyte disturbances. Heart attacks (Myocardial Infarction) are a common cause of death globally and most injury patterns of heart attacks are seen within ST segment changes in an ECG. Below are a few of the most common and their causes: [2]

I. Myocardial Ischemia (IHD) is when blood flow to the heart tissue has decreased, which can result in hypoxia. This is associated with depression in the ST segment and is caused by a partial or full blockage in a coronary artery due to coronary artery disease or blood clots. Common symptoms include chest, neck, jaw, shoulder, arm, or back pain, shortness of breath, a faster heartbeat, fatigue, etc. [2]

II. Myocardial Infarction (Heart Attack / MI) is a more severe form of ischemia, where hypoxia from a coronary artery blockage causes myocardial cell death. This is associated with elevation of the ST segment and is caused by prolonged oxygen deprivation and untreated Myocardial Ischemia. [2]

2.3 Machine Learning Algorithms

Recently, sophisticated deep learning algorithms have proven to offer support to clinical experts for premature medical diagnoses. Raw ECG waveform data contains thousands of data points that are not immediately accessible to physicians. Due to the vast amount of data, deep learning algorithms on ECG wave forms

Segment	Physiology	Common Arrhythmia's Depression = lower amplitude Elevation = higher amplitude Wide = Longer interval Short = Shorter interval
P Wave	Atrial depolarization and initiation of the SA node in the right and left atrium	Atrial enlargement, atrial fibrillation
PR Interval	Atrial depolarization and ventricular depolarization at the AV node	Wide – AV block Short – Accelerated AV conditions (WPW and LGL syndromes)
QRS Complex	Depolarization of ventricles as current passes the AV node	Wide – hyperkalemia, bundle branch block, premature ventricular contraction Elevation – left ventricular hypertrophy Short – pericardial effusion, myocardial disease
ST Segment	Final ventricular depolarization and beginning of ventricle repolarization	Elevation –myocardial infarction, angina, pericarditis, myocarditis, hyperkalemia, pulmonary embolism, hemorrhage, left bundle branch block, metabolic abnormalities Depression – myocardial ischemia, angina, digoxin toxicity, hypokalemia, hypothermia, tachycardia
QT Interval	Start of depolarization to end of repolarization of ventricles, highly varied interval	Long – ventricular tachycardia, ventricular fibrillation Short – hypercalcemia, hyperkalemia, hyperthermia
U Wave	Delayed repolarization of muscles	Hypokalemia
J Wave	Point between QRS and ST	Hypothermia

Figure 4: A Table reviewing segments of an ECG and associated arrhythmias

Parameter	Meaning
sampling_rate	Sampling rate, default 250 Hz
duration	Default 10s
gamma	A (12,5) matrix to modify each lead's five spikes' amplitudes
mu_hr_1	The mean of the heart rate
sigma_hr_1	The variance of heart rate
min_noise_1,max_noise_1	The max value and min value of noise
t	The starting position along the circle of each interval in radius
a	The amplitude of each spike;
b	The width of each spike

Figure 5: A Table of Parameters for waveform generation.

have flourished. Deep learning analysis of electrocardiography has accurately detected atrial fibrillation [1, 7], Aortic stenosis [5], Atrial/Mitral regurgitation [4], and arrhythmia [3], among others. Such a diagnostic test could be leveraged as an inexpensive screening tool for premature heart disease. Model predictions could result in focused, careful follow-ups with asymptomatic individuals or a referral to further comprehensive screening. Many, if not all of the authors utilized a CNN for prediction, over other machine learning methods, such as clustering or regression. We anticipate we will do the same.

Machine learning techniques that have been used to detect heart arrhythmia's include Naive Bayes with a weighted approach, 2 SVM's with XGBoost based prediction, Improved SVM (ISVM) based on duality optimization (DO) and XGBoost based prediction [6]. It was found that the accuracy, precision, recall and F1-measure (a function of precision and recall) of the XGBoost algorithm was the highest for all parameters. After finding this out we looked deeper into the XGBoost algorithm for inspiration for our own project. What we found is that the XGBoost have four stages - the pre-processing, the R-peak detection, feature extraction, and then classification of the ECG data to find out if the patient suffers from an arrhythmia [7]. In this paper, we present the construction of our deep learning model with the goal to accurately identify normal and abnormal ECG waveforms.

