CMPE 255 - Data Mining Spring 2018

Project Report

Canvas Group:

Project Groups Section 01 15 Project Groups Section 01

Student Names:

a. Uttara Vishwas Kulkarni (012420811)
b. Nikita Sengupta (012413011)
c. Ujjval Soni (012475645)

Team Name:

Team Insights

Project:

Title: Audio Classification (Multi-Class)

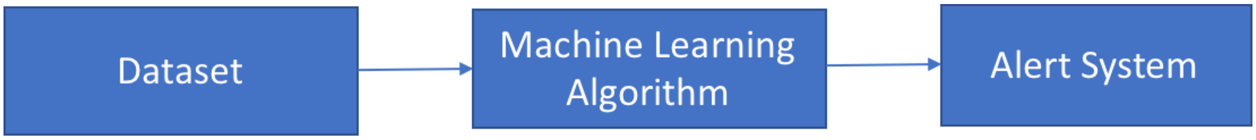
Section 1 Introduction:

- Motivation:
  A person (hearing impaired) will not be able to detect various types of sounds such as a door opening, listening to a fire alarm go off in his building, etc. In some scenarios it could turn out to be a real danger. For instance, assume if a person who has a hearing problem is sleeping in a house alone and someone tries to break in the house and try to harm the person, it would be difficult for him to know what's going on and before he figures something is happening it would already be too late.

- Objective:
  The objective of the project is to resolve the above mentioned issue. We have analyzed the data (sound clippings) of different everyday things such as alarms or cars honking or dogs barking,etc. We have used these sounds and trained our model (Multi-Class Classification) in a way that it can understand a sound in the real world and send a notification on the users phone that the fire alarm is going off in his building or if the door is being opened. And thus the user can take necessary precautions before it's too late.

Section 2 System Design & Implementation detail:

System Design:



Overview of System Design includes, the input Dataset, Machine Learning Logic and the Alert System.

Input Dataset: It consists of Audio Files (50 different types of Sounds / Classes). Mainly these consists of ambient sounds. Data is pre-processed and then fetch to Machine Learning Model.

Machine Learning Model consists of Classification Algorithms to classify audio files into various classes. For this project, we have used classifiers such as SVM, KNN, SG Classifier, etc.

Alert System consists of an Email - Altering System. This system is used to alert an user by sending an email to his id. This email will consists of a type of sound and a message saying there could be a possible event of danger.

Implementation:

- Import Audio Files
- Import librosa library for manipulation and handling of audio data.
- First we used the librosa.display.waveplot function to visualise how different the .wav files are from each other .
- Then we used librosa's in-built functions to extract features from sound wave.
- We used the tempogram, zero-crossing rate, mel-scale spectrogram, spectral centroid, tonal centroid, chromagram , spectral flatness , roll-off frequency to extract features from sound wave .
- We took mean of each individual feature for every sound file present and kept on appending a column in the dataframe .
- After getting the feature values , we normalised them using min-max normalisation so that all the various values gets converted between range 0-1.

- Then we performed various types of split to see which one works best for us .
- We tried train-test split, cross validation , k-fold (k in range 5 -10) , Shufflesplit before proceeding with prediction .
- Once this was done , we fed the data to training model for getting desired production .

