

Speech Process

Handwriting Recognition space

↓ ↓ ↓

This is a Bangalore doll

alphabets → building blocks

a, b, c, ... }

Speech basic units → Phoneme

phoneme → spoken alphabets

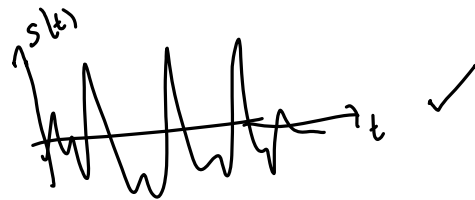
good → { |g| |uh| |d| }  
          ↑     ↑     ↑  
          ग     अ     ड

45 - 60 → phonemes

दे    घर    जा    रहा    है  
      ↑    ↑    ↑  
      space  
अ    आ    इ    ई    ..  
क    ख  
-----  
ग    → Tamil alphabet

Hindi Alphabet

English } a, b, .. z  
Spanish } → é  
German }    ö ←



I am going to Bangalore

I am going to Bangalore! ✓✓  
↓ OCR

I am going

maiin ghar ja

---

Week- 5 (Speech)

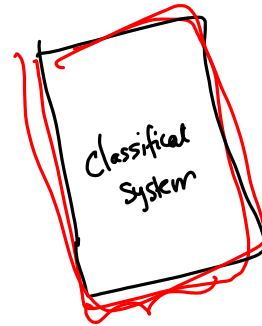
Speech- Expt 1

Given audio clip → Identify Language

# PATTERN CLASSIFICATION



① Features  $\Rightarrow$  help me distinguish languages.



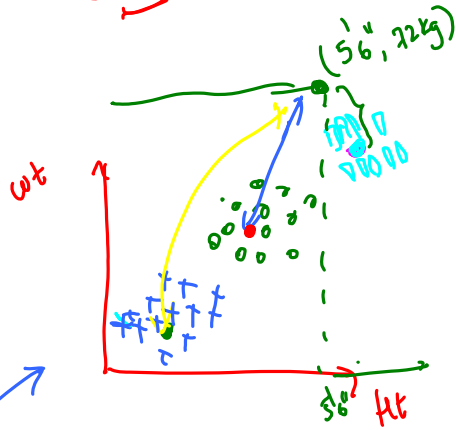
Language

Hindi  
Tamil  
Telugu

Labelled Train Data	
AC1	Tamil
AC2	Telugu
AC3	Tamil
AC4	Hindi
AC5	Tamil

TRAIN

Ht, wt  
Features



Train  
Labelled Data

TRAINING

5'6" 72 kg

Pattern  
Classification  
Problem

Inferencing step  
Testing step

VALID } hold out

Problem we will address in Colab

→ Hugging Face → Audio Clips from many language  
Dataset

→ 50 Audio Clips of English

→ 50 Audio Clips of Spanish

→ 50 Audio Clips of Romanian

Training  
Data

Labelled

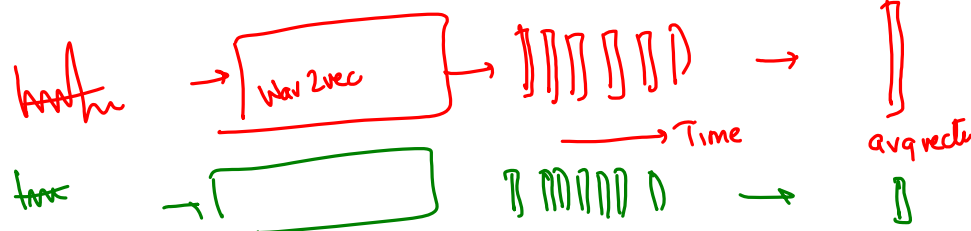
Features

Representation Learning

BERT } Pre-trained  
↓  
Representation

Speech  
→ Pre-trained

Wav2vec 2.0 Model → For Extracting Speech Representations  
Neural Network ⇒ Vector / Speech Representation



Representation

English }  
Spanish }  
Neural Network  
classify