3 Methods

ECG data was generated through a variation of Pierre Elias's library for machine learning on ECG waveforms. This was a python adaptation of ECGSYN from the

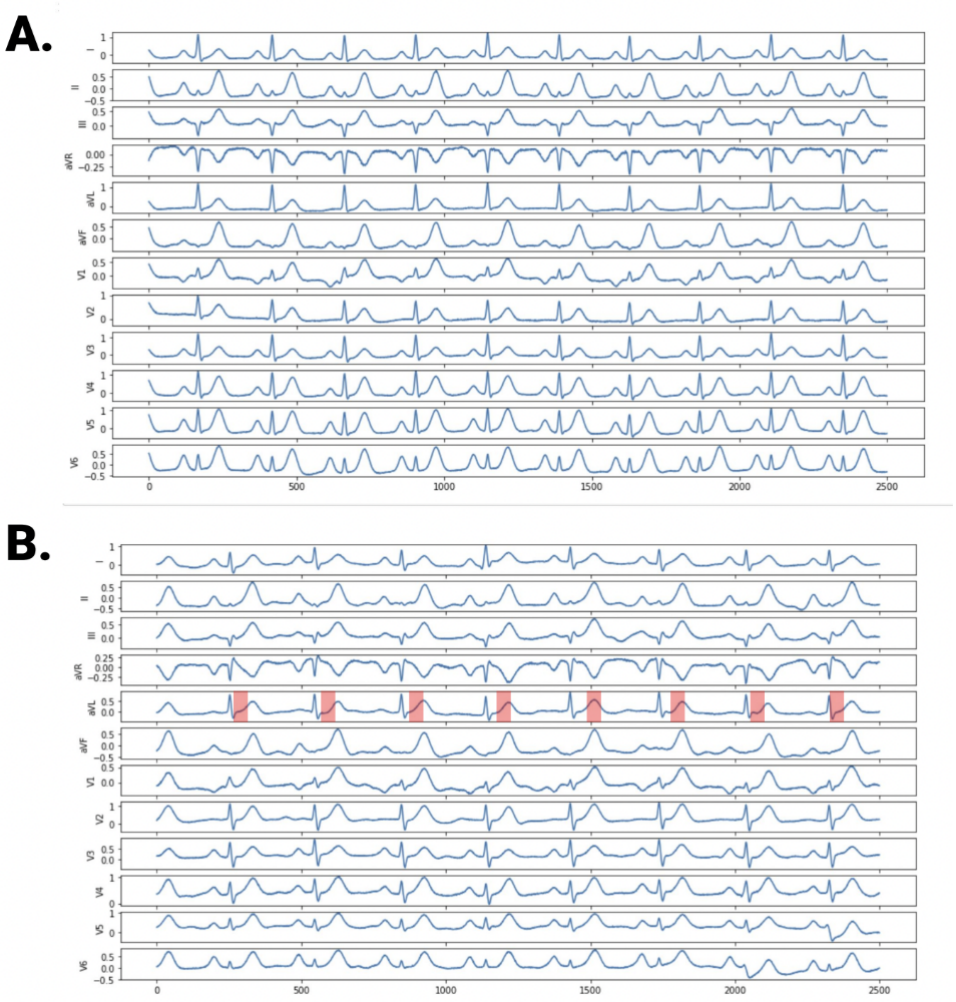

Figure 6: Examples of 12-lead synthetic wave-forms. (A) presents a typical read. In (B), red highlights display mild ST segment elevation in the aVL lead.

physionet database, a web-based resource designed to support current research of clinical data. The parameters of interest were sampling rate, duration, gamma, mu-hr-1, sigma-hr-1, min-noise-1, max0noise-1, t,b, and a (Figure 4). Generating a depression in the ST segment, as well as T segment elevations were done through mutations to the amplitude (a). Using this approach, we generated 60,000 examples.

However, lacking any significant hardware firepower, we were restricted to 12,000 waveforms for training and testing purposes. Of these waveforms, 7000 were normal and 5000 were abnormal. In addition to training hardware, we were restricted to file sizes. Data was compressed using python pickle and burrows-wheeler transforms. It was compressed after generation, and decompressed when accessing it for future use which allowed fluid use on a personal laptop.

Throughout our training process, multiple arrangements of ECG input data were tested. Data was arranged in the following order: ['I', 'II', 'III', 'aVR', 'aVL', 'aVF', 'V1', 'V2', 'V3', 'V4', 'V5', 'V6']. The ECG was recorded over time each example consisted of 12 leads 10 seconds at 250 Hz for a shape of (12,2500). We found that transposing data presented a more realistic and accurate shape to real world data, so the shape of our final input example was (2500,12). For the training set, we split our data in x/x/x, for training/validation/testing, respectively. Neither patient consent nor administrative approval was required, as this data was purely synthetic.

