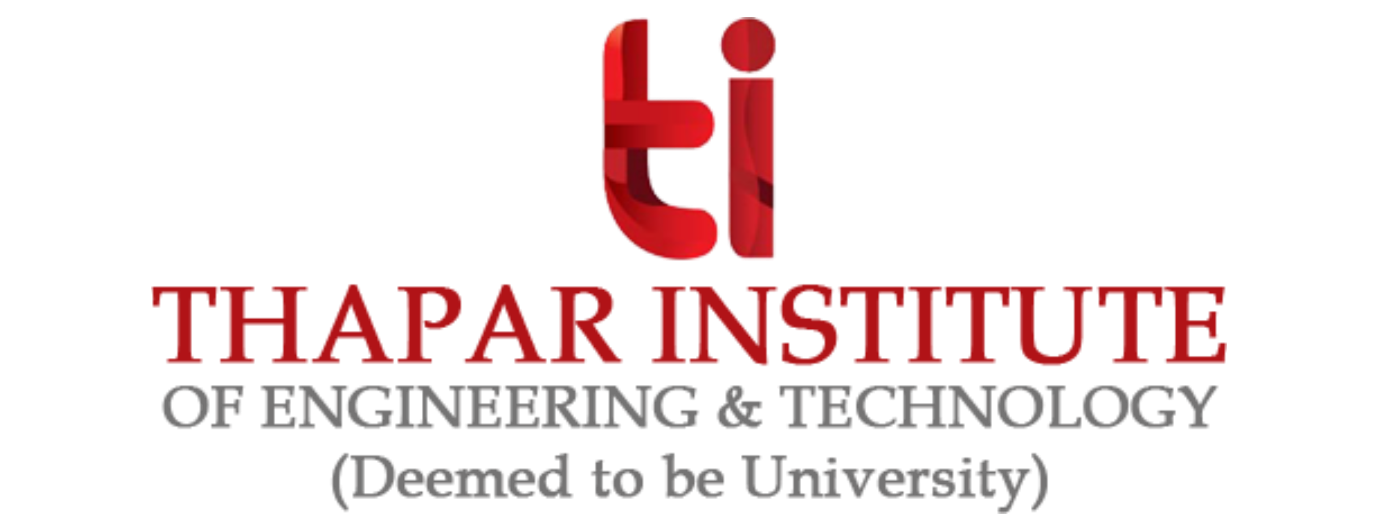
**RECOGNISE MY VOICE COMMANDS**

**SPEECH PROCESSING AND SYNTHESIS**

****

**Name: Atin Arora**

**Roll Number: 102103548**

**Submitted to:**

**Dr. Raghav B. Venkataramaiyer**

**THAPAR INSTITUTE OF ENGINEERING AND TECHNOLOGY,**

**(A DEEMED TO BE UNIVERSITY),**

**PATIALA, PUNJAB**

**INDIA**

**SUMMARY**

The paper introduces the "Speech Commands Dataset," a collection of spoken utterances for training keyword detection systems. Since effective models are intended for low-resource devices, they should be able to understand instructions such as "Yes," "No," and "Stop." The dataset enhances the reliability and comparability of the model. Baseline models with an accuracy of 88.2% are included. More than 100,000 words from 2,618 speakers make up this collection.

I've changed this project so that it can identify my voice using 900 audio samples from my own dataset. It reaches a 91% maximum frequency.