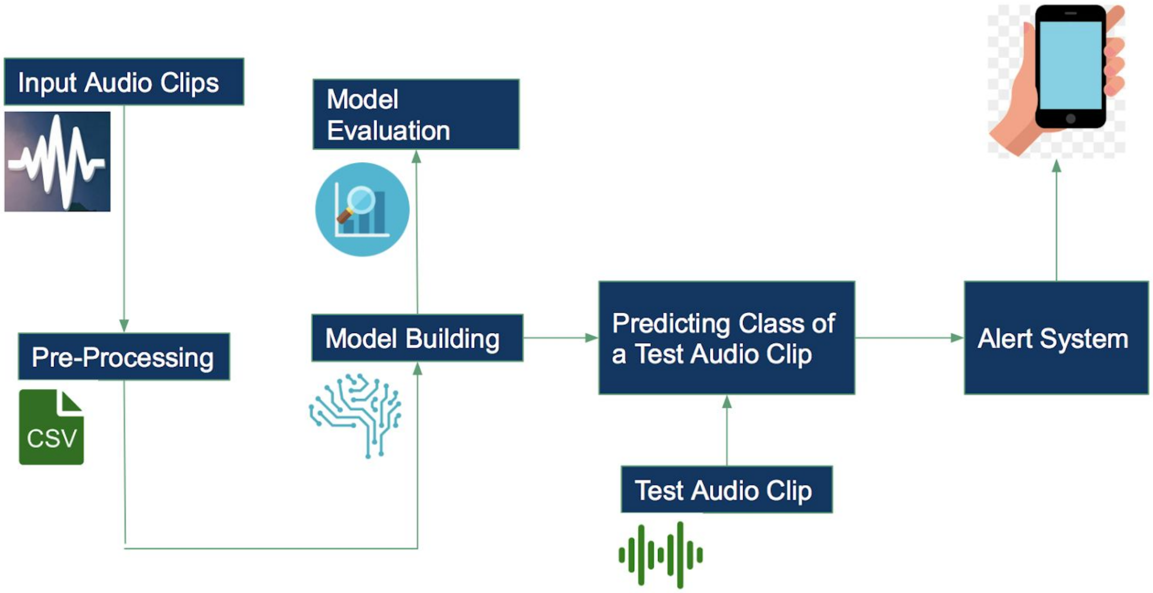
Algorithms Considered:

KNN, SGD Classifier, SVM, Neural Networks (TensorFlow).
Reason: These classifiers are very much relevant to the course work of 255.
Additional Knowledge area: Explored tensorflow model.

Technologies & Tools used:

- Librosa Module (Python): To convert Audio clippings into ML compatible input format.
- Languages and Libraries (Pre-processing, Model Building): Python, Pandas, Numpy, Scikit-Learn Classification

Architecture:

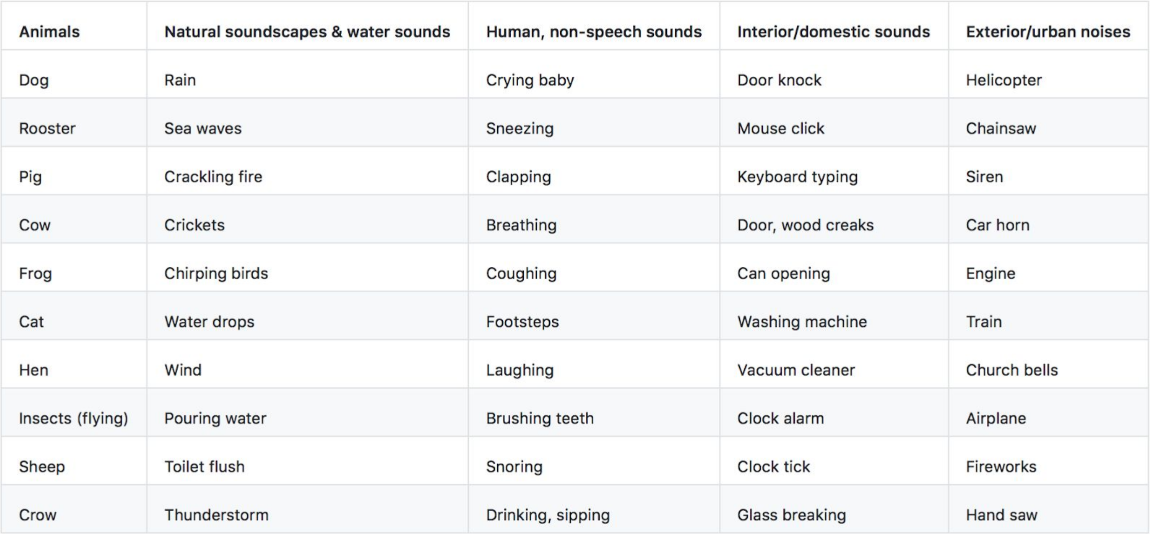


- For this application we have decided to first import the audio clips in our application using Librosa Module (Python).
- We have pre processed the data by removing the fields that were not necessary for our application and then applied feature extraction on those audio files.
- We have used various feature extraction methods as given in the documentation for the librosa module.
- After taking the mean of those features we exported them to a .csv file thus in case if we ran into an issue we won't have to extract those features again.
- Then we split the training data into two parts, one for testing and the other one for training and used various classifier models to find the model that gives our dataset the maximum accuracy.
- For the final step we are sending an email to the user stating there might be an error nearby and thus he/she can take necessary precautions to avoid the danger.