3.0.1 The Algorithm

Our model is an innovative residual neural network. It was implemented using the Keras framework with a Tensorflow backend. Our input shape was of (2500,12). The longer axis (2500) represented the axis of waveform, and most efforts were directed towards convolutions to extract morphological features, while the shorter axis represented the lead axis and was used to fuse data from all leads.

Throughout our training process, we observed a multitude of deep learning models. We began training with a simple densely connected layer, which we found could not accurately account for the size of the data and performed no better than a random classifier. We then tried a standard convolutional neural network, which provided accuracy a few percentage points better, but also with not much luck. We finally settled on a residual neural network.

Our model was composed of 4 residual blocks (Figure 7). 1-dimensional convolutions compress our multitude of data points, while skip connections maintain input from previous layers. Each residual block consisted of two blocks which contained a batch normalization layer to normalize data distribution, a ReLU activation function, and an Add to combine the layers (Figure 7). The first unit in each residual block contained a skip connection which allowed skip connections. Skip connections further allow our relatively large network to be trained easily without vanishing gradients. After the last residual block, our input was fed to a global average pooling layer, which fed to a densely connected neuron and a subsequent prediction.

Our model was trained on a MacBook Air (M2, 2022), with the M2 chip and 8 GB of memory. We acknowledge this hardware as a major limitation to our research. While training, we used an Adam optimizer with an initial learning rate of 0.001. After the development of our model, we input feature data and ECG raw signal of each patient in the validation data into the developed algorithms. We used binary cross entropy as a loss metric, with binary accuracy and AUC as additional metrics.

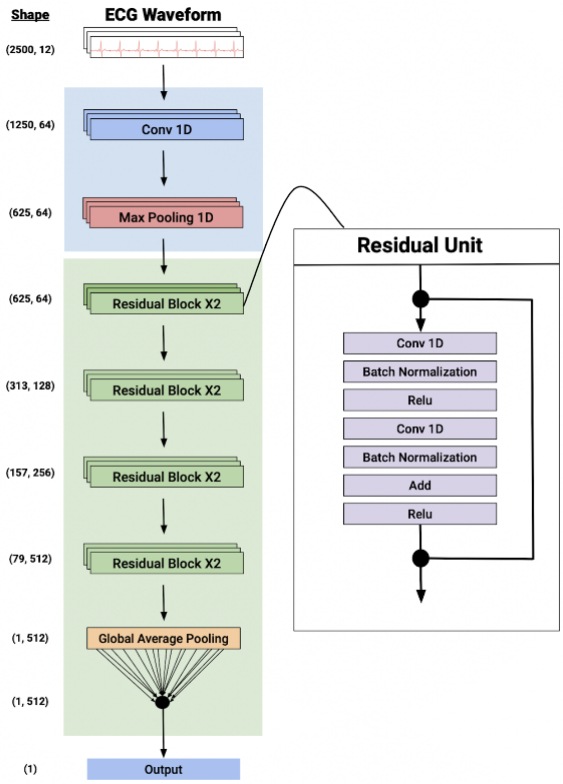

Figure 7: Model structure developed to detect cardiac arrhythmia. Inputs are a synthetic ECG waveform (12 leads 10 seconds at 250 Hz) of shape (2500,12). Outputs binary prediction of abnormality.