**DATASET SUMMARY**

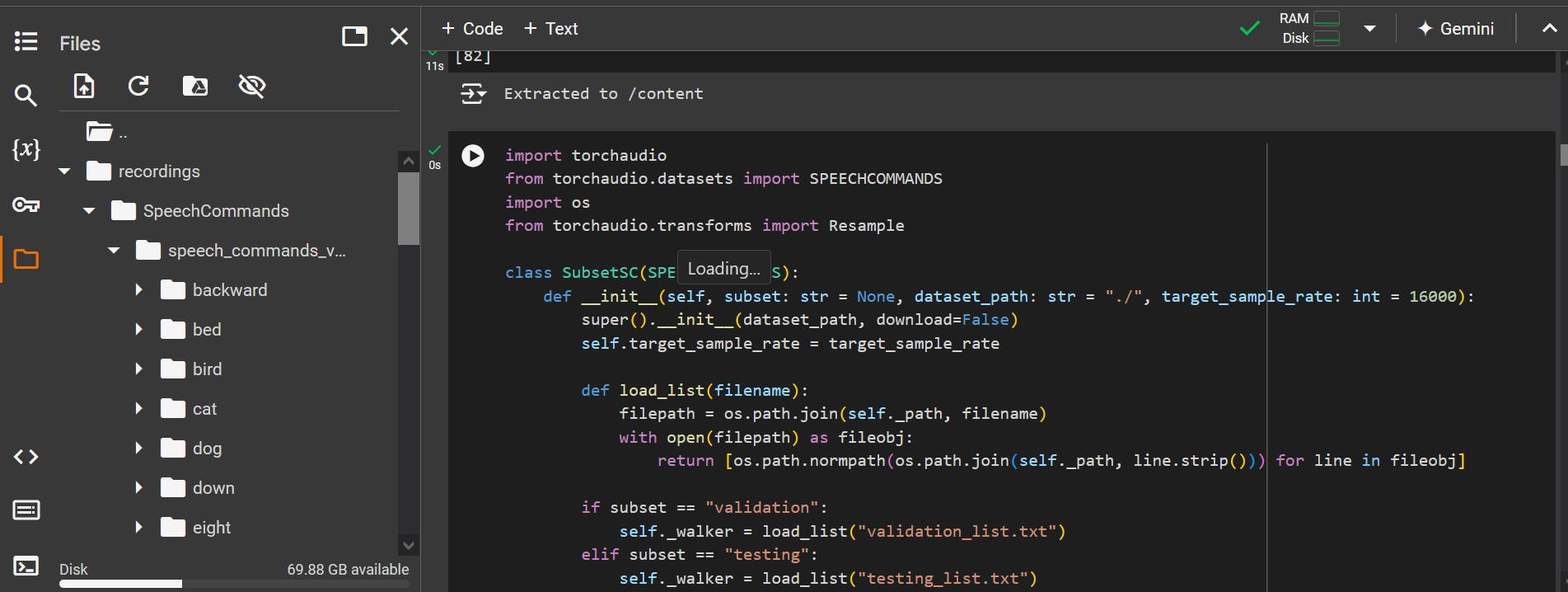
The study paper's Speech Commands Dataset includes over 105,829 audio recordings of 35 distinct words. Every audio file in the WAV format has a single spoken word that is sampled at 16 kHz. Since 2,618 speakers provided the dataset, a wide range of accents and pronunciations were guaranteed. With words like "Yes," "No," "Up," "Down," "Left," "Right," "On," "Off," "Stop," "Go," and other words like numerals (zero to nine), it primarily focuses on small-vocabulary keyword detecting tasks.