Dataset:

- The ESC-50 dataset is a labeled collection of 2000 environmental audio recording.
- The dataset consists of 5-second-long recordings organized into 50 semantical classes (with 40 examples per class) loosely arranged into 5 major categories.
- With over 1900 classified sample recordings for testing.
- Source: Download ESC-50 dataset

Screenshot of the dataset

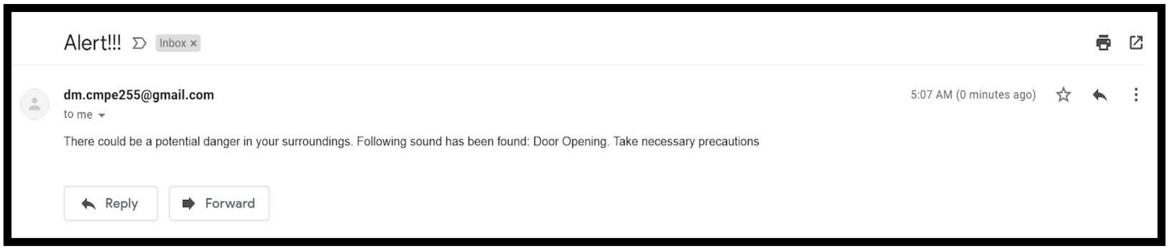


Animals	Natural soundscapes & water sounds	Human, non-speech sounds	Interior/domestic sounds	Exterior/urban noises
Dog	Rain	Crying baby	Door knock	Helicopter
Rooster	Sea waves	Sneezing	Mouse click	Chainsaw
Pig	Crackling fire	Clapping	Keyboard typing	Siren
Cow	Crickets	Breathing	Door, wood creaks	Car horn
Frog	Chirping birds	Coughing	Can opening	Engine
Cat	Water drops	Footsteps	Washing machine	Train
Hen	Wind	Laughing	Vacuum cleaner	Church bells
Insects (flying)	Pouring water	Brushing teeth	Clock alarm	Airplane
Sheep	Toilet flush	Snoring	Clock tick	Fireworks
Crow	Thunderstorm	Drinking, sipping	Glass breaking	Hand saw

Use cases / GUI / screenshots:

Use Case: Altering system for hearing impaired people.

Screenshots:



Alert!!!

dm.cmpe255@gmail.com

to me

5:07 AM (0 minutes ago)

There could be a potential danger in your surroundings. Following sound has been found: Door Opening. Take necessary precautions

Section 3 Experiments / Proof of concept evaluation:

- **Methodology followed:**
  Dataset split techniques used: Train-Test, K-Fold, Cross_Val_Score, Shuffle_split
  For K-fold, we tried k=5 to k=10.

- **Graphs showing different parameters/algorithms evaluated in a comparative manner, along with some supportive text.**

```
<matplotlib.collections.PolyCollection at 0x10279240>
```

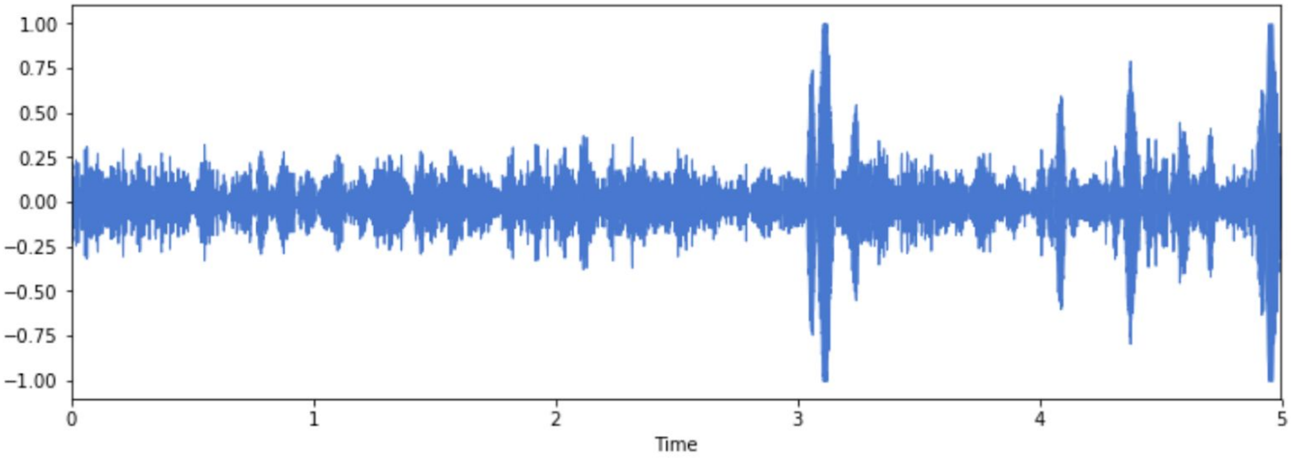


Fig 1. Waveform of wind

```
<matplotlib.collections.PolyCollection at 0x100d8550>
```