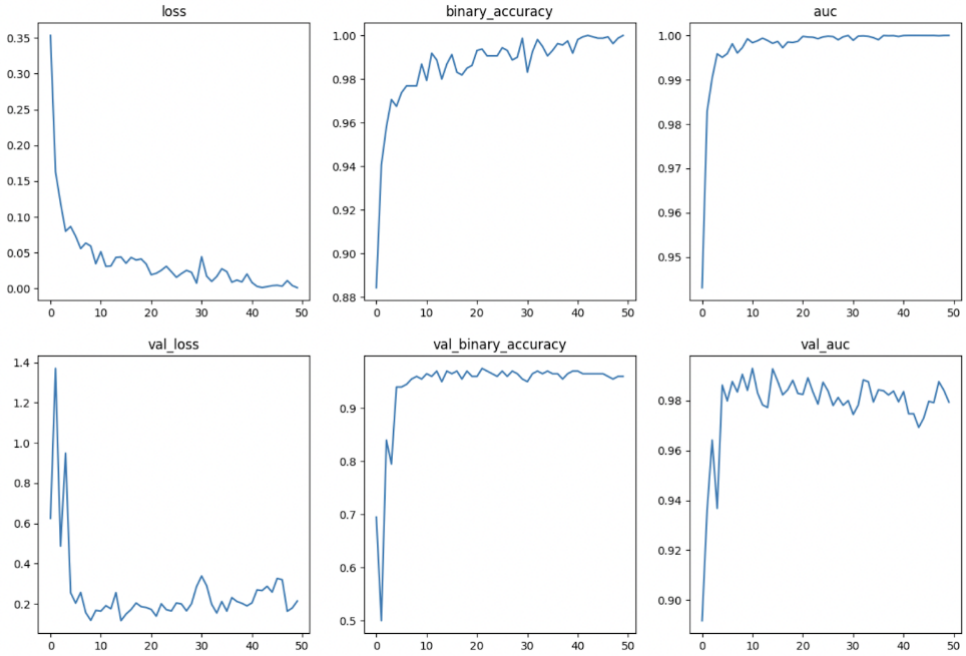

Figure 8: Training metrics for our dataset over 50 epochs. Loss metric is binary cross entropy. Metrics of binary accuracy and AUC were used. Metrics were used on validation sets with 15 percent of total data.

4 Results

We trained the model with (85 percent of size) unique 12-lead ECGs for 25 hours. When our model is applied to an ECG, it outputs a binary prediction score, with a value of 1 corresponding to an abnormal waveform and a 0 signifying normal activity. During internal validation, the metrics of our algorithm were improved over 50 epochs of training (Figure 8). While binary cross entropy continued to decrease over the 50 epochs of training, binary-accuracy peaked around 25 epochs, which is the number of epochs we decided to train our final model on.

We tested our model with 85 percent of random examples from our dataset, split 70 percent testing and 15 percent validation. The performance of our algorithm was confirmed using a test set of 15 percent of our data.

For further exploration, we tested our model on accentuated waveforms, or ECG's that had abnormalities noticeable to the trained eye. For this data, our model achieved an accuracy of 97 percent (Figure 9), and an AUC of 0.98 (Figure 11) after 50 epochs. We found this result was promising and detection of more obvious abnormalities was effortless. While our model accuracy was slightly lower for our first dataset, as abnormalities became more obvious our model excelled.

5 Discussion and Future Work

This paper presents a deep residual neural network trained on a dataset of ECG waveforms. With this model we were able to achieve near perfect accuracy with obvious abnormalities, and mildly decreasing accuracy as fluctuations become minute.

	Binary Cross Entropy	Binary Accuracy	AUC
Miniscule point abnormalities	0.6760	65.8%	0.739
Accentuated abnormalities	0.1248	97.5%	0.989

Figure 9: Table representing metrics on our testing data.

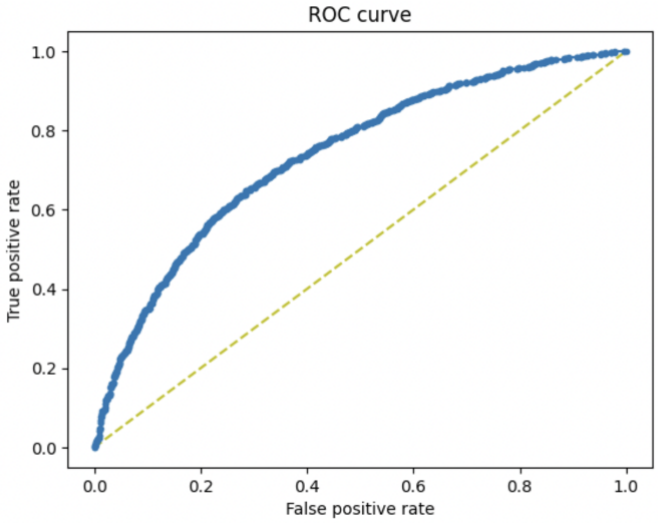

Figure 10: ROC curve for testing dataset. AUC is 0.738.

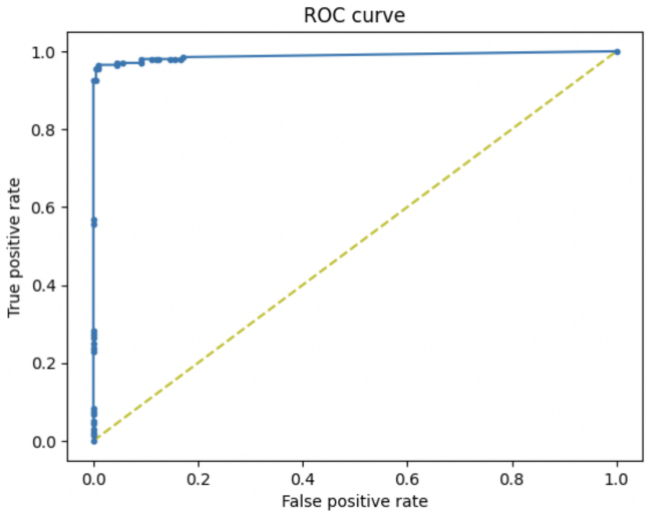

Figure 11: ROC curve for accentuated dataset. AUC is 0.98.