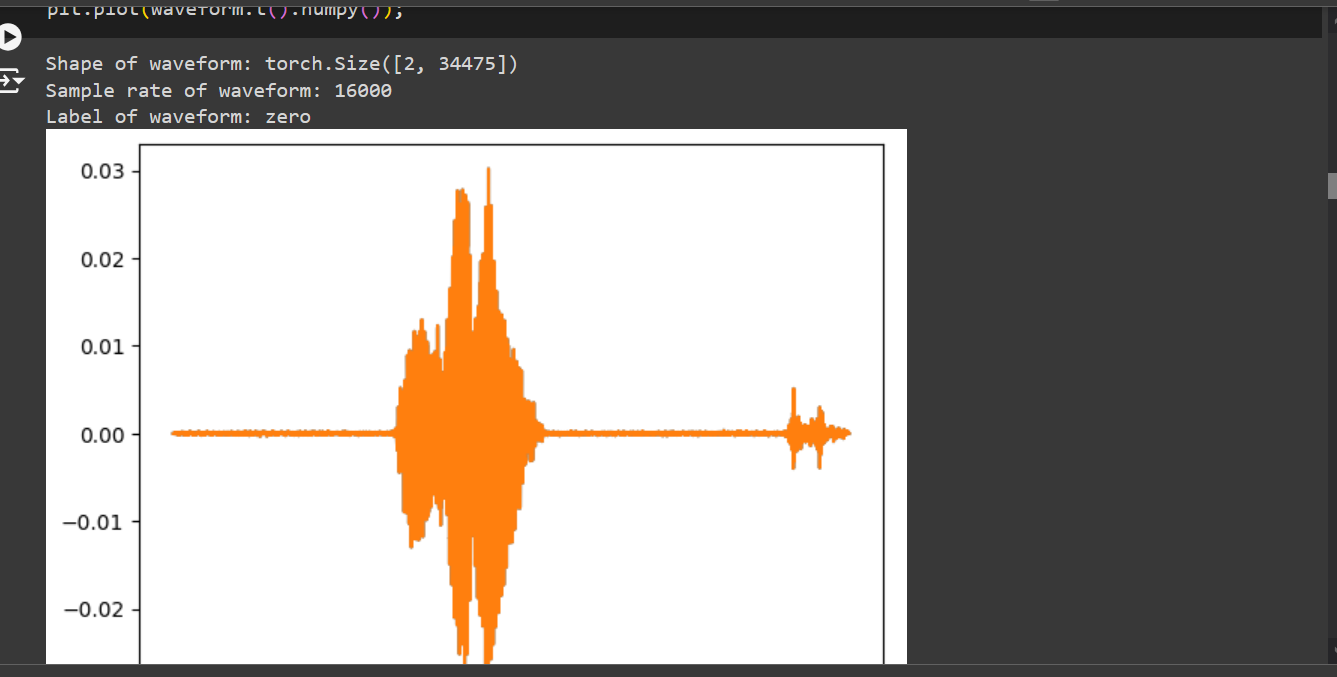
When uncompressed, the complete dataset is about 3.8 GB, or 2.7 GB when saved as a tar archive that has been compressed using gzip. It also comes with files including background noise to mimic actual situations and increase the robustness of the model.

