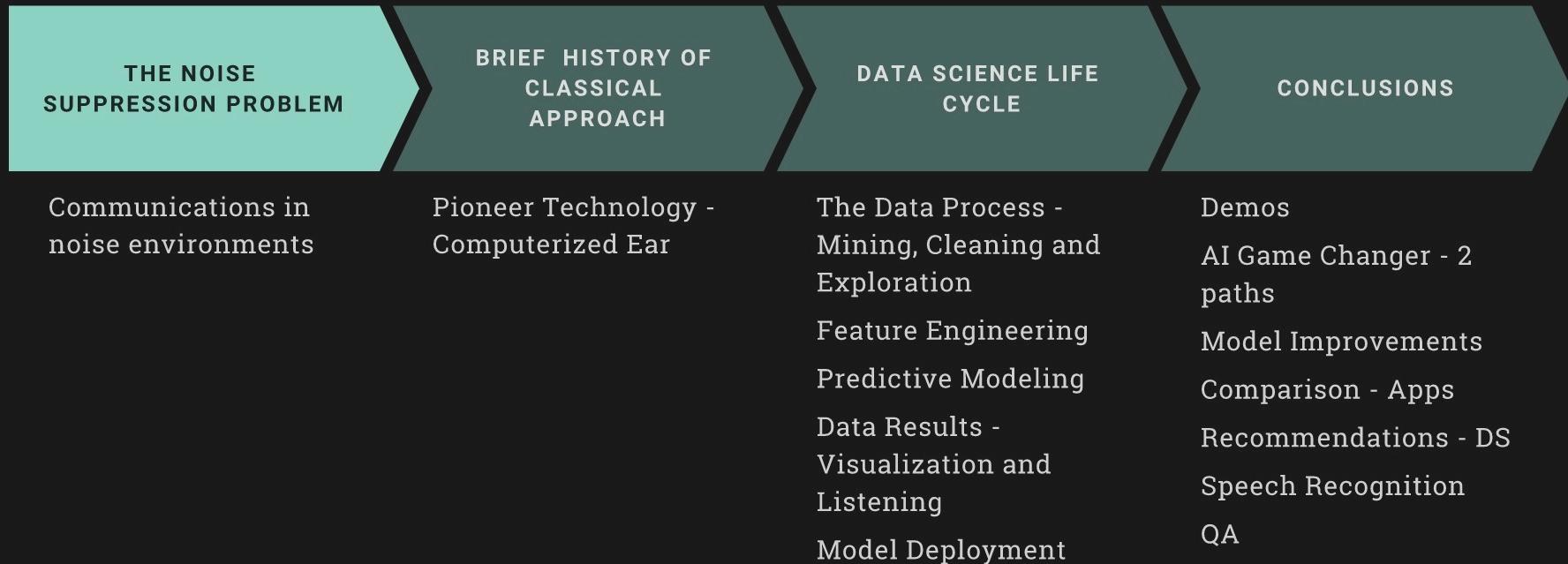


# Noise Suppression AI

Oscar Caraballo

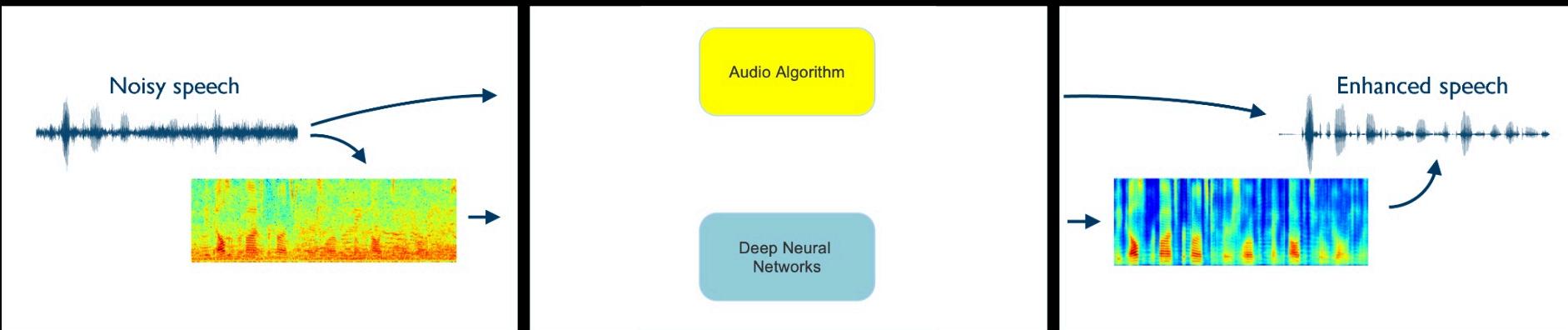
# Agenda

---



# Problem Statement

## Removing Noise From a Speech Signal



Classic Computational Problem -  
Solved by Algorithm Creation

A 10 years journey

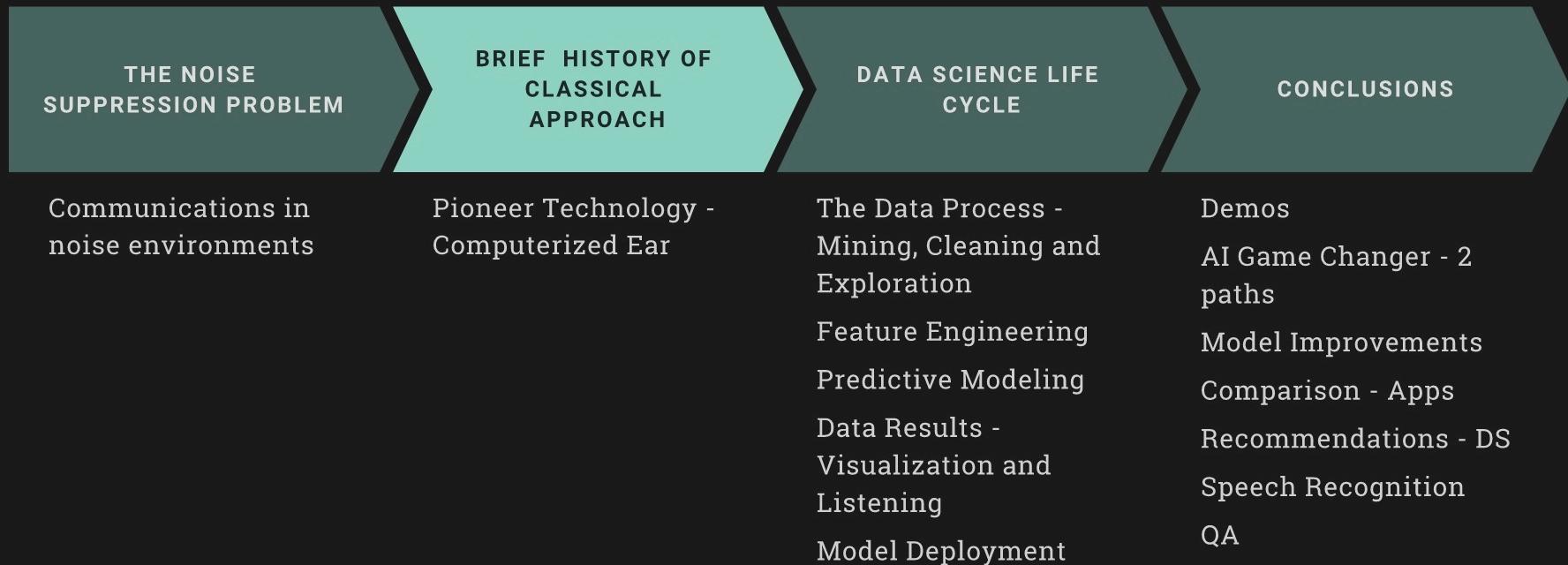


Data Science Model Problem -  
Solved by AI - Neural Networks

A 10 Days Journey

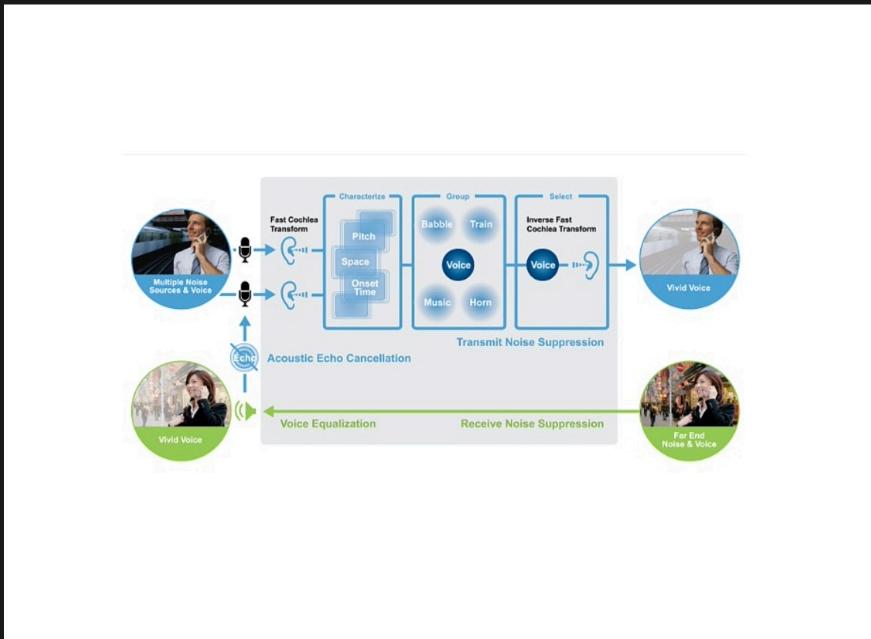
# Agenda

---



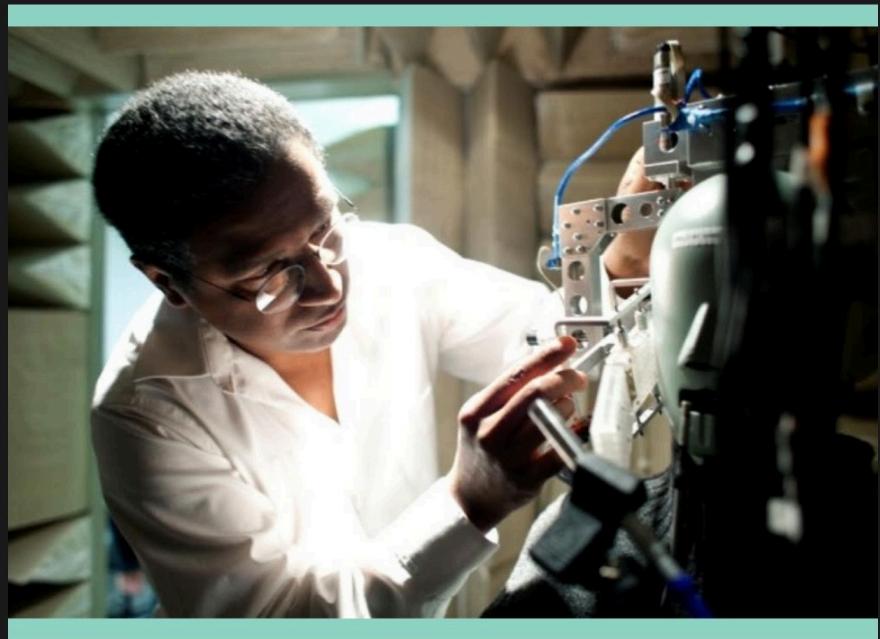