Classically, barriers to entry in machine learning in medicine have been incredibly high. Patient info is secure and available to healthcare providers, making the nature of machine learning in healthcare and medicine rather secluded behind closed doors and high walls. In this paper, we hoped to illuminate an alternative to these hurdles. As we have seen before (cite video game paper), when synthetic data is used in tandem with real world health data, it can be leveraged to provide meaningful premature classification of many sly problems.

Though our classifier achieved successful results, our project contained multiple pitfalls.Ensuring equal and fair classification is incredibly important when considering sweeping impacts of deep learning models in medicine. (Find article for bias) Our model currently lacks demographic and regional data to evaluate our algorithm on different groups, which could (is this good or bad)? We lacked significant computing power and time to train incredibly deep models on increasingly large data sets, which would have further improved model performance. We anticipate doing so would allow robust classification of the mildest arrhythmia.

This project has multiple future extensions. Rather than binary classification, outputting a risk score between 0 and 1 of how likely a patient is could assign soft probabilities of likelihoods and establish a pecking order for priority. As medical doctors oversee dizzying numbers of patients, doing this would allow resource priority to go to those with the highest risk. A second feature to improve our model would be to transition to a categorical classifier of different conditions. Currently, our model only detects abnormalities, but being able to compartmentalize and predict certain conditions is advantageous. Finally, using our trained model on real data to deploy into practice is the ultimate goal. We view this as a stepping stone for advanced machine learning techniques for even more complex cardiac classification in future research.

References

[1] Z. I. Attia, P. A. Noseworthy, F. Lopez-Jimenez, S. J. Asirvatham, A. J. Deshmukh, B. J. Gersh, R. E. Carter, X. Yao, A. A. Rabinstein, B. J. Erickson, S. Kapa, and P. A. Friedman, "An artificial intelligence-enabled ecg algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction," vol. 394, no. 10201, 2019, pp. 861–867. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140673619317210

[2] E. C. de Bliek, "St elevation: Differential diagnosis and caveats. a comprehensive review to help distinguish st elevation myocardial infarction from nonischemic etiologies of st elevation," vol. 18, no. 1. Turkish Jounal of Emergency Medicine, 2018, p. 10.

[3] R. M. Devadas, "Cardiac arrhythmia classification using svm, knn and naive bayes algorithms," vol. 1229, no. 1. IOP Publishing, May 2021, p. 012074.

[4] P. Elias, T. J. Poterucha, V. Rajaram, L. M. Moller, V. Rodriguez, S. Bhave, R. T. Hahn, G. Tison, S. A. Abreau, J. Barrios, J. N. Torres, J. W. Hughes, M. V. Perez, J. Finer, S. Kodali, O. Khalique, N. Hamid, A. Schwartz, S. Homma, D. Kumaraiah, D. J. Cohen, M. S. Maurer, A. J. Einstein, T. Nazif, M. B. Leon, and A. J. Perotte, "Deep learning electrocardiographic analysis for detection of left-sided valvular heart disease," vol. 80, no. 6, 2022, pp. 613–626. [Online]. Available: https://www.jacc.org/doi/abs/10.1016/j.jacc.2022.05.029

[5] K. J. Y. L. K. K. J. P. B. O. Joon-Myoung Kwon, Soo Youn Lee and M. Lee, "Deep learning–based algorithm for detecting aortic stenosis using electrocardiography." Journal of the American Heart Association, 2020. [Online]. Available: https://doi.org/10.1161/JAHA.119.014717

[6] U. Nagavelli, "Machine learning technology-based heart disease detection models," vol. 1, no. 1. IOP Publishing, February 2022, p. 10. [Online]. Available: https://doi.org/10.1155/2022/7351061

[7] Z. Yue and Z. Jinjing, "Atrial fibrillation detection based on EEMD and XGBoost," vol. 1229, no. 1. IOP Publishing, may 2019, p. 012074. [Online]. Available: https://doi.org/10.1088/1742-6596/1229/1/012074