First Increment Report

An Intelligent Context Aware Audio Classification System

By

Zeenat Tariq (22)

&

Sayed Khushal Shah (20)

Team (6-1)

School of Computing and Engineering Department

University of Missouri-Kansas City

Table of Contents

List of Figures

1. Introduction

The report of first increment done for our project is explained in this document. The research project proposes an intelligent audio classification application based on deep learning techniques. During the first increment, we were more focused in finding the material related to existing papers and implemented projects. Currently, we found few implementations in terms of machine learning and shallow learning.

The representation of our intelligent system is described in the document. The designing phase mentioned in the document is only the basic level. This can be changed with the research implementation by adding more features.

1.1. Objectives

- Create a context aware sound application using deep learning techniques
- Real time audio classification using context aware models
- To create an android client application for the real time audio classification
- Amplify the sound according to the hearing capability of the listener

1.2. Approach

A large scale of computations is required to apply Machine Learning Techniques on Mobile application. In this project, we will implement an intelligent mobile application for context aware audio classification. The API will have audio classification features based on deep learning models. We will be using the Convolution neural network model and Neural network model and the result will be compared

2. Related Work

Android platform will be used for the project which is based on Linux for mobile applications. On daily basis, 0.355 million approx. are users added to android platform. The cost of applications on android platform is cheap than other platforms such as iOS. Sound hound is an application that is created for displaying lyrics to the songs that are available in the database. Deep sound is a deep learning model based on tensor flow. Deep Sense project was implemented by using Tensor flow.

We found few models based on Machine learning techniques. The existing implemented models were executed for Audio classification using Decision Tree Algorithm. The data was recorded for 4 seconds which was stored in buffer using spark server. The classification was done for the existing 4 categories which are Classroom, Home, Outdoor and Office. For each category, we have 5 classifications of audios. For example, in home, we have telephone ringing, door knock, doorbell, dog bark and siren. Similarly, for each category the data was trained. The basic filtration techniques were applied using the high pass libraries.

3. Research Design and Specifications

The workflow of the system model is shown in figure 1. Android platform and Tensor Flow will be used for our project. The client sent the real time audio about the product using the android phone. Deep Learning Algorithms will be applied for the classification of audio signals. The audio signal will be transformed and filtered out by applying the filters that are Band pass, Low pass and High pass filters and the distortion will be removed. The transformed data will be sent to the library for feature extraction. This

can be pitch, bandwidth, loudness etc. The supervised models will be applied on the given data. We have created a basic design for our project. The design of the system can be changed with time by adding more features.

