# Voice Command Feature

## Basics

Introducing a voice-based command system to stop the robot while disinfecting. We will be using Arduino Nano, with a built-in microphone to do real time inference on voice data.

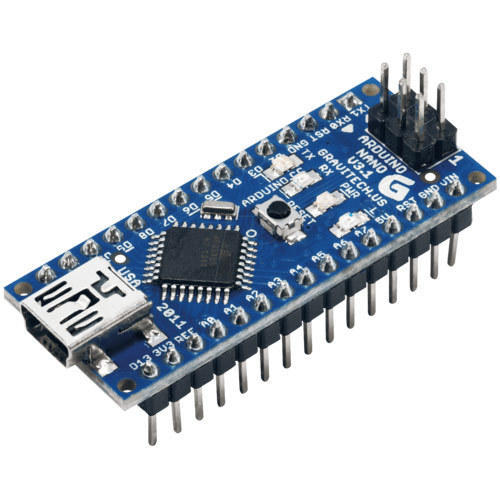


Figure 1 Arduino Nano

## Intro

* We have trained a neural network model that recognizes wake words such as “STOP” for a voice enabled emergency stop while the robot is disinfecting.
* The network is small 20 KB, so that it can be ported to light hardware’s such as Arduino Nano
* We have made use of the TensorFlow Lite suite to convert TF models to TFLite models that reduce model size significantly.
* With real time inferencing, we will integrate this functionality in our robot



Figure 2- Aiming to stop the robot using voice commands such as "stop"

## Details

* We trained the model on GPU with Google Speech Commands dataset with the required wake words.
* Preprocessing of voice must be done first. It converts voice sample (1 sec) to MFCC (Mel Frequency Cepstral Coefficient) like image. This is the feature extraction preprocessing.
* Sample Window= 0.5sec
* Stride= 0.5sec
* For a 0.5 s voice sample it takes 0.25s for it to be preprocessed on Raspberry Pi. (to be calculated on Arduino Nano)
* Next, they are converted to a header file (.h file) using the help of a TFLite converter.
* And finally ported to the microprocessor for inference

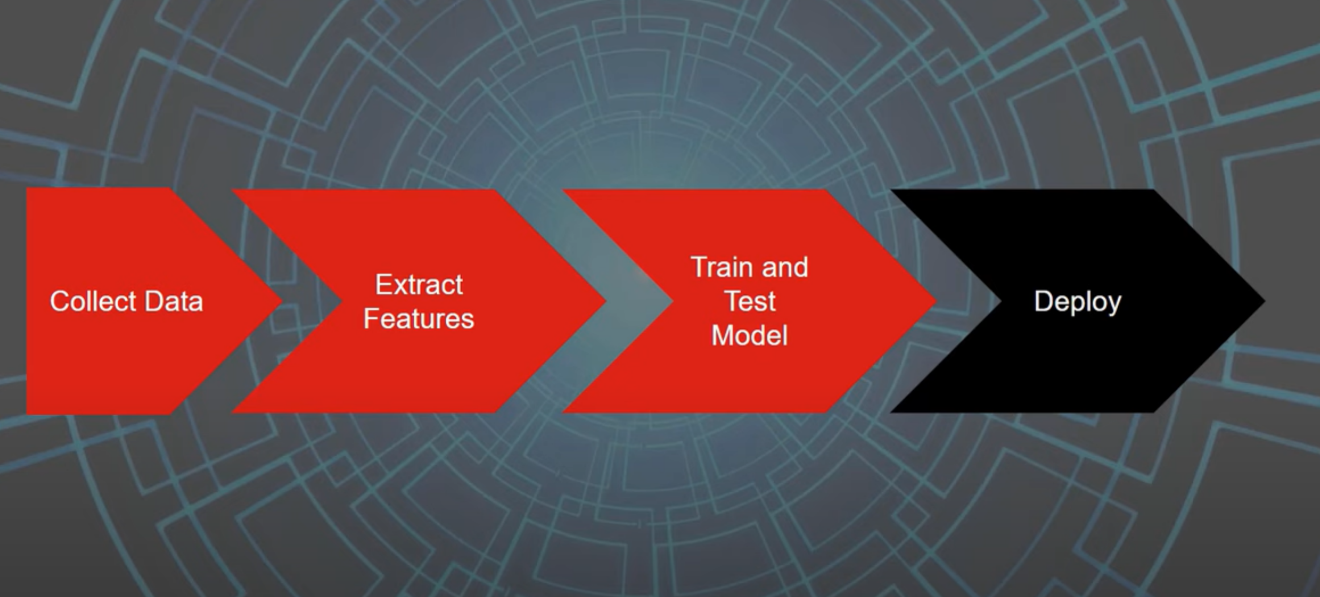


Figure 3-Schematic of the overall process

## The First Attempt

* **DATASET**: Google Speech Commands Dataset v0.02
* **Trained Classes**: ***STOP, NO, OFF,*** SILENCE, UNKNOWN WORD
* **Training Performance**:
  + **Overall Accuracy on Test set**: 87.5%