# Noise Suppression AI

## Classical Computation vs Data Computation



AUDIO ALGORITHM STRUCTURE

A 10 years journey



AUDIO ALGORITHM STRUCTURE

Audio Calibrations

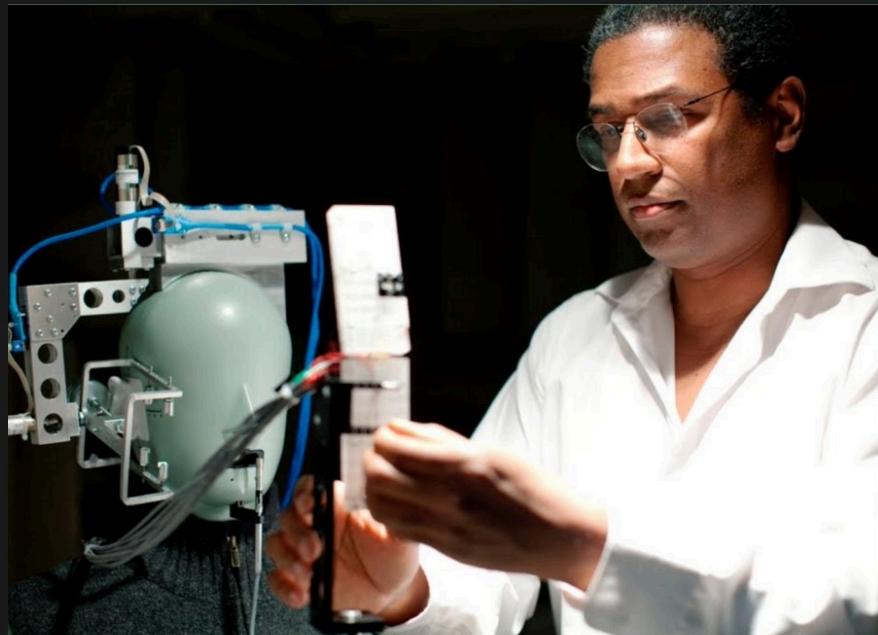
# Noise Suppression AI

## Classical Computation vs Data Computation

The screenshot shows the Audience earSmart website. At the top, there's a navigation bar with links for Demos, About Us, Partners, and Press. Below the navigation is the Audience logo and the earSmart logo with the tagline "The world's most intelligent voice processor™". A large image of a smiling woman talking on a phone is on the right. The main headline reads "The future of voice is hear.™". Below the headline are four menu items: What is earSmart, Our Technology, Our Processors, and Products with earSmart. A call-to-action button "Discover How" is visible. On the left, there's a section titled "Learn why earSmart™ is the world's most intelligent voice processor." featuring a video thumbnail with a man and a robot arm.