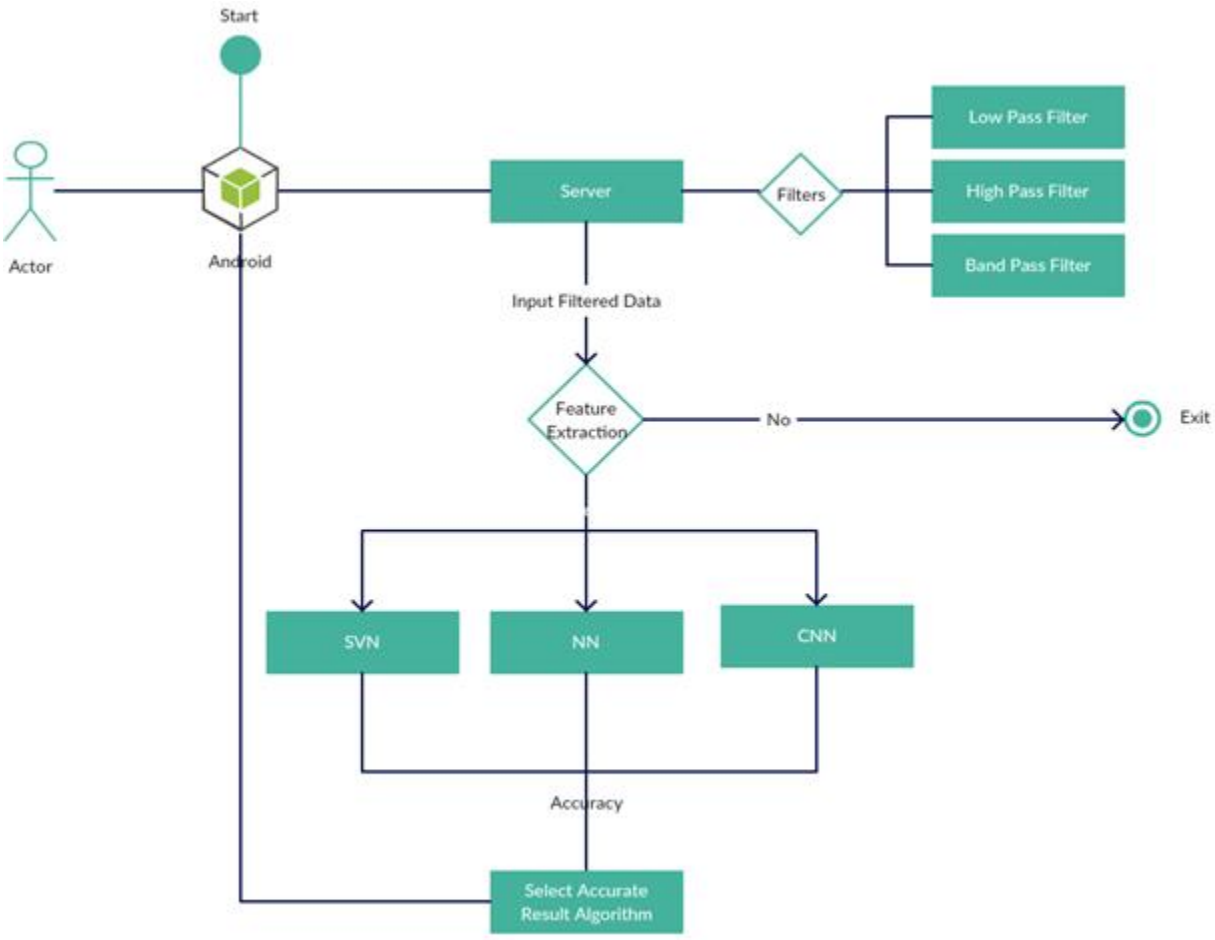

Fig 1. Workflow for Intelligent System

4. Expected Results

- We expect the following results from our project.
- Reduced Time for real time classification of audio signals.
- Accurate Results
- Efficiently trained Deep Learning Models

5. Tentative Project Plan

The following images show the project board, timeline created in Github.