### Confusion Matrix

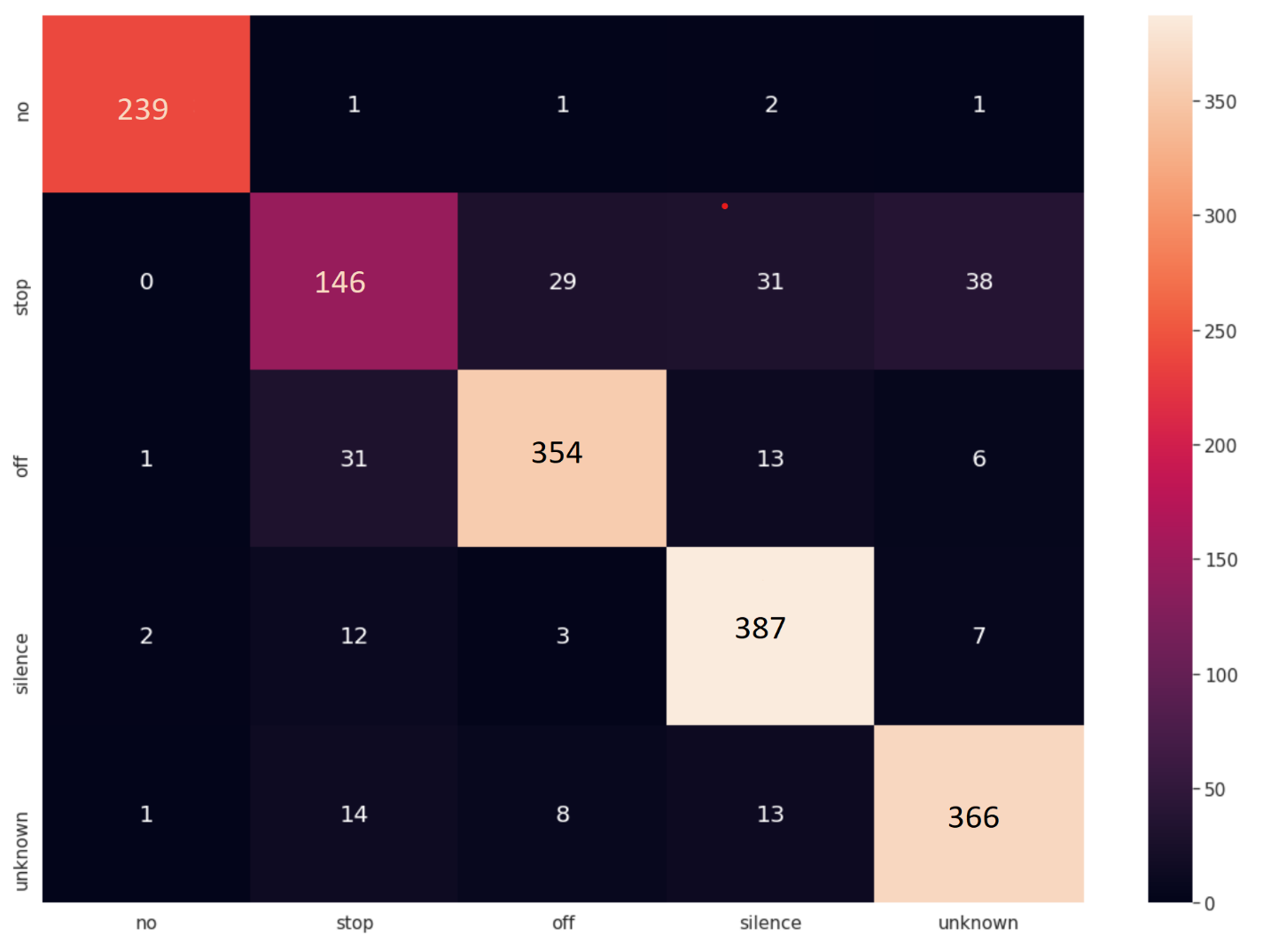


Figure 4- Confusion matrix on test set

## The Second Attempt

* **DATASET**: Google Speech Commands Dataset v0.02
* **Trained Classes**: ***STOP,*** SILENCE, UNKNOWN WORD
* **Training Performance**:
  + **Overall Accuracy on Test set**: 93.9%
  + **Precision (For STOP**): 97.5%
  + **Recall (For STOP)**: 96.5%

### Confusion Matrix



Figure 5-Confusion matrix on test set

## Conclusion

The results post training looks very promising. With such great accuracy and precision, this feature seems to be an appropriate addition to our robot. Now the only thing left is to test the TFLite model on Arduino NANO for real life scenarios inference. We might even fine tune the model with new training data as we please.