KNOWLES APPLE GOOGLE SAMSUNG

Telecommunication and Carriers Perspective

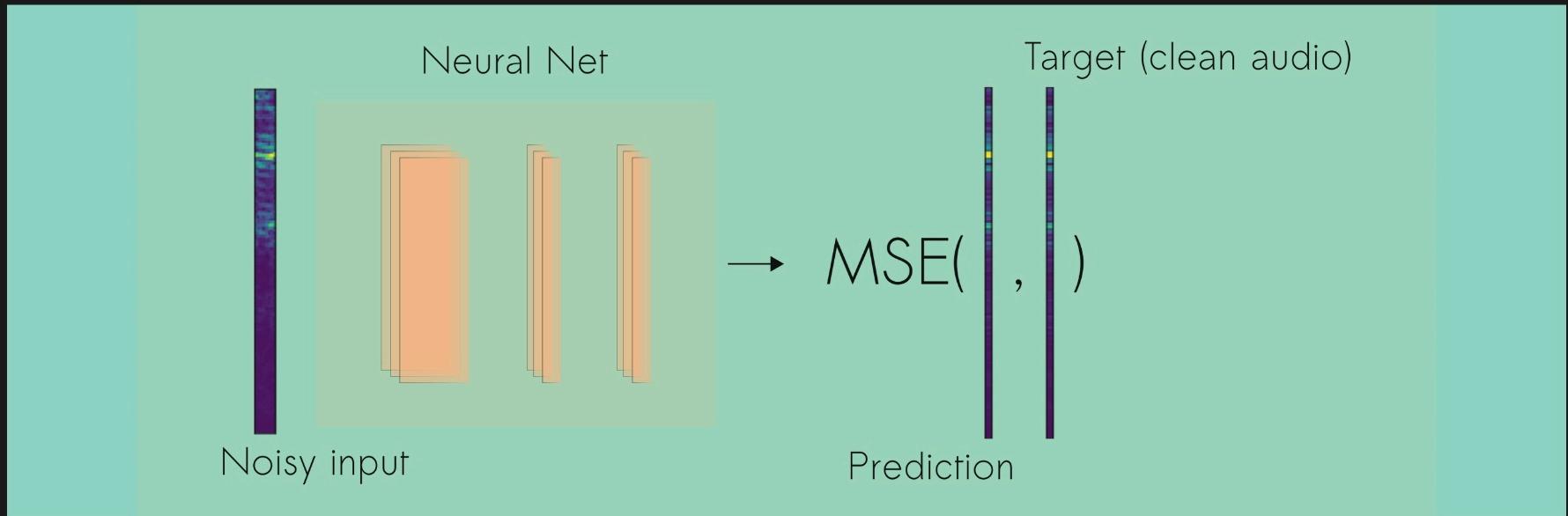


EXPERIMENTAL DATA COLLECTION

Generative Models

# Noise Suppression AI

## Classical Computation vs Data Computation



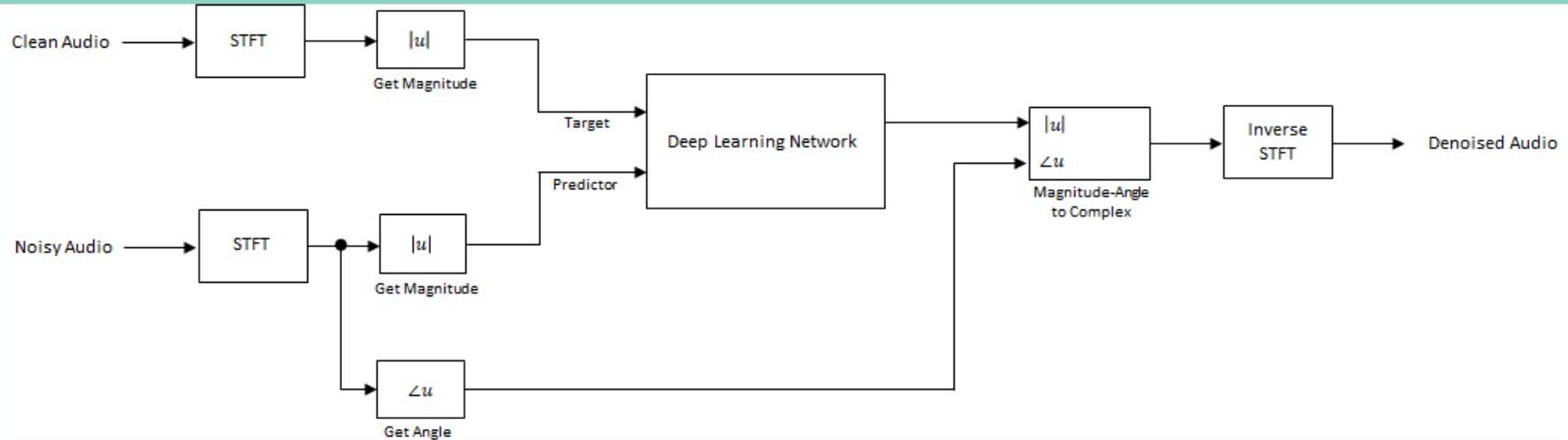
### DATA SCIENCE APPROACH

Build a statistical model that can extract the clean signal from a complex signal

Measured MSE, RMSE and LOSS

# Deep Learning System Overview

---

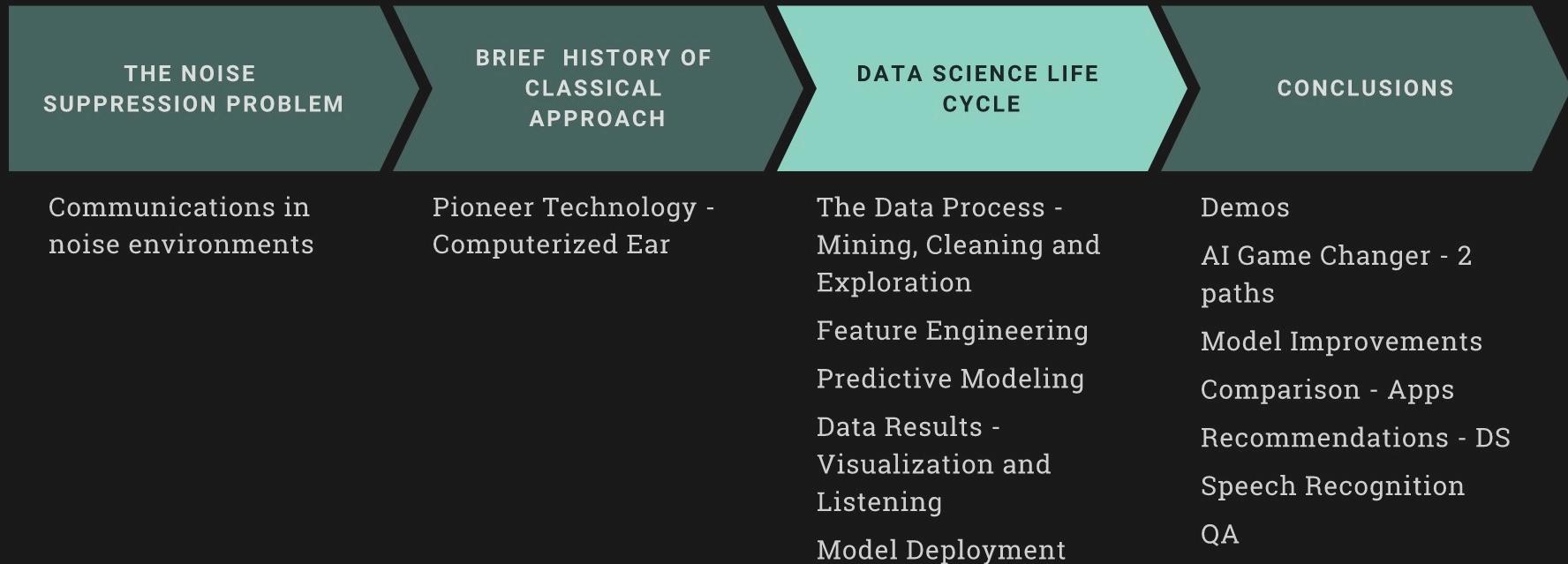


## SIGNAL FLOW INPUT TO CONVOLUTION NEURAL FEATURE ENGINEERING

Digital Signal Processing to Deep Learning Networks to Digital Signal Processing

# Agenda

---



# Data Processing

We're building an open source, multi-language dataset of voices that anyone can use to train speech-enabled applications.