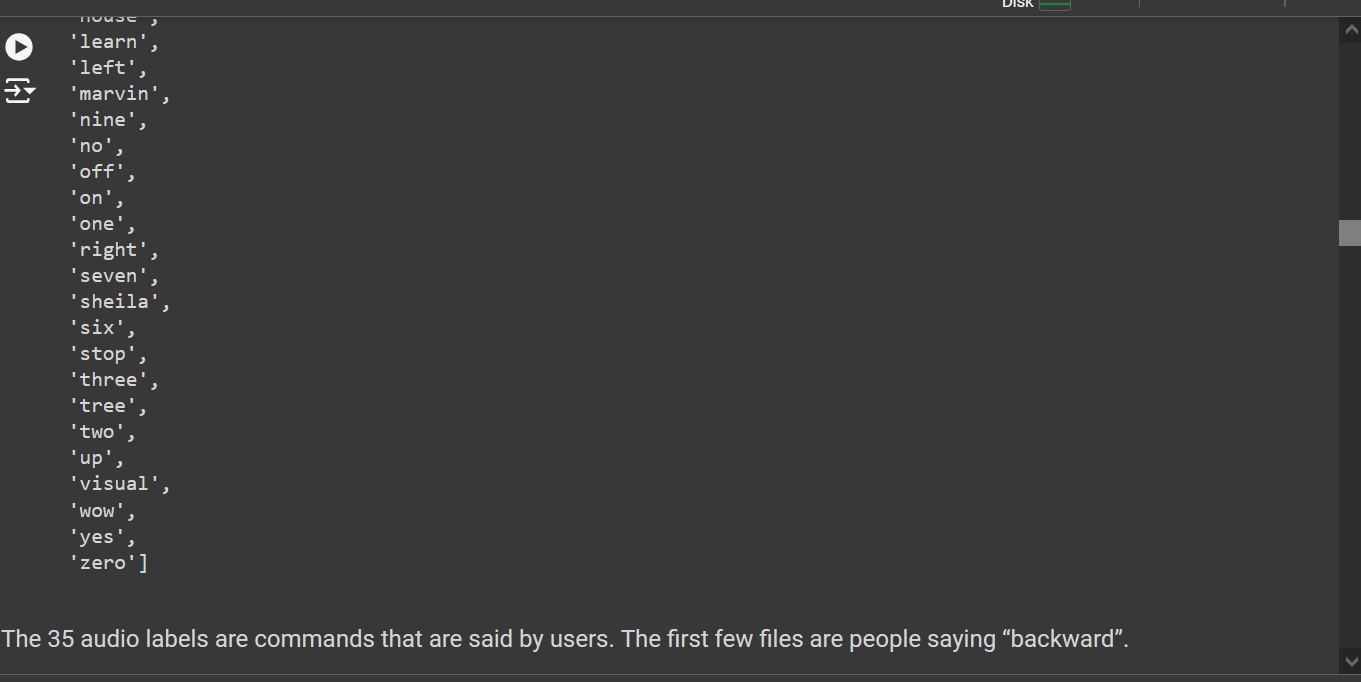
**PERFORMANCE MEASURES**

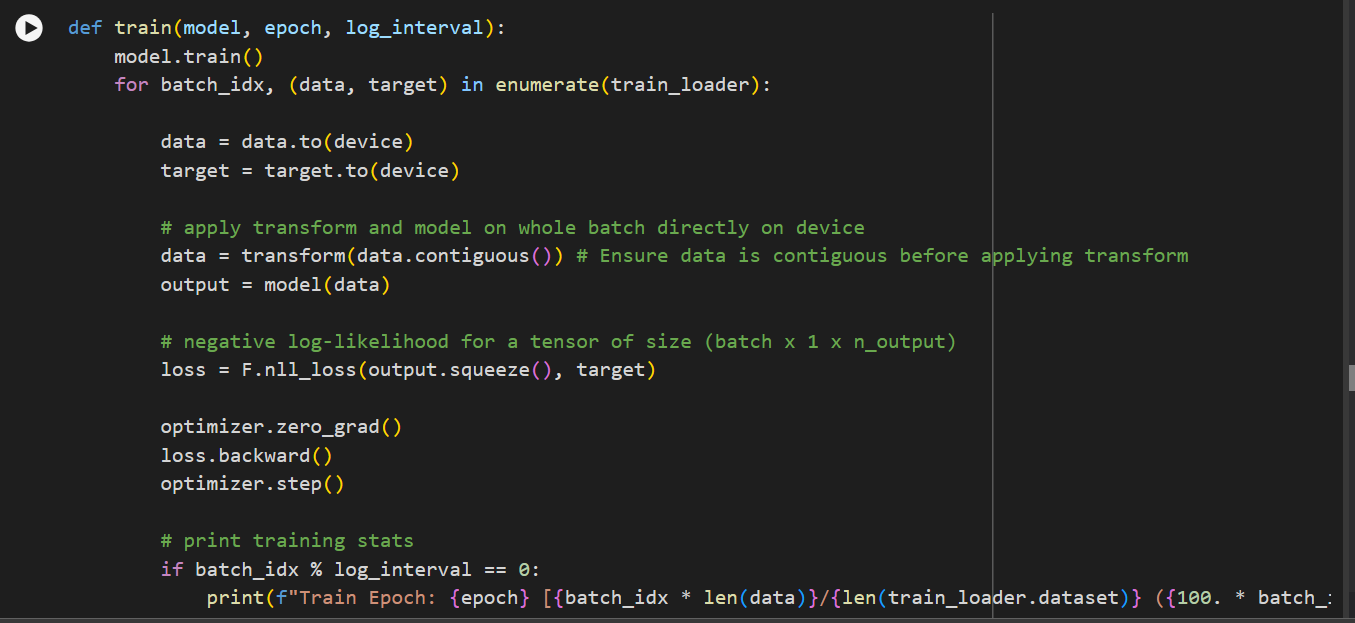
The accuracy of the model came out to be maximum of 91% after 20 epochs.