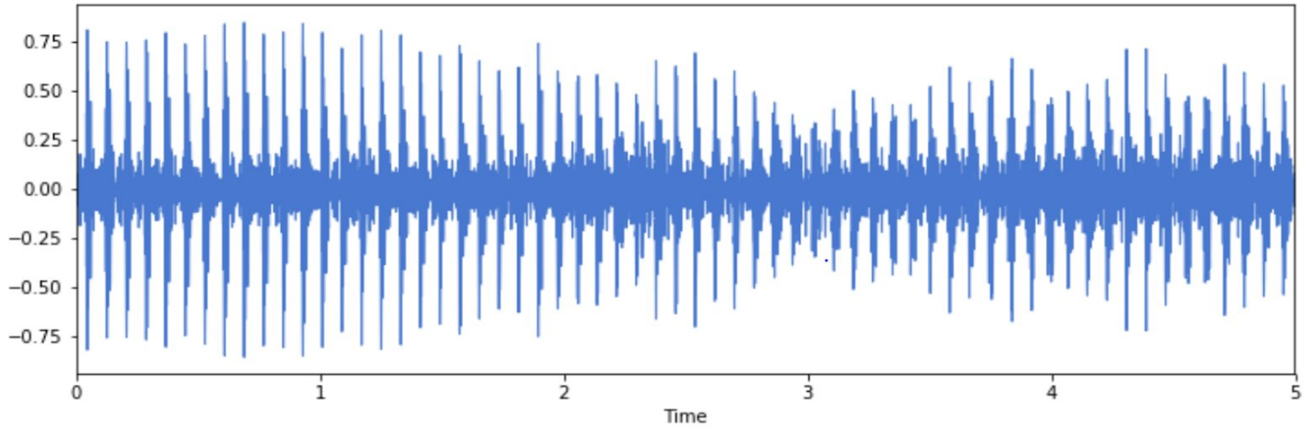


Fig 2. Waveform of Helicopter

```
<matplotlib.collections.PolyCollection at 0xff964e0>
```

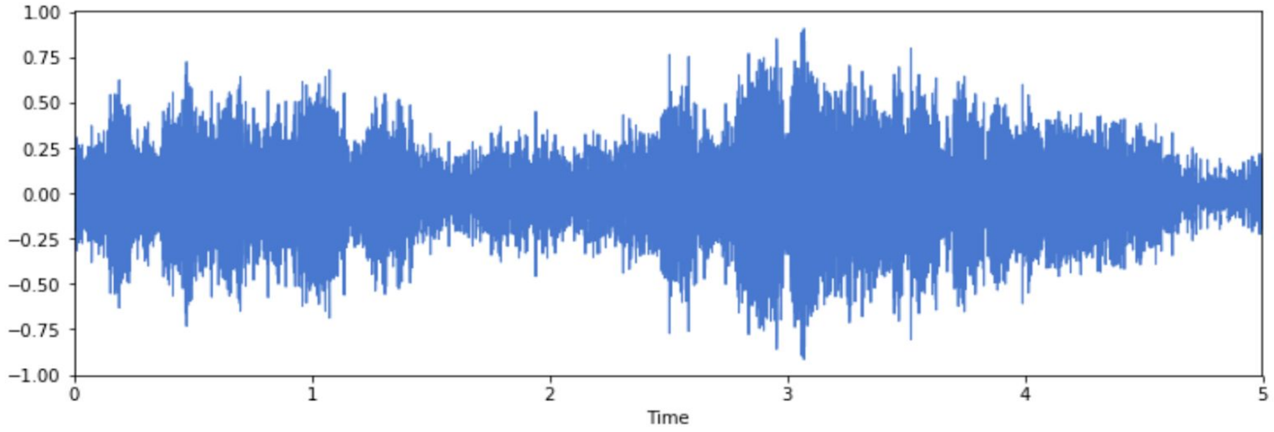


Fig 3. Waveform of Cow Mooing

```
<matplotlib.collections.PolyCollection at 0x14b84e48>
```