We believe that large, publicly available voice datasets will foster innovation and healthy commercial competition in machine-learning based speech technology.

Common Voice's multi-language dataset is already the largest publicly available voice dataset of its kind, but it's not the only one.

Look to this page as a reference hub for other open source voice datasets and as Common Voice continues to grow, a home for our release updates.

**Version:** Common Voice Corpus 6.1  
**Language:** English  
**Date:** 2020-12-11  
**Size:** 56 GB  
**Version:** en\_2181h\_2020-12-11  
**Validated HR Total:** 1,686  
**Overall HR Total:** 2,181  
**Licence:** CC-0  
**Number of Voices:** 66,173  
**Audio Format:** MP3  
**Splits:**  
Gender:  
24% United States English,  
8% England English,...  
Age:  
24% 19 - 29, 14% 30 - 39,...  
47% Male, 15% Female  
[Enter Email to Download](#)

## DATA SOURCES

Speech Data Common Voice - Mozilla  
2000 Male and Female Speakers  
Corpus Creation - Speech Data Preparation

### Ontology > Channel, environment and background >

Noise

A sound that has no perceptible structure and that typically interferes with the perception of more interesting or important sounds.  
349 annotations in dataset

Background noise

Sounds present alongside a sound of interest but typically weaker in energy and generally not of interest.

Hubub, speech noise, speech babble

Loud, disordered, unintelligible speech noise from many sources.  
1,480 annotations in dataset

Cacophony

A loud, confusing, disagreeable sound, such as many voices shouting simultaneously.  
856 annotations in dataset

White noise

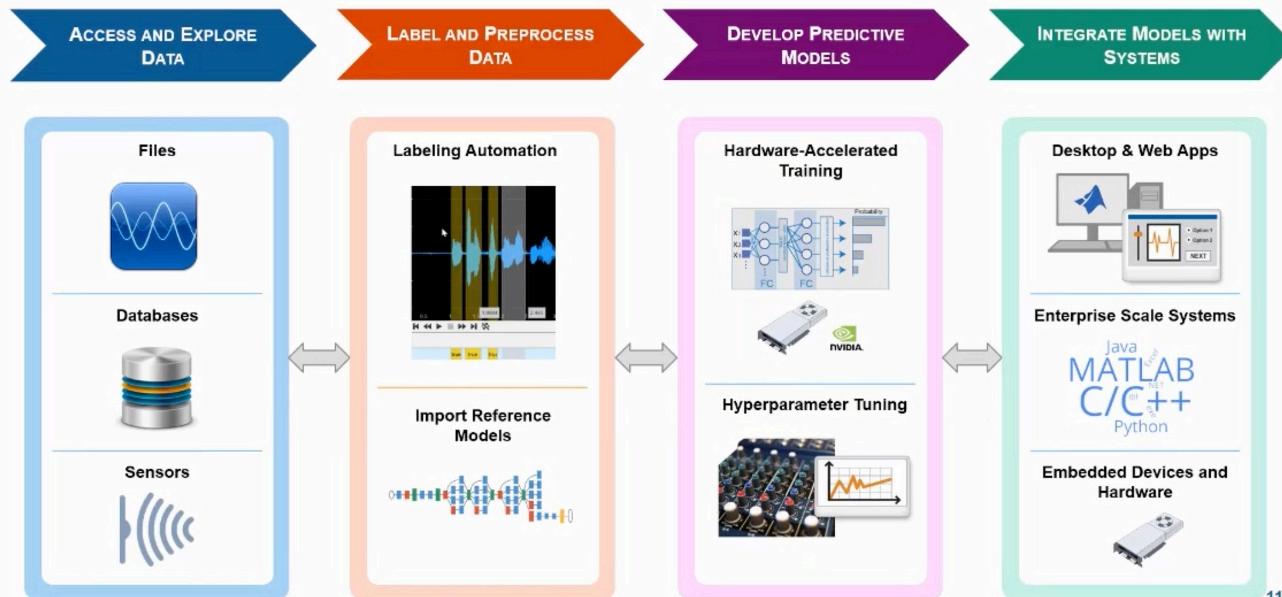
A random, unstructured sound in which the value at any moment provides no information about the value at any other moment. White noise has equal energy in all frequency bands.  
1,800 annotations in dataset

## NOISE DATA SOURCES

Ontology - Most Critical Noise Environments  
5 Noise Environments  
Corpus Creation - Noise Data Preparation

