

**T.C.**

**MARMARA UNIVERSITY**

**FACULTY of ENGINEERING**

**COMPUTER ENGINEERING DEPARTMENT**

CSE4088 Introduction to Machine Learning Project Final Report

Title of the Project

*“Sound Signal Classification “*

Group Members

Hasan Mert Yalçın – 150119647

Mertcan Çiy – 150119908

Emirhan Meral - 150119649

1. **Abstract**

Sound Signal Classification is one of the most widely used applications in Deep Learning. It contains learning to categorize sounds and to give notice of the category of that sound.Sound signal classification means categorizing certain sounds in some categories, like environmental sound classification. In modern urban environments, many social signals are disappearing both physically and acoustically with rising environment noise levels. The aim of the project is to ensure that these lost sounds or the sounds that are wanted to be separated are separated and recognized with the help of certain algorithms.

1. **Subtask**

**1)** Import the Libraries.

**2)** Digitase the audio files.

**3)** Extract MFCC's for every audio file in the dataset.

**4)** Iterate through every audio file and extract features using Mel-Frequency Cepstral Coefficients

**5)** Creating X\_train and y\_train Data Structures.

**6)** Perform Label Encoding since we need one hot encoded values for output classes in our model (1s and0s)

**7)** We split dataset as Train and Test

**10)** Predicting the given audio file.

**11)** Plotting the given audio file.

1. **Overview**

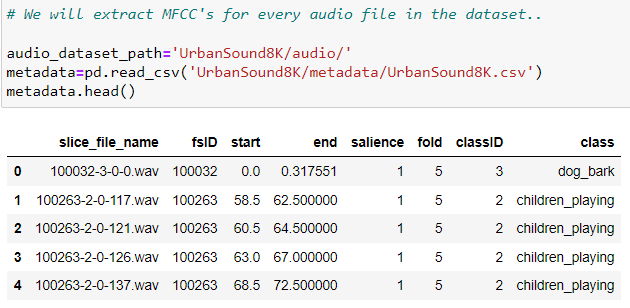
**What is Sound Signal Classification?**

The audio signal classification listens to the audio files, analyze the audio files and extracts the mel frequency cepstral coefficients of these audio files and provides an estimation of the sounds using an estimation method.

**What is Environmental Sound Classification?**

The environmental sound classification process is the subject of our project. In the environmental sound classification, the sounds are taken from the urban life.

