Machine Learning Nanodegree
Udacity Connect Intensive

Capstone Proposal

Using Machine Learning to Identify Heart Sounds

Spandan Chakraborty

Resources

For the Capstone Proposal:

- Capstone Proposal Rubric — The criteria against which your capstone project proposal will be assessed.
- Capstone Proposal Template — An outline that guides you through the capstone proposal process. The objectives for each of the seven points are laid out, along with a target length for each section.

For the Capstone Report:

- Capstone GitHub Repo — A folder containing the capstone proposal & report templates, along with example reports.
- Capstone Report Rubric — The criteria against which your Capstone Project final report will be assessed.
- Capstone Report Template — An outline for your Capstone Project final report. It’s useful to take a look at this document before doing all of your coding. Writing sections of your report as you accomplish them can make the Capstone Project report less daunting!

Suggested Problem Areas: from the Capstone Project description: *"Below are a few suggested problem areas you could explore if you are unsure what your passion is"*

- Robot Motion Planning
- Healthcare
- Computer Vision
- Education
- Investment and Trading

Additional Resources : Platforms hosting data science competitions

https://www.kaggle.com/competitions-

https://devpost.com/hackathons

Capstone Proposal : Project Specifications

CRITERIA	MEETS SPECIFICATIONS
Domain Background	Student briefly details background information of the domain from which the project is proposed. Historical information relevant to the project should be included. It should be clear how or why a problem in the domain can or should be solved. Related academic research should be appropriately cited. A discussion of the student's personal motivation for investigating a particular problem in the domain is encouraged but not required.
Problem Statement	Student clearly describes the problem that is to be solved. The problem is well defined and has at least one relevant potential solution. Additionally, the problem is quantifiable, measurable, and replicable.
Datasets and Inputs	The dataset(s) and/or input(s) to be used in the project are thoroughly described. Information such as how the dataset or input is (was) obtained, and the characteristics of the dataset or input, should be included. It should be clear how the dataset(s) or input(s) will be used in the project and whether their use is appropriate given the context of the problem.
Solution Statement	Student clearly describes a solution to the problem. The solution is applicable to the project domain and appropriate for the dataset(s) or input(s) given. Additionally, the solution is quantifiable, measurable, and replicable.
Benchmark Model	A benchmark model is provided that relates to the domain, problem statement, and intended solution. Ideally, the student's benchmark model provides context for existing methods or known information in the domain and problem given, which can then be objectively compared to the student's solution. The benchmark model is clearly defined and measurable.
Evaluation Metrics	Student proposes at least one evaluation metric that can be used to quantify the performance of both the benchmark model and the solution model presented. The evaluation metric(s) proposed are appropriate given the context of the data, the problem statement, and the intended solution.
Project Design	Student summarizes a theoretical workflow for approaching a solution given the problem. Discussion is made as to what strategies may be employed, what analysis of the data might be required, or which algorithms will be considered. The workflow and discussion provided align with the qualities of the project. Small visualizations, pseudocode, or diagrams are encouraged but not required.
Presentation	Proposal follows a well-organized structure and would be readily understood by its intended audience. Each section is written in a clear, concise and specific manner. Few grammatical and spelling mistakes are present. All resources used and referenced are properly cited.

Presentation

Proposal follows a well-organized structure and would be readily understood by its intended audience. Each section is written in a clear, concise and specific manner. Few grammatical and spelling mistakes are present. All resources used and referenced are properly cited

Machine Learning Engineer Nanodegree: Capstone Proposal Template

Domain Background
↓
Problem Statement
↓
Datasets and Inputs
↓
Solution Statement
↓
Benchmark Model
↓
Evaluation Metrics
↓
Project Design

Domain Background

Student briefly details background information of the domain from which the project is proposed. Historical information relevant to the project should be included. It should be clear how or why a problem in the domain can or should be solved. Related academic research should be appropriately cited. A discussion of the student's personal motivation for investigating a particular problem in the domain is encouraged but not required.