# Data Processing

## Deep Learning Workflow for Signals

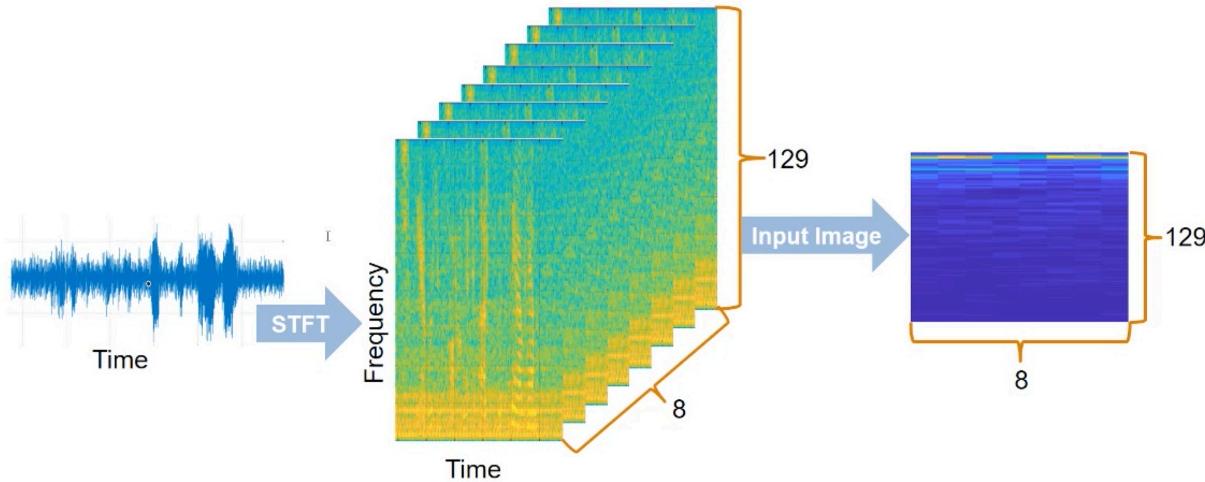


11

# Data Processing

---

## Image Input to the Convolutional Neural Network



### SIGNAL FLOW INPUT TO CONVOLUTION NEURAL NETWORK

Digital Signal Processing Intro

Time Domain to Frequency Domain - Time Series

Noise Suppression AI

## Data Science Life Cycle

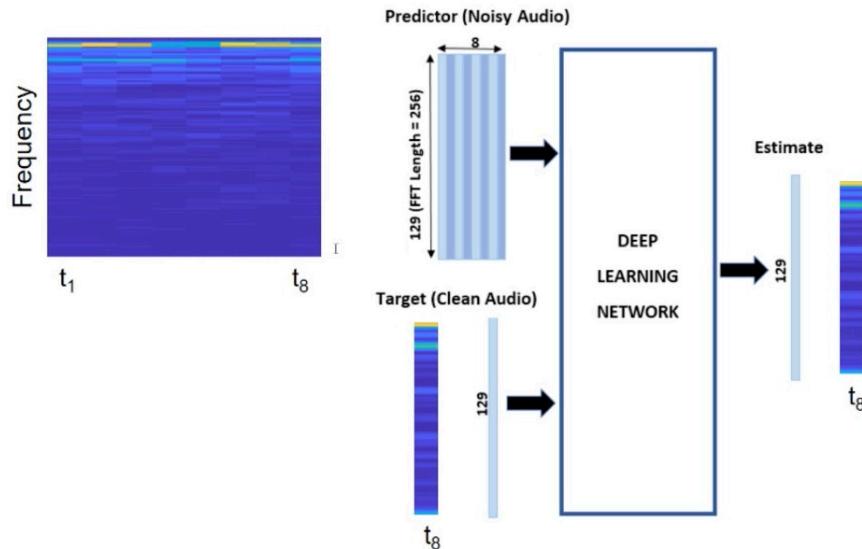
---

**FEATURE ENGINEERING**

Target and Predictors

# Feature Engineering

## Short Time Fourier Transform (STFT)



## SIGNAL FLOW INPUT TO CONVOLUTION NEURAL FEATURE ENGINEERING

Target and Predictors

Noise Suppression AI

# Data Science Life Cycle

---

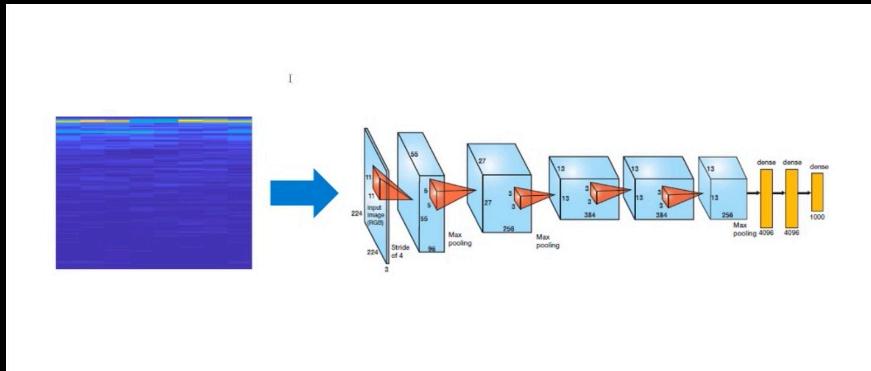
## PREDICTIVE MODELING

Predictive Modeling

## Neural Networks

# Predictive Modeling

### Neural Network Structure for Signal Processing



#### NEURAL NETWORK PROCESSING

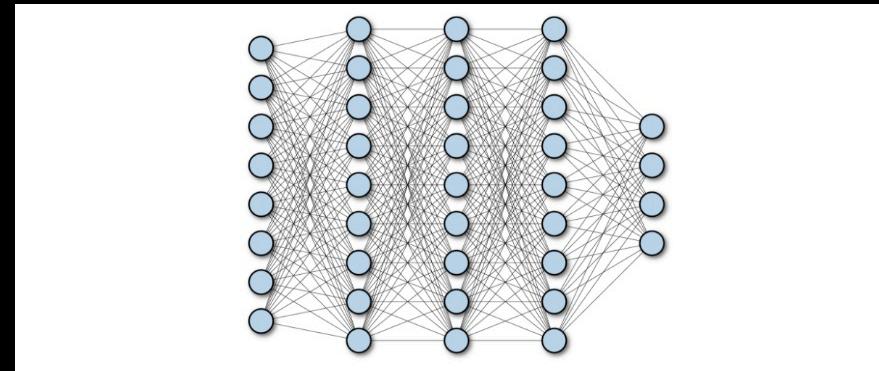
Input size 129-by-8

