

SUCCESS STORY

# AUDIOFARM TOOL

An Audio ML-based Model Generator for  
Customized Voice Commands





## INTRODUCTION

A Tier-1 semiconductor manufacturer approached Ignitarium to create an AI-based audio application for microcontroller-based boards. The requirement was to bring Machine Learning capabilities for audio keyword and command detection. To address this, we created an Audio Keyword Detection Platform using Machine Learning which would provide customized ML models for deployment.

Spoken commands are gaining momentum with Consumer devices / appliances and ML-based solutions for speech recognition and keyword detection have been proven to work well. Most consumer appliance vendors do not have in-house Machine Learning and Data science teams, especially in the computer audio domain. This solution was aimed at alleviating this and allows the client team to focus on their key strengths.

### Solution in Focus

## VOICE COMMAND ENGINE TOOL



#### Industry

Semiconductors used in Consumer Appliances



#### Scope

Producing an ML model specific to the deployment target ranging from Cloud to pruned/quantized model for Embedded devices

## THE CHALLENGE

The objective of the project is to provide customized ML models for deployment. The customer uploads audio data specific to their use case and lets the platform model learn the keywords. The service will produce a model specific to the deployment target ranging from Cloud to pruned / quantized model for Embedded devices.



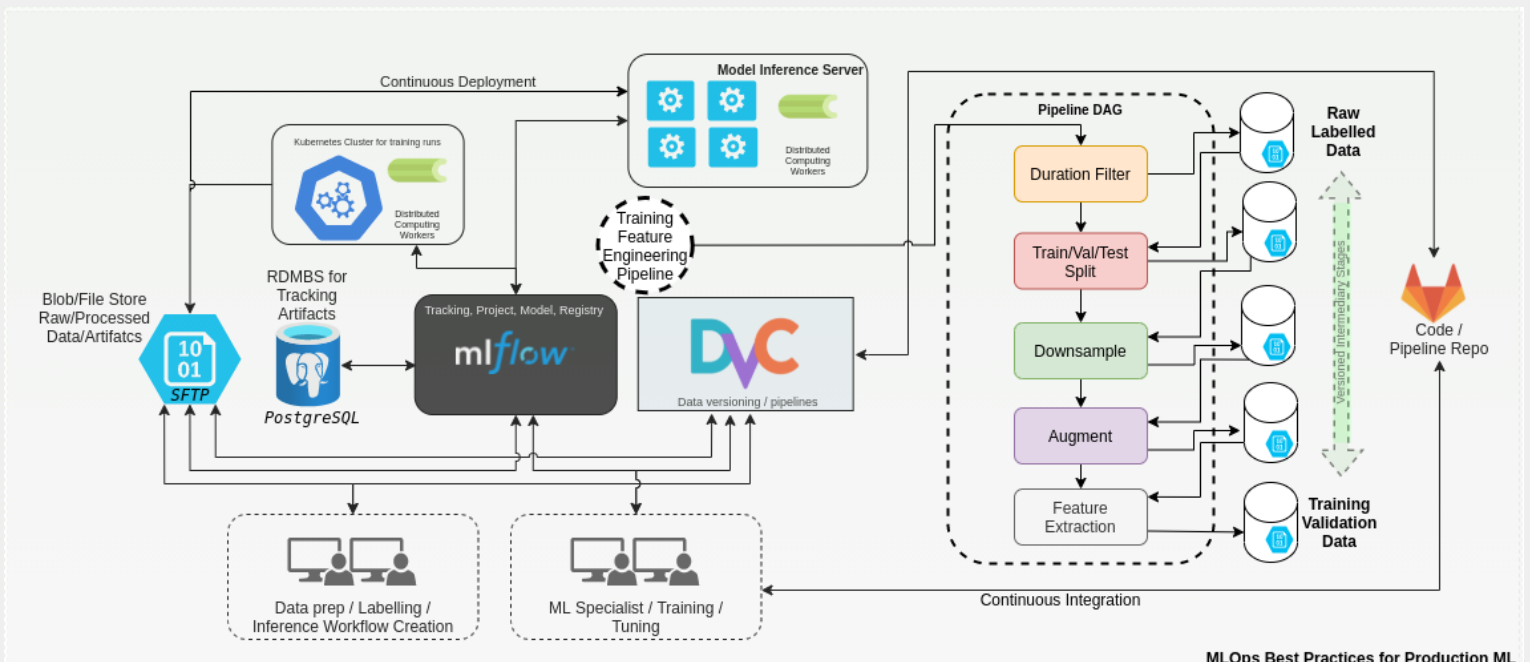


While designing the solution, Ignitarium was required to address the following challenges:

- Constraints with respect to deployment on low-memory devices
- Keywords in all Indian accents to be correctly predicted
- All words except the original keywords should be predicted in an 'unknown' category
- Predict voice commands correctly in a noisy environment
- Voice command prediction is activated only when any voice activity is detected

## IGNITARIUM'S APPROACH

Ignitarium's audio engineering team conceptualized a solution which would allow the customer / tenant to upload audio data using the web front end and label the audio clip for the keyword it represents. The tenant can also crowdsource the data collection process. The platform can automatically filter these data clips for duration, clean and split the training and validation data, augment the data for real world scenarios like noise, ambient sounds, hum and recurrent beeps. The platform also supports curated bespoke services for data collection and labelling based on ethnicity, gender, age groups, etc. The tenant can also have customized pipelines for specialized cases like in-car, kitchen, etc.



Our approach entailed the following tasks:

- Choosing different keyword set and data collection
- Identifying the best data augmentation and data cleaning technique
- Testing custom neural network developed for voice activity detection and voice command engine
- Implementing pre-processing and post-processing flow for the inputs generated for a custom neural network built for voice command engine. Task involved modifying the process flow retrieving intermediate outputs from the network so as to enable post-processing to happen on the host
- Porting the custom neural network and successfully deploying the voice activity detection and voice command engine on to customer hardware
- Evaluating neural network performance
- Creating a custom tool for data collection, cleaning, augmentation and training

MIC TESTING

help ?

PRESS TO RECORD

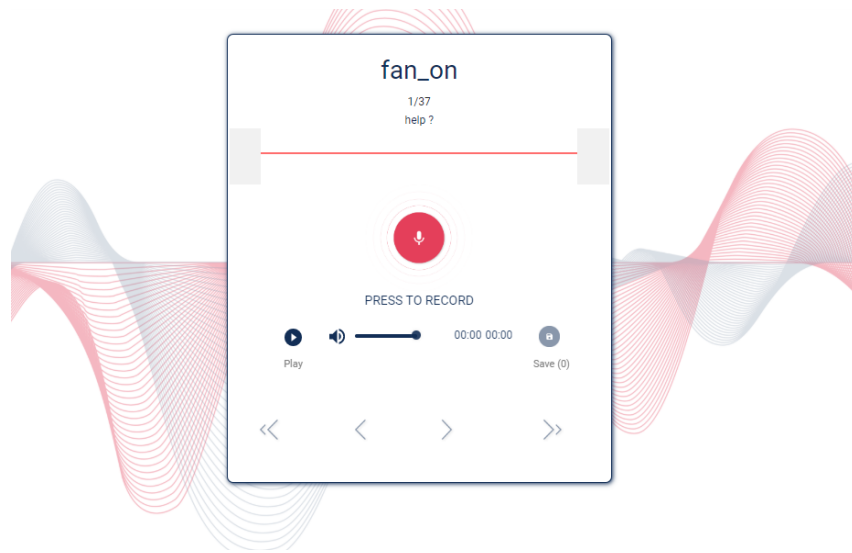
00:00 00:00

USER DETAILS ?

Age Gender Country

State Language Keyword List

Start Recording



## SOLUTION

- Scalable MLOps pipeline in the backend which can scale to many thousands of ML engineers.
- Easily handle / version several TBs of data
- Exposes experimentation results, and retraining results w.r.t accuracy and similar metrics to tenants
- The solution can share instances with multiple tenants with isolation of data at the DB or Bucket level
- Can be deployed as a service specific to a tenant for data privacy sensitive domains or as on-prem subject to requirements w.r.t hardware and software for IaaS or Virtualization platform

## BUSINESS IMPACT



Customized data collection and cleaning tool can be used for different audio ML applications



50% faster time to market



25% cost reduction

Drop us a line to get in touch with our experts.