Identify Heart Sound Locations

- According to the World Health Organization, cardiovascular diseases (CVDs) are the number one cause of death globally: more people die annually from CVDs than from any other cause. An estimated 17.1 million people died from CVDs in 2004, representing 29% of all global deaths
- Hence any method that helps in pre screening for CVDs is very important
- Recording heart sound with smart phones and digiscope, can be very convenient and inexpensive pre screening mechanism
- A challenge may be general public will not have proper training to place their phone/digiscope to pick up heart sounds
- In this process if machine learning model can assist to locate heart sounds apart from ambient noise and other noise generated inside body from lungs /stomach
- This can be stepping stone to build more sophisticated pre screening mechanism to detect heart conditions very early and get proper medical intervention

https://www.kaggle.com/kinguistics/heartbeat-sounds
http://www.peterjbentley.com/heartchallenge/

Problem Statement

Student clearly describes the problem that is to be solved. The problem is well defined and has at least one relevant potential solution. Additionally, the problem is quantifiable, measurable, and replicable.

The ultimate goal of this project:

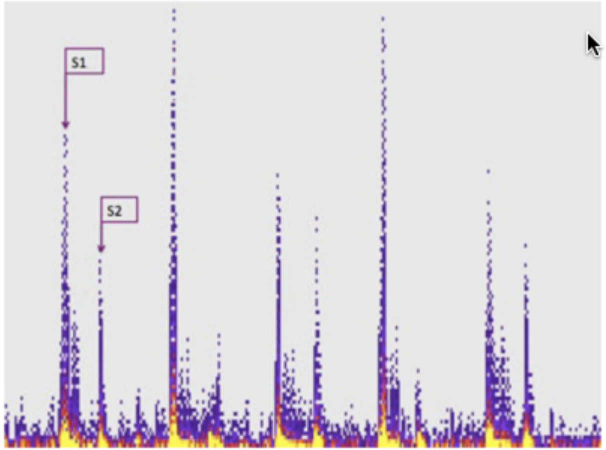
Develop an Sound Classification model that can locate & label two distinctive Heart Sounds (HS) in an audio file

- The proposal is to build a model to locate precise heart sounds (HS) in a audio file
- The audio can be recorded by any one with a smart phone /Digiscope at their own convenience. Since it can pick up ambient noise and heart sound can be very faint in it, the task of the machine learning model is to locate precise locations of heart sounds
- There are two distinctive heart sounds generated from closing valves in heart
- We are all familiar with lub..dub....lub..dub rythm
- lub: corresponds to 1st heart sound: S1
- dub: corresponds to 2nd heart sound : S2
- Precise location of knowing S1 and S2 is the first step to understand if heart function is normal vs any on set of cardiac issues
- Any one can record their Heart Sound through smart phone. The model will classify & locate their heart sounds, the same can be send via email to their physicians to take a look & decide if further consultation is needed
- The model may valuable to build a pre screening app, to discover early symptoms of Cardiac issues.

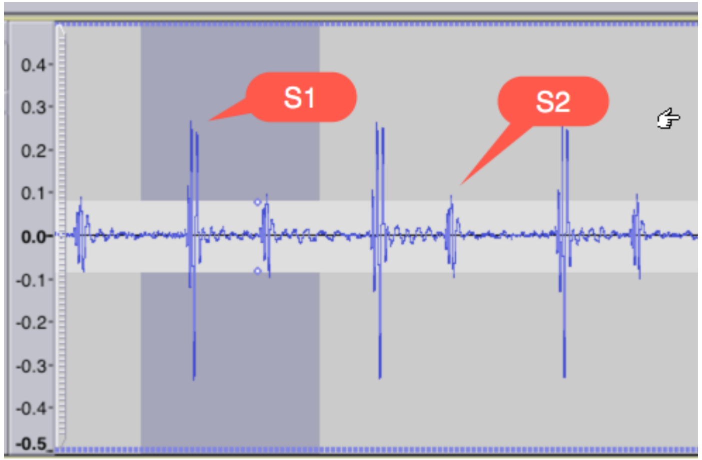
Datasets and Inputs

The dataset(s) and/or input(s) to be used in the project are thoroughly described. Information such as how the dataset or input is (was) obtained, and the characteristics of the dataset or input, should be included. It should be clear how the dataset(s) or input(s) will be used in the project and whether their use is appropriate given the context of the problem.