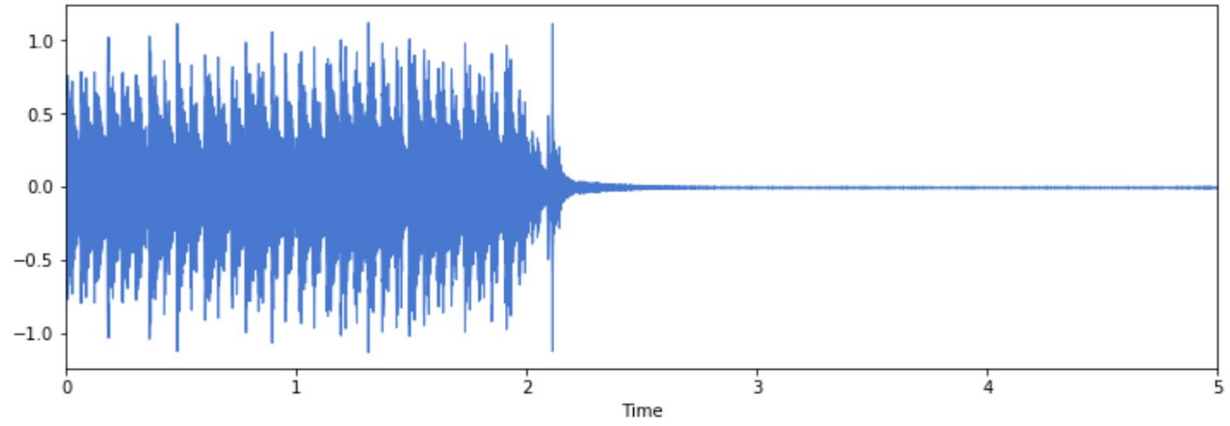
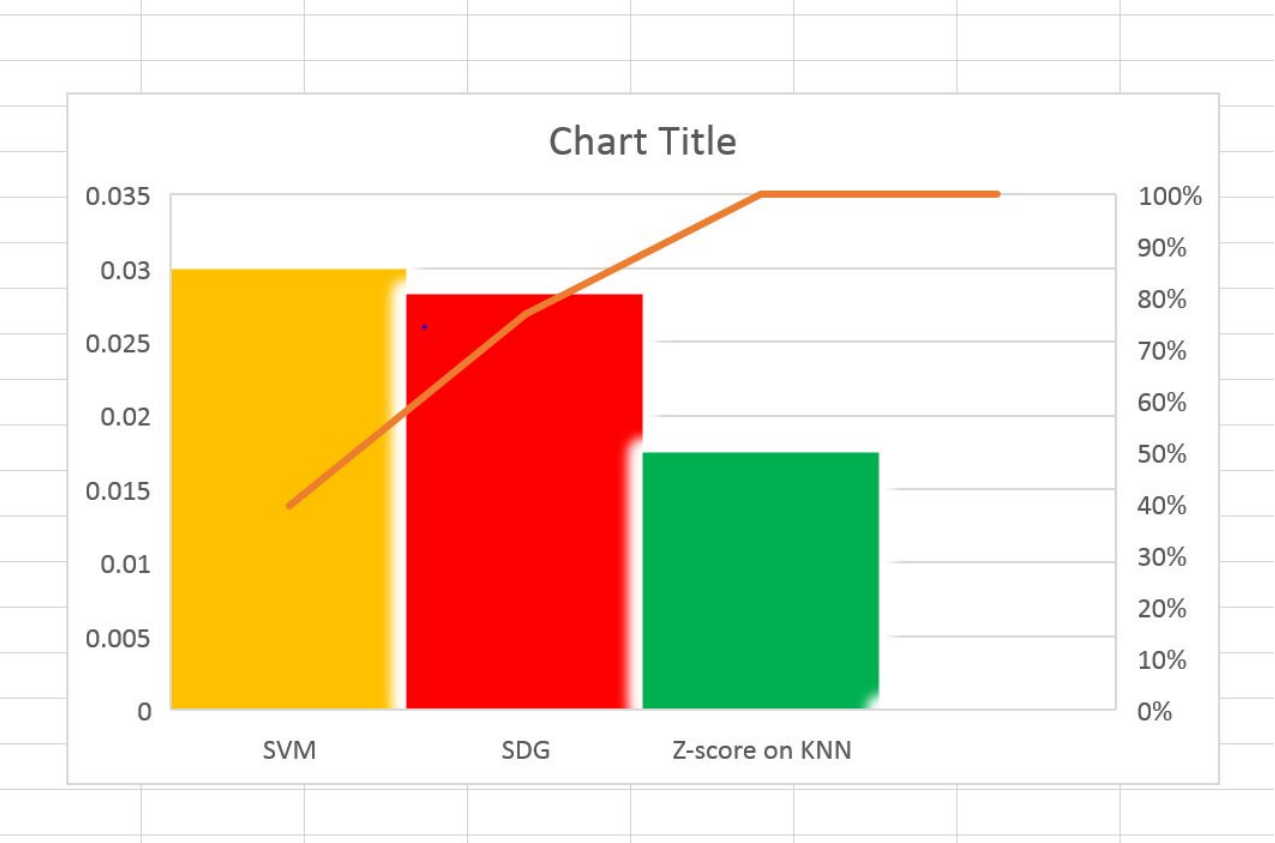