Number Number of Hidden Layers 16 with 1024 neurons + Relu

Type of Layers Input Normalization Relu Regression

CNN Efficiency vs FCNN

Sound Quality

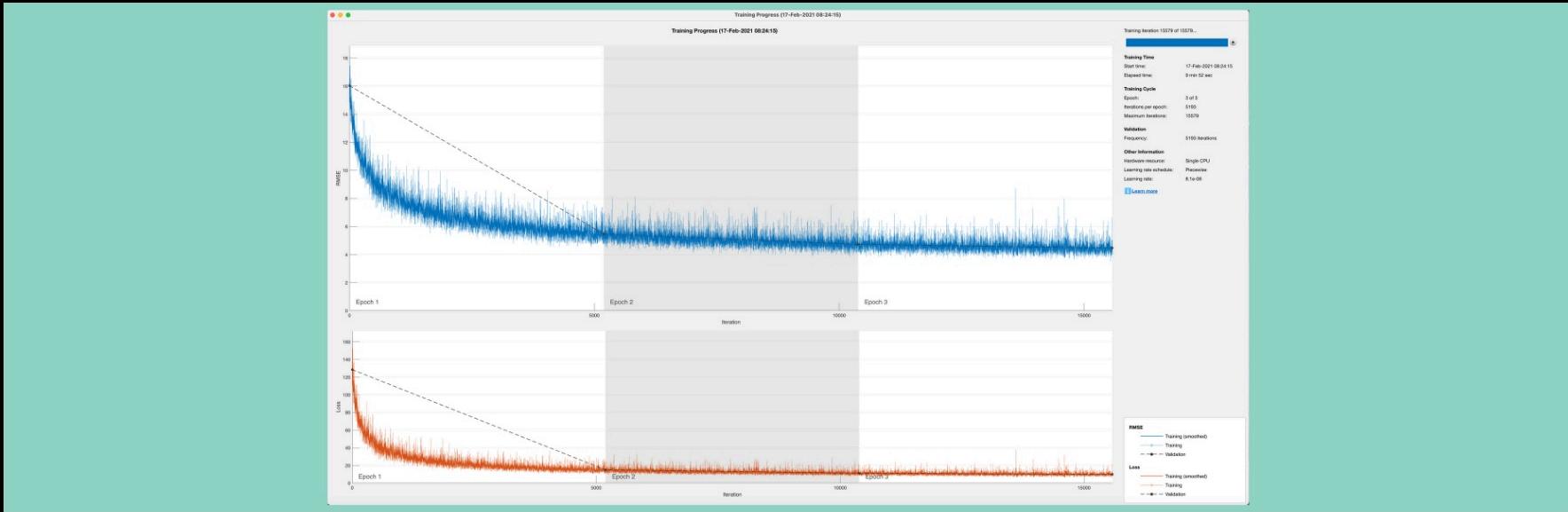


#### FCNN NETWORKS

```
layers = [ imageInputLayer([numFeatures,numSegments])  
          fullyConnectedLayer(1024) batchNormalizationLayer  
          reluLayer fullyConnectedLayer(1024)  
          batchNormalizationLayer reluLayer  
          fullyConnectedLayer(numFeatures) regressionLayer ];
```

# Predictive Modeling

---



## TRAINING AND TEST EVALUATION OF NEURAL NETWORKS

Results of MSE, RMSE and LOSS - Min Values

CNN loss : 0.034183663230164854 CNN acc : 0.9867579853041704

FCNN loss : 0.08263185542204896 FCNN acc : 0.9764345177383123

Noise Suppression AI

# Data Science Life Cycle

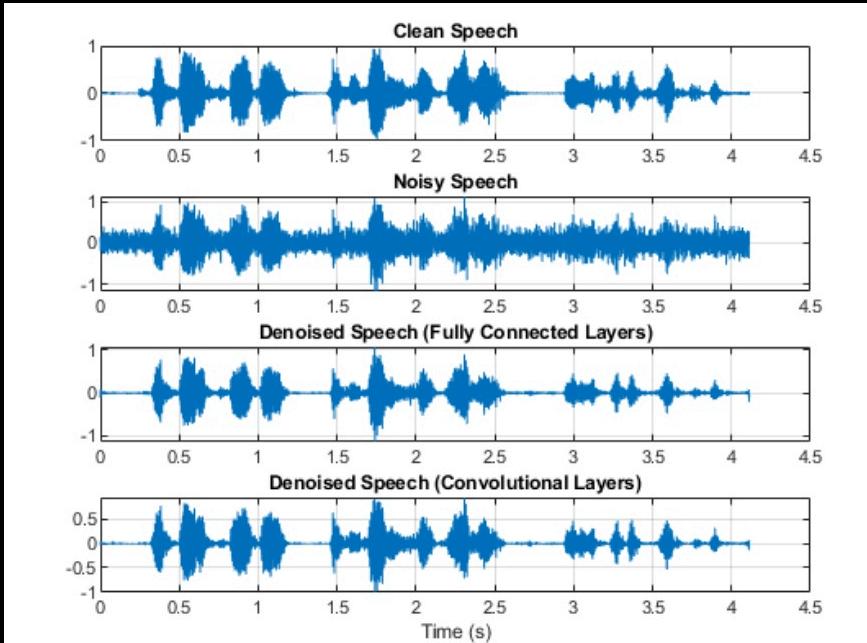
---

## **RESULTS - VISUALIZATION AND LISTENING**

Data Signal Results - Visualization and Listening  
Subjective and Objective Metrics