The dataset used for this project is from Kaggle Competition :
Heartbeat Sounds: Classifying heartbeat anomalies from stethoscope audio
https://www.kaggle.com/kinguistics/heartbeat-sounds



Source: http://www.peterjbentley.com/heartchallenge/



Source: http://www.peterjbentley.com/heartchallenge/

Datasets and Inputs

The dataset(s) and/or input(s) to be used in the project are thoroughly described. Information such as how the dataset or input is (was) obtained, and the characteristics of the dataset or input, should be included. It should be clear how the dataset(s) or input(s) will be used in the project and whether their use is appropriate given the context of the problem.

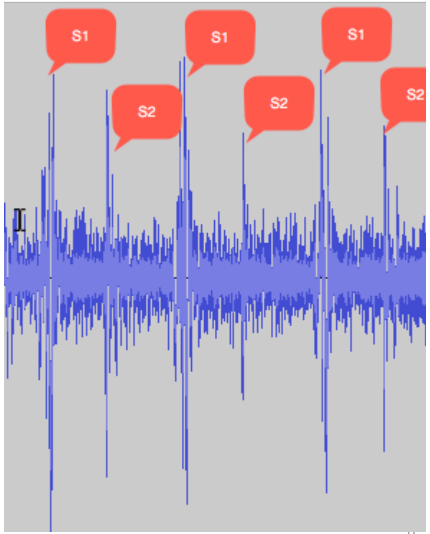
The dataset used for this project is from Kaggle Competition :
Heartbeat Sounds: Classifying heartbeat anomalies from stethoscope audio
https://www.kaggle.com/kinguistics/heartbeat-sounds

Example of Input audio Files:

Index	File Name	Frames	duration in seconds	rate in KHz
0	normal__2011010 70538.wav	391787.0	8.884059	44.0
1	normal__2011011 51127.wav	396900.0	9.000000	44.0
2	normal__2011020 81152.wav	315088.0	7.144853	44.0
3	normal__2011020 81321.wav	347886.0	7.888571	44.0
4	normal__2011022 01230.wav	396900.0	9.000000	44.0
5	normal__2011022 60502.wav	277130.0	6.284127	44.0
6	normal__2011022 70940.wav	396900.0	9.000000	44.0

Location of HS in each audio:

Index	fname	cycle	sound	location	Time In Sec
0	set_a/normal__2011 02081321.wav	1	S1	10021	0.227234
1	set_a/normal__2011 02081321.wav	1	S2	20759	0.470726
2	set_a/normal__2011 02081321.wav	2	S1	35075	0.795351
3	set_a/normal__2011 02081321.wav	2	S2	47244	1.071293
4	set_a/normal__2011 02081321.wav	3	S1	62992	1.428390
5	set_a/normal__2011 02081321.wav	3	S2	73729	1.671859
6	set_a/normal__2011 02081321.wav	4	S1	88761	2.012721



Datasets and Inputs

The dataset(s) and/or input(s) to be used in the project are thoroughly described. Information such as how the dataset or input is (was) obtained, and the characteristics of the dataset or input, should be included. It should be clear how the dataset(s) or input(s) will be used in the project and whether their use is appropriate given the context of the problem.

The dataset used for this project is from Kaggle Competition :
Heartbeat Sounds: Classifying heartbeat anomalies from stethoscope audio
https://www.kaggle.com/kinguistics/heartbeat-sounds

In this project my aim is to tackle the first challenge: Locate HS in given wave files.

- In Data set A there are 21 audio (.wav) files with labeled HS (both S1 and S2)
- To train data I will like to use 10 audio files
- Will keep 5 audio files for validation set
- Will use rest 6 audio files for testing purpose

Solution Statement

Student clearly describes a solution to the problem. The solution is applicable to the project domain and appropriate for the dataset(s) or input(s) given. Additionally, the solution is quantifiable, measurable, and replicable.