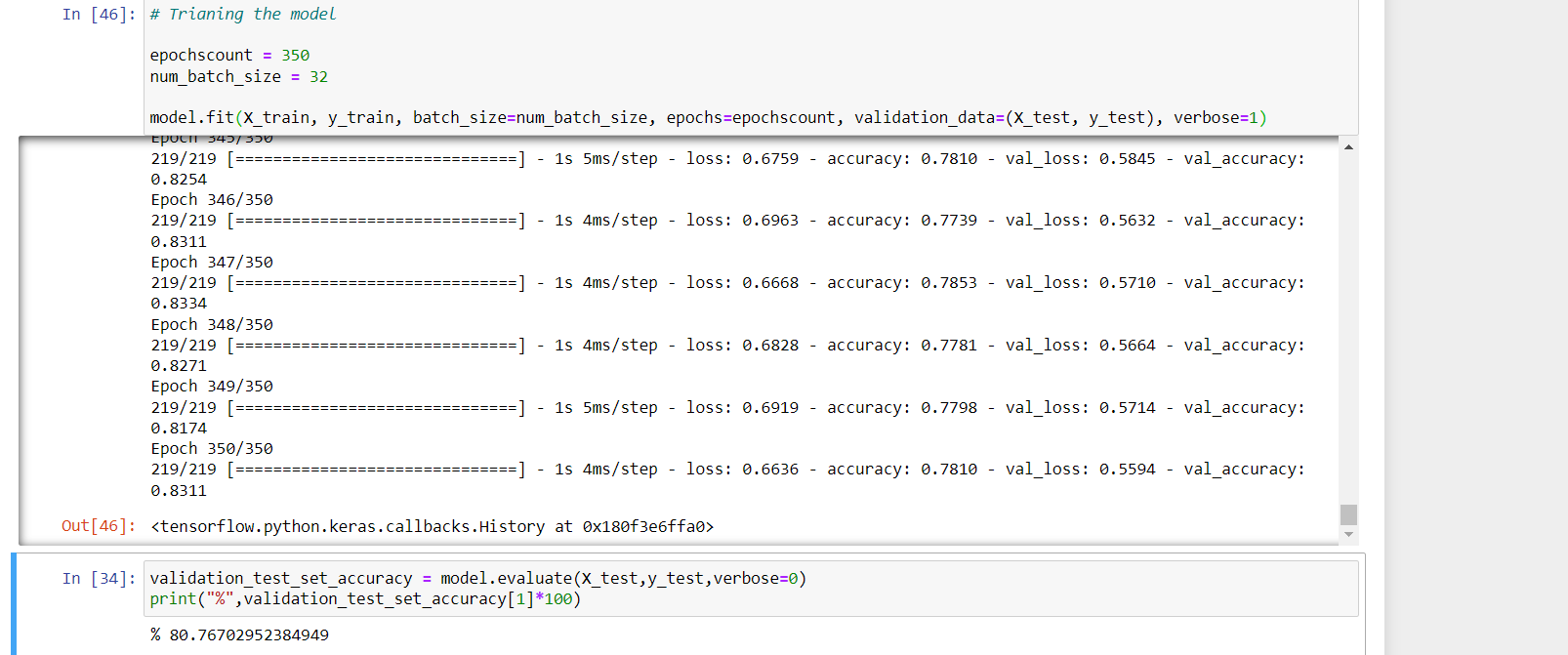
**Data Set**

The image below shows how our 8732 audio files are kept in the file and its format.

**Project Accomplishement**

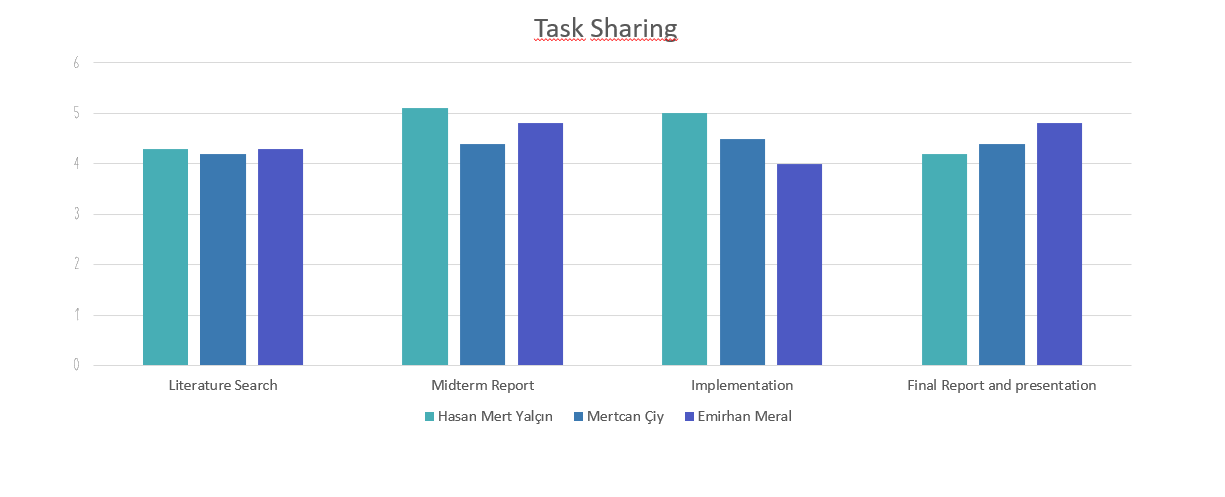
In the project, we implemented each of the subtasks given above step by step. We explained what we learned at the end of the project in the Summary section.

In the project, we extracted the mel frequency cepstral coefficients of each sound file. Based on this attribute, we tried to guess the audio files. At the end of the project, we achieved validation accuracy of 81%.



1. **Task Sharing**

In the image below, the distribution of tasks done while developing the project is given:



**TasksAccomplished:**

Researches and studies carried out since the date of the project are given below as sub-tasks.