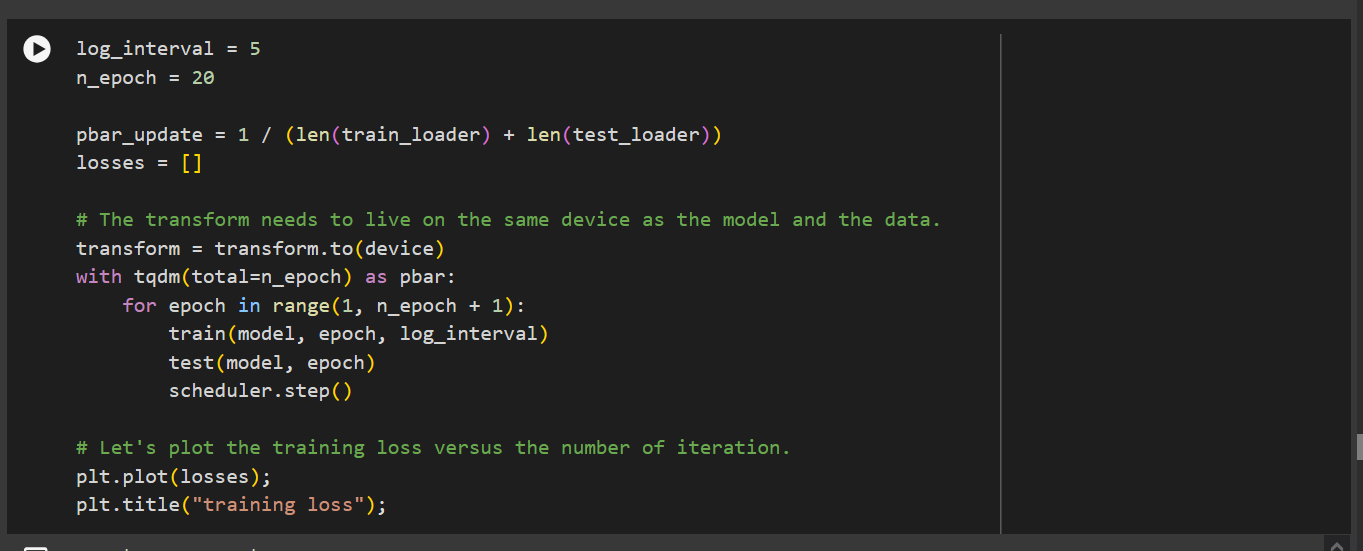
**CODE SNAPSHOTS**

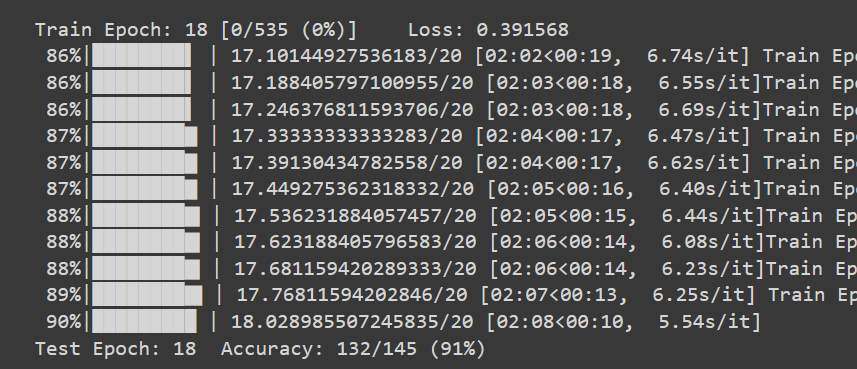