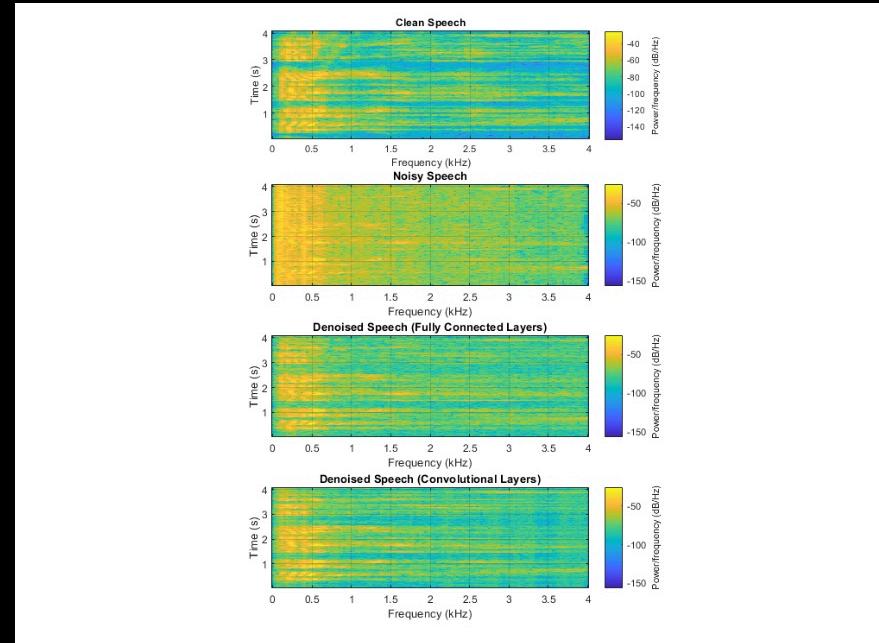
# Results

---



**SIGNAL FLOW INPUT TO CONVOLUTION NEURAL FEATURE ENGINEERING**

Target Speech Results



**SPECTRAL COMPARISONS**

Noise Suppression AI

# Data Science Life Cycle

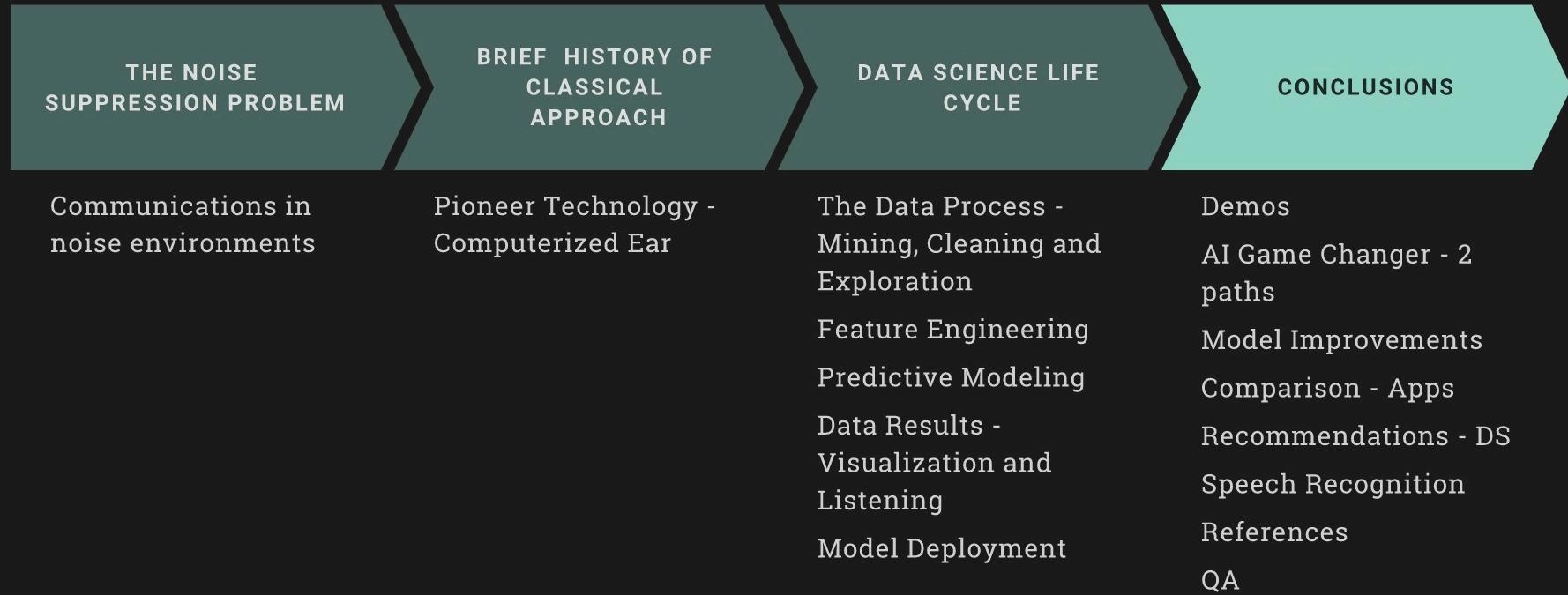
---

## MODEL DEPLOYMENT

Model Deployment - App

# Agenda

---



# **Video**