Utilize CNN to locate Heart Sound

One of many ways to extract features from audio data in order to train classifiers, is MFCC (Mel-frequency cepstrum), which characterizes human vocal tract generated sounds. MFCCs is also widely used in automatic speech and speaker recognition. We will use the same algorithm to extract sound features from audio given heart sound audio files.

1. Since we know the location of S1 and S2 for training data set, we will label the respective MFCC record with one hot encoding

2. The labeled mfcc data set to be loaded in panda dataframe

3. A Convolutional Neural net will be trained on this training data set (mfcc as feature and labels one hot encoded to represent existense S1, S2 and rest as noise as labels in a frame): will be focused on detecting S1 or S2 in audio file utilizing TensorFlow example for MNIST.

4. Then when a new audio file is presented , mfcc features will be extracted from the audio and based on the trained model the frame containing S1 and S2 will be identified

References:
https://stackoverflow.com/questions/22471072/convolutional-neural-network-cnn-for-audio
https://www.tensorflow.org/get_started/mnist/pros
This follows the same process as described in to detect Foosball goal:
https://humblesoftwaredev.wordpress.com/2016/05/02/an-audio-dataset-and-ipython-notebook-for-training-a-convolutional-neural-network-to-distinguish-the-sound-of-foosball-goals-from-other-noises-using-tensorflow/
More about Mel Frequency Cepstral Coefficient (MFCC):
http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/

Benchmark Model

A benchmark model is provided that relates to the domain, problem statement, and intended solution. Ideally, the student's benchmark model provides context for existing methods or known information in the domain and problem given, which can then be objectively compared to the student's solution. The benchmark model is clearly defined and measurable.

Benchmark: model >= other experiments done in same area

To understand how well the model is doing, one way is to compare with results claimed by others in Kaggle competition or general researchers analyzing the same problem.

I have found research articles claiming F measure score as 0.91, reference below.
Hence I will target for 91% F Score as a bench mark

References:
https://www.researchgate.net/publication/237822466_Detection_and_identification_of_first_and_second_heart_sounds_using_empirical_mode_decomposition

Evaluation Metrics

Student proposes at least one evaluation metric that can be used to quantify the performance of both the benchmark model and the solution model presented. The evaluation metric(s) proposed are appropriate given the context of the data, the problem statement, and the intended solution.

The evaluation metrics typically used in pattern recognition and information retrieval are precision, recall, and F-measure. These metrics serve as the standards for evaluating the effectiveness of a pattern recognition system

TABLE I
EVALUATION METRICS

	Actual Class (Observation)	
Predicted Class (Expectation)	Tp (True positive) Correct result	Fp (False positive) False alarm
	Fn (False negative) Missed detection	Tn (True negative) Correct absence of result

Equations (13)–(15) show the definitions of these metrics.

$$\text{Precision} = \frac{T_p}{T_p + F_p} \tag{13}$$

$$\text{Recall} = \frac{T_p}{T_p + F_n} \tag{14}$$

$$F = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{15}$$

Since it is hard to tell which one is more desirable : high precision Vs high recall. Will use F score as evaluation metrics

References:
https://www.citi.sinica.edu.tw/papers/yu.tsao/5206-F.pdf

Project Design

Student summarizes a theoretical workflow for approaching a solution given the problem. Discussion is made as to what strategies may be employed, what analysis of the data might be required, or which algorithms will be considered. The workflow and discussion provided align with the qualities of the project. Small visualizations, pseudocode, or diagrams are encouraged but not required

1. Obtain the dataset: 21 Labeled Audio Files containing Heart Sound
2. Extract mfcc features for each apprx. 480 frames (our audio files have 44100 frames per second)
3. The same mfcc audio features are used in voice recognition machine learning area
4. Normalize the mfcc data from different audio files
5. Download and install model and dependencies: follow TensorFlow MNIST steps
6. Modify the code with project specifications
7. Train the model with default hyper-parameters
8. Test to evaluate model performance
9. Adjust parameters to maximize performance
10. Implement model