****

****

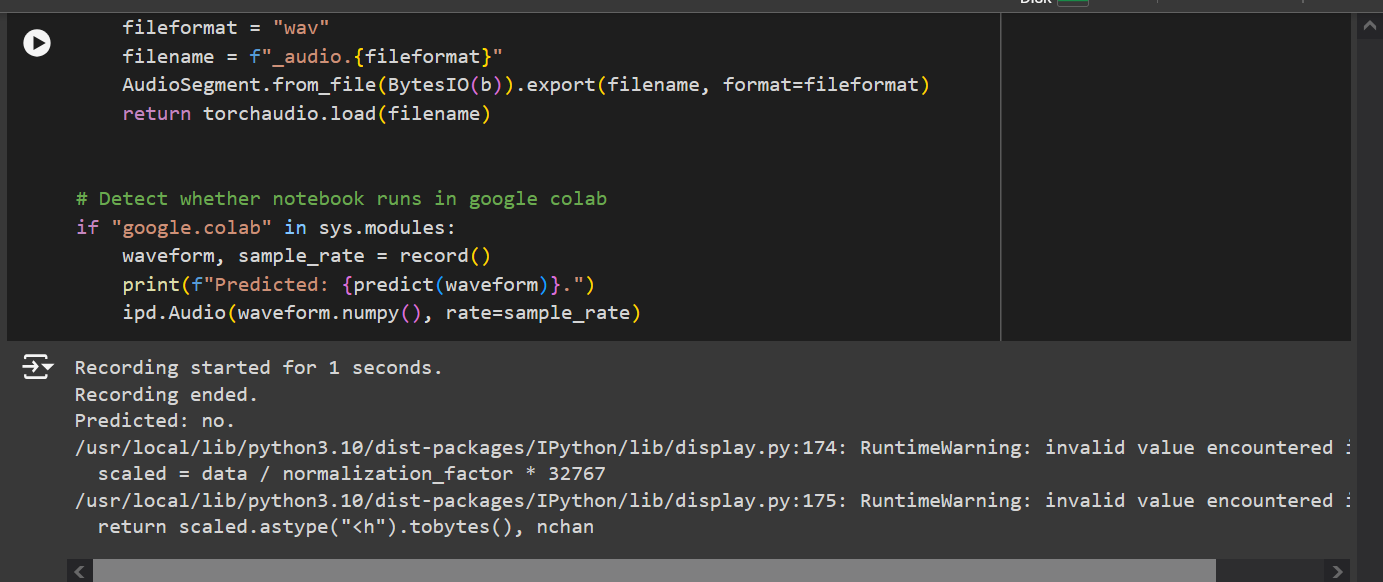
****

****

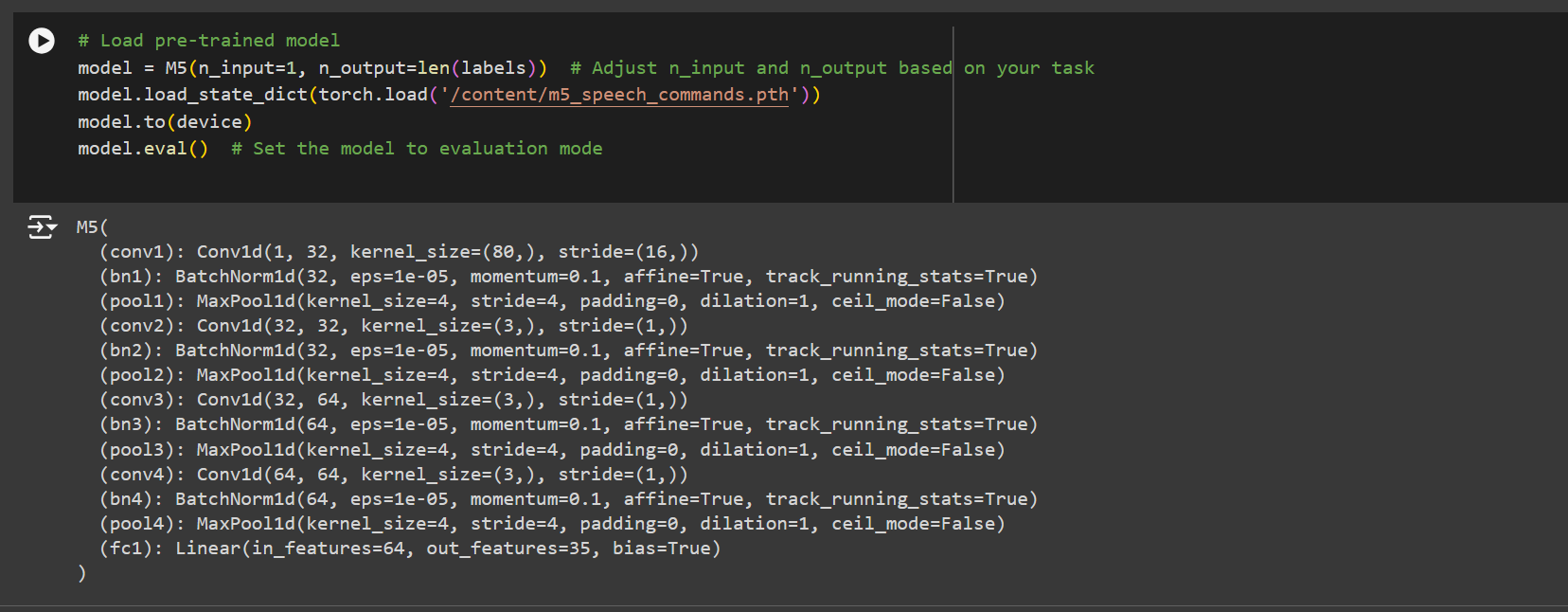
****

