Machine Learning Engineer Nanodegree

Capstone Proposal

Akshay Bhatia

June 2017

Domain Background

Cardiac Arrhythmia is a condition where a person suffers from an irregular or abnormal heart rhythm. It is due to the malfunction in the electrical impulses within the heart that coordinate how it beats. As a result, the heart beats too fast, too slowly, or irregularly. The rhythm of the heart is controlled by a node at the top of the heart, called the sinus node, which triggers an electrical signal that travels through the heart – causing the heart to beat, pumping blood around the body. Excess electrical activity in the top or bottom of the heart means that the heart doesn't pump efficiently. The most common symptoms of Arrhythmia include shortness of breath, fainting, an unexpected loss of heart function and unconsciousness that leads to death within minutes unless the person receives emergency medical treatment to restart the heart. So, it's vital to know about and understand the condition, what danger signs to look out for and how to diagnose it early.

To diagnose Cardiac Arrhythmia early, doctors need to carefully evaluate heartbeats from different locations of the body accurately. Reviewing these fundamental heart sounds (FHSs) for every patient manually is very time consuming for medicians. A potential solution to this is to provide automated diagnosis on the mobile phone. Hence classification of heart sound recordings using Machine Learning techniques could help overcome this problem. Being a patient of Cardiac Arrhythmia, I was inspired to take up this challenge and apply Machine Learning techniques to this domain.

The aim of this project is to build a convolutional neural network that differentiates between normal and abnormal heart sounds.

Problem Statment

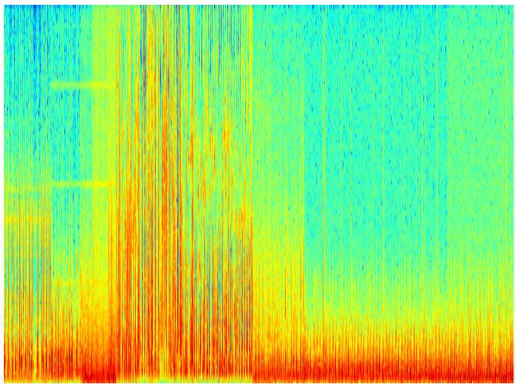
The objective of this project is to classify Phonocardiogram (PCG) recordings or heartbeat recordings as normal or abnormal to quickly identify patients who would require further diagnosis. This is a supervised learning problem since we already know if the heart sound in training dataset is normal or abnormal. The basic idea is to convert each heart sound recording(wav file) to a spectrogram image and train a Convolutional Neural Network. Then given a new PCG recording, we will be able to classify it as normal or abnormal.

Datasets and Inputs

The dataset used for this capstone is available freely as part of the PhysioNet / Computing in Cardiology Challenge 2016 which focuses on automatic classification of normal / abnormal phonocardiogram (PCG) recording. The dataset has 4,430 recordings taken from 1,072 subjects, collected from both healthy subjects and patients with a variety of conditions such as heart valve disease. Along with clean heart sounds, the dataset also contains some noisy recordings. The samples have been obtained from both normal subjects and pathological patients, providing a variety of signal sources. All recordings have already been resampled to 2,000 Hz and are provided in .wav format. Each recording contains only one PCG lead. Each recording has been labelled as normal or abnormal in a separate file(.hea format). Since the dataset is very large in size, I decided to work with a smaller subset. Therefore I choose the training set to be about 1000 labeled recordings and the validation set to be 100 recordings.

Solution Statement

Given a heart sound recording, we will try to classify it as either normal or abnormal. Recent algorithms applied to Cardiology challenges include Heart sound segmentation, transformation of one-dimensional waveforms into two-dimensional time frequency heat map representations using Mel-frequency coefficients and Classification of MFCC heat maps. But given the success of Deep Neural Networks in computer vision, I will try to use CNNs for this project. The recordings in the dataset are based on time-frequency features of the raw signal. So to capture the intensity and the pitch of the recording, I will use spectrograms. Due to its robustness and interpretability, Spectrograms have been applied to a wide range of areas since it a visual way of representing the signal strength, or “loudness”, of a signal over time at various frequencies present in a waveform. I will also perform some data augmentation on the spectrograms obtained from the sound recordings before feeding them into the CNN.



Benchmark Model

Baseline model for this project would be an accuracy of about 60%. This is because randomly selecting a sample from the training dataset as normal would result in accuracy of 60% since the training set has about 600 samples as normal and 400 as abnormal. Also the Physionet Cardiology Challenge 2016 has provided description of the baseline classification method with a score of 0.71 or 71%.

Evaluation Metric

The metrics to determine how well the model performs on the entire dataset is logarithmic loss which is also named as ‘categorial cross entropy’.

$$\text{log-loss} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}\log\left(p_{ij}\right)$$

This can be described as negative the log likelihood of the model given each observation is chosen independently from a distribution that places the predicted probability mass on the corresponding class, for each observation. In our case, each recording is already labeled with one true class and for each one, a set of predicted probabilities is considered an outcome. Here *N* is the number of images in the test set, *M* is the number of image class labels i.e 2, *log* is the natural logarithm, *Yi,j* is 1 if observation belongs to class and 0 otherwise, and P*i,j* is the predicted probability that observation belongs to given class..

Since the dataset is imbalanced i.e number of normal samples is greater than number of abnormal samples in the training dataset, accuracy is not only the metric we will consider.

So the model will be evaluated on ‘precision’ which can be described as the ratio of correct positive predictions made out of the total positive predictions made, ‘Recall ’ which is the ratio of correct positive predictions made out of the actual total that were positive. We also calculate another metric known as ‘fbeta score’ which is given by the weighted average of precision and recall.

Project Design

The project workflow would be as follow :

- **Download the dataset**: The first step of the project will be to download dataset from the PhysioNet website.
- **Convert to spectrograms** : Then write a script to convert the recordings (wav files) to corresponding spectrograms.
- **Pre-Processing (data augmentation)**: This step will include resizing and rescaling images so that they can be easily fed to the CNN
- **Train the model**: Once the images are reshaped and all the necessary pre-processing has been performed, we train the CNN from scratch.
- **Model evaluation**: Finally we evaluate how well the model generalises over new data. So we check performance of the model on the test dataset.

References

https://physionet.org/challenge/2016/

https://www.kaggle.com/kinguistics/heartbeat-sounds

http://www.heart.org/HEARTORG/Conditions/Arrhythmia/AboutArrhythmia/About-Arrhythmia_UCM_002010_Article.jsp#.WTeSZROGMUs

http://www.kdnuggets.com/2016/09/urban-sound-classification-neural-networks-tensorflow.html

https://www.physionet.org/challenge/2016/papers/challenge2016.pdf