DATS 6203 –Speech Command Recognition

Team Members:

1. Adina Dingankar
2. Rehapriadarsini Manikandasamy

Project Description:

Keyword spotting (KWS) technology is a potential technique to provide fully hands-free interface, and this is especially convenient for mobile devices compared to typing by hands. And it is also the desired technique for situations like driving or some emergency cases. Since speech command recognition system usually runs on smartphones or tablets, it therefore must be low-latency, and must have a very small memory footprint, and require only very small computation.

The dataset used in this project is sourced from Kaggle which has audio files (.wav) of duration of one second. The dataset has approximately 30K audio files classified based on 14 major labels. The dataset will be separated into train-test sets manually and processed to be trained on the deep learning models. The project will use deep learning models like 2D CNN+LSTM, CNN to find the best suitable algorithm for keyword spotting. The performance of the model will be evaluated based on the accuracy of prediction of keywords.

Project Schedule:

DEADLINES	TASKS
03/29/2021	Proposal
04/04/2021	Data Preprocessing and Analysis
04/12/2021	CNN Model
04/21/2021	2D CNN+LSTM Model
04/24/2021	Predictions on the Test Data
04/25/2021	Project Documentation
04/26/2021	Demo

Dataset source:

Source: Kaggle
Name: Speech Command: Application of speech recognition
Source Link: https://www.kaggle.com/datasets/venkatkumar001/speechcommands

Group GitHub:

https://github.com/Rehamanikandan/Final-Project-Group-2