· **Literature Research**

We spent a lot of time reviewing articles and resources available on the internet to get ideas for our project or to learn how to approach it. Since sound signal classifcation is one of the difficult projects to be done in the field of machine learning, we had no trouble finding more than one source. While reading these articles, we observed what steps to follow, what resources we should use and what difficulties we would encounter while doing the project. One of the important points we made while reading the articles was that the audio files should be in a format that can be used in deep learning algorithms. In other words, we should not expect the algorithm to recognize an audio file directly. We need to convert the audio file to a suitable format. In the articles, we saw that another approach is to convert each of the audio files to the frequency domain using the MFC algorithm and to calculate the MFC attributes of these audio files, and we decided to prepare the project in this way.

· **Existing Project Research**

We searched the internet for similar projects to the one we were going to do. As a result, we found similar projects and examined how they approached the problem according to the articles we researched. In this way, we observed how to approach our project and which Python libraries we should use. In addition, we found open source data sets in the projects that we were examine and we decided to use these source sound files in our project.

· **Bringing The Dataset To The Appropriate Format**

Preparing audio files for use in deep learning algorithms and preparing their formats correctly.

· **Training and Testing**

Separating the data set containing the audio files into train and test validation and optimizing the algorithm according to these results.

· **Architectural Building**

Building CNN architecture using Tensorflow and Keras libraries

· **Accuracy Calculation**

We will train our model with audio signals and calculate validation accuracy by applying the validation test with audio signals.

**Task Dates**

Week 1 [Oct-Nov 31-04]: Preparing the infrastructure of the

model to be formed and conducting research.

Week 2 [Nov 07-11]: Preparing the infrastructure of the model to

be formed and conducting research.

Week 3 [Nov 14-18]: Preparing the infrastructure of the model to

be formed and conducting research.

Week 4 [Nov 21-25]: Midterms Midterms Report.

Week 5 [Nov 28-02]: Data Analysis.

Week 6 [Dec 05-09]: Implementing algorithms and techniques.

Week 7 [Dec 12-16]: Implementing algorithms and techniques.

Week 8 [Dec 19-23]: Implementing CNN Algorithm.

Week 9 [Nov 26-30]: Testing.

Week 10 [Nov 05-11]: Testing.

Week 11 [Dec12-18]: Begin preparations for the final report and presentation.

1. **Summary**

* The first thing we learned while doing this project was how audio files are analyzed in deep learning projects.
* We have obtained many articles and information on this subject on the Internet. We learned that audio files are not used directly in deep learning algorithms, and we have to extract the MFC attributes of each audio file.
* In this project we learned to process audio signals. We learned to do this by detecting the different frequencies in the audio signals and using many mathematical algorithms and using the Python library (Librosa).

1. **References**

* [1]“UrbanSound8K,” *Urban Sound Datasets*. [Online]. Available: <https://urbansounddataset.weebly.com/urbansound8k.html>.
* [2] K. Doshi, “Audio deep learning made simple: Sound classification, step-by-step,” *Medium*, 21-May-2021. [Online]. Available: https://towardsdatascience.com/audio-deep-learning-made-simple-sound-classification-step-by-step-cebc936bbe5. [Accessed: 30-Nov-2022]
* [3] N. Seth, “Introduction to audio classification -emergency vs nonemergency vehicle,” *Analytics Vidhya*, 12-Jul-2021. [Online]. Available:https://www.analyticsvidhya.com/blog/2021/07/introduction-to-audio-classification-emergency-vs-nonemergency-vehicle/. [Accessed: 30-Nov-2022]
* [4] H. Subramanian, *AUDIO SIGNAL CLASSIFICATION*, Nov. 2004 [Online]. Available: https://www.ee.iitb.ac.in/~esgroup/es\_mtech04\_sem/es\_sem04\_paper\_04307909.pdf. [Accessed: 30-Nov-2022]
* [5] W. Mu, B. Yin, X. Huang, J. Xu, and Z. Du, “Environmental sound classification using temporal-frequency attention based convolutional neural network,” *Scientific Reports*, vol. 11, no. 1, 2021 [Online]. Available: https://www.nature.com/articles/s41598-021-01045-4. [Accessed: 30-Nov-2022]
* [6] S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” *2017 International Conference on Engineering and Technology (ICET)*, 2017 [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8308186. [Accessed: 30-Nov-2022]