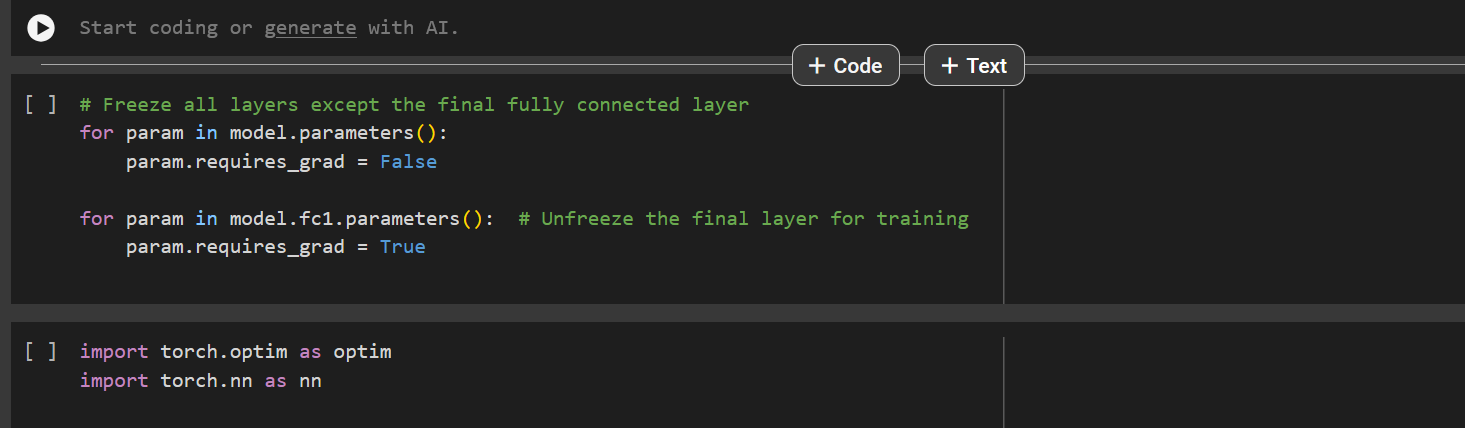
****

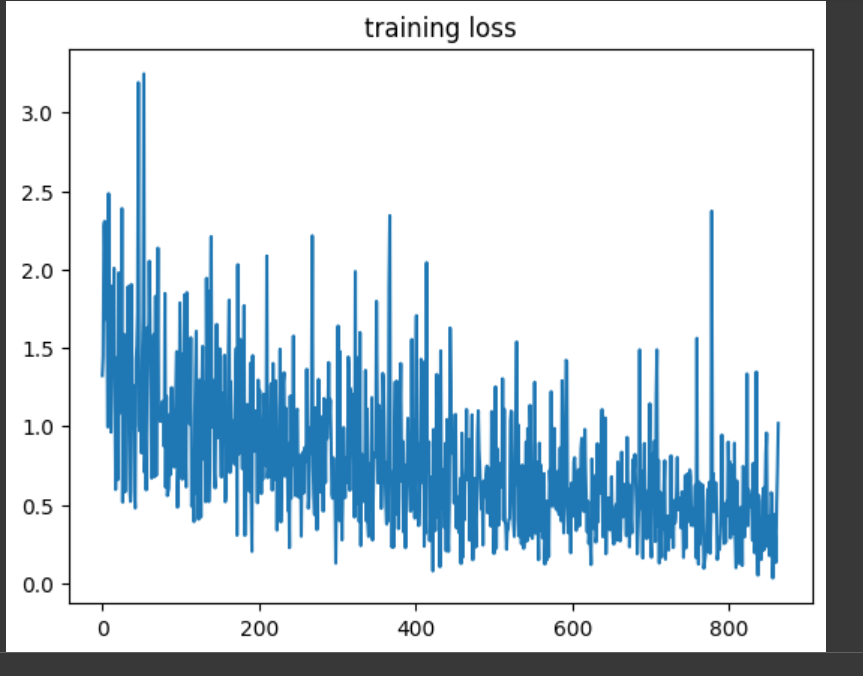
****

****

**FINE TUNING**

****

****

****