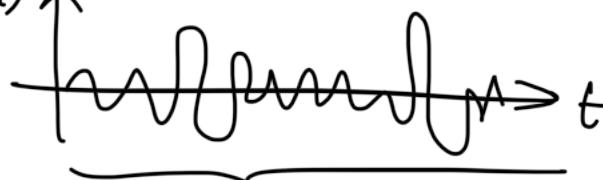


L1 Intro to Speech

- completely diff. from NLP.
- we generate a waveform.
↳ signals $s(t)$ ↑


continuous across time

- pitch variations can cause a comp. different sense. Speech is different than audio (can't show as a sparse matrix, harmonics are smooth).
- speech features can be correlated to handwriting recognition.

speech the building blocks are phonemes. (spoken alphabet) good → /g/ /u/ /d/	handwriting This is a doll. ↓ space ↓ alphabets are the building blocks.
---	--