Fig 4. Waveform of clock alarm

- **Analysis of results**
    - After applying different classifiers like KNN, SVM and SDG we have come a conclusion that the SDG classifier works best for our dataset.



Section 4 Discussion & Conclusions:

- **Decisions made:**
    - In the beginning we decided how our application should be developed, we started out by deciding on basic pre processing techniques.
    - Extracting features extraction methods given on official librosa documents.
    - Training the model using various classifiers such as SVM, SDG, KNN and neural networks using tensorflow.
    - And for the final step using the classifier that gives the best accuracy.
- **Difficulties faced**
    - As slowly we started working on the application, we found out it was more difficult than we imagined.
    - Firstly for preprocessing we had decided to take metadata files provided by the dataset, but we failed to see that metadata files had train and testing dataset used by the makers of the dataset but to us on the training data was provided.

    - We had to figure out what files were provided and what weren’t so to be sure when the application is used we have all the files necessary.
    - For the feature extraction step, there were about 10 different features extraction methods that we could use on our dataset. But each of those features had 216 columns in a dataframe and varied number of columns for every method we choose.
    - So that was a difficulty on how to choose the methods and find out the best accuracy for our dataset.
    - We wanted to implement the neural networks using Tensor flow, but we failed to implement it as we weren’t able to solve dimensionality related problems that tensor brought to our application.

- **Things that worked**
    - The preprocessing in the end worked out pretty good, as we got what data we wanted in the meta files.
    - In the feature extraction step we ended up choosing the methods that were best suited for our application.
    - And finally we ended up implementing the all other classifiers besides the neural networks using tensorflow.
    - Also the implementation of sending an email notification worked as planned.

- **Things that didn’t work well :**
    - As mentioned before we ended up completely discarding implementation of neural networks using tensorflow.
    - Also we did mention we only used a few feature extraction methods instead of using all the methods because of the curse of dimensionality.
- **Conclusion**
    - In the end we are pretty satisfied with as, what we had decided in the beginning more things ended up working out and than the things that didn’t.

Section 5 Project Plan / Task Distribution:

- For our application as it was uncertain what classifier or what features or what preprocessing would work, so we experimented with a bunch of different combinations.

TASK TITLE	TASK OWNER	START DATE	DUE DATE
Discuss project objectives	Team	04/02/2018	04/06/2018
Research for Dataset	Team	04/02/2018	04/06/2018
Research technologies and approaches	Individual	04/02/2018	04/06/2018
Research classification algorithms	Individual	04/06/2018	04/13/2018
Research existing tools	Team	04/06/2018	04/13/2018
Assessment of the dataset	Uttara,Nikita,Ujjval	04/14/2018	04/20/2018
Converting Audio files into data	Uttara and Nikita	04/14/2018	04/20/2018
Model Building - KNN	Uttara	04/21/2018	04/30/2018
Model Building - SVM	Nikita	04/21/2018	04/30/2018
Model Building - SDG	Ujjval	04/21/2018	04/30/2018
Post-Processing	Uttara and Nikita	05/01/2018	05/03/2018
Develop the alert system	Ujjval	05/01/2018	05/03/2018
Project Report and Presentation	Team	05/04/2018	05/06/2018