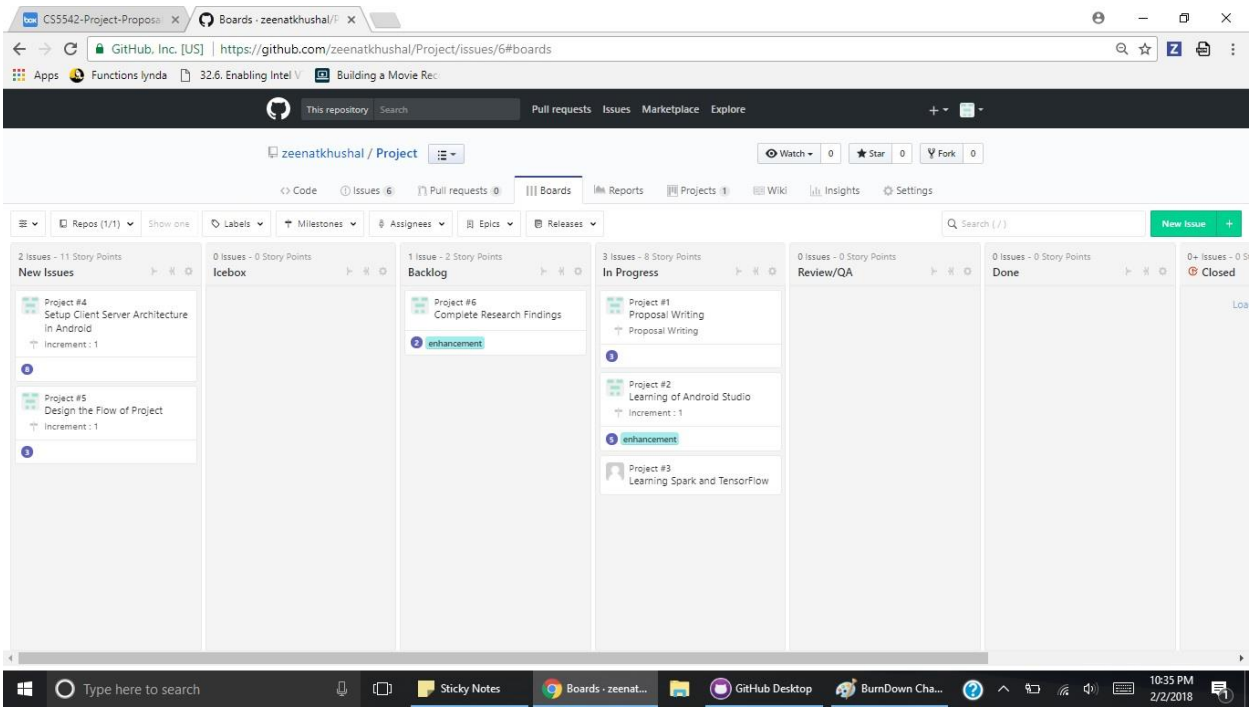

Fig. 2. Project Board in GitHub

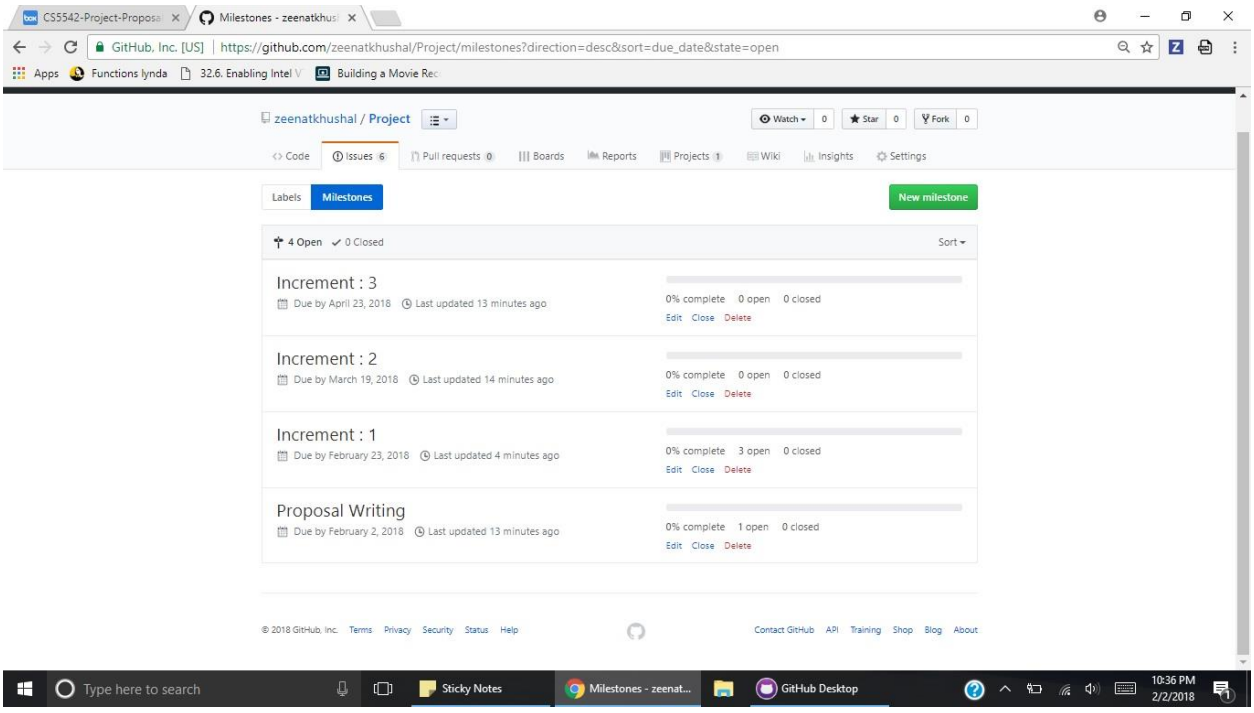

Fig 3. Project Timeline created in GitHub

6. Project Management

6.1 Contribution of Each Member

Sayed Khushal Shah – 50%

Zeenat Tariq – 50%

References

1. https://soundhound.com/
2. https://github.com/deepsound-project/pggan-pytorch
3. https://www.statisticbrain.com/android-phone-statistics/
4. https://github.com/zeenatkhushal/Project/tree/master/Project/Documentation#boards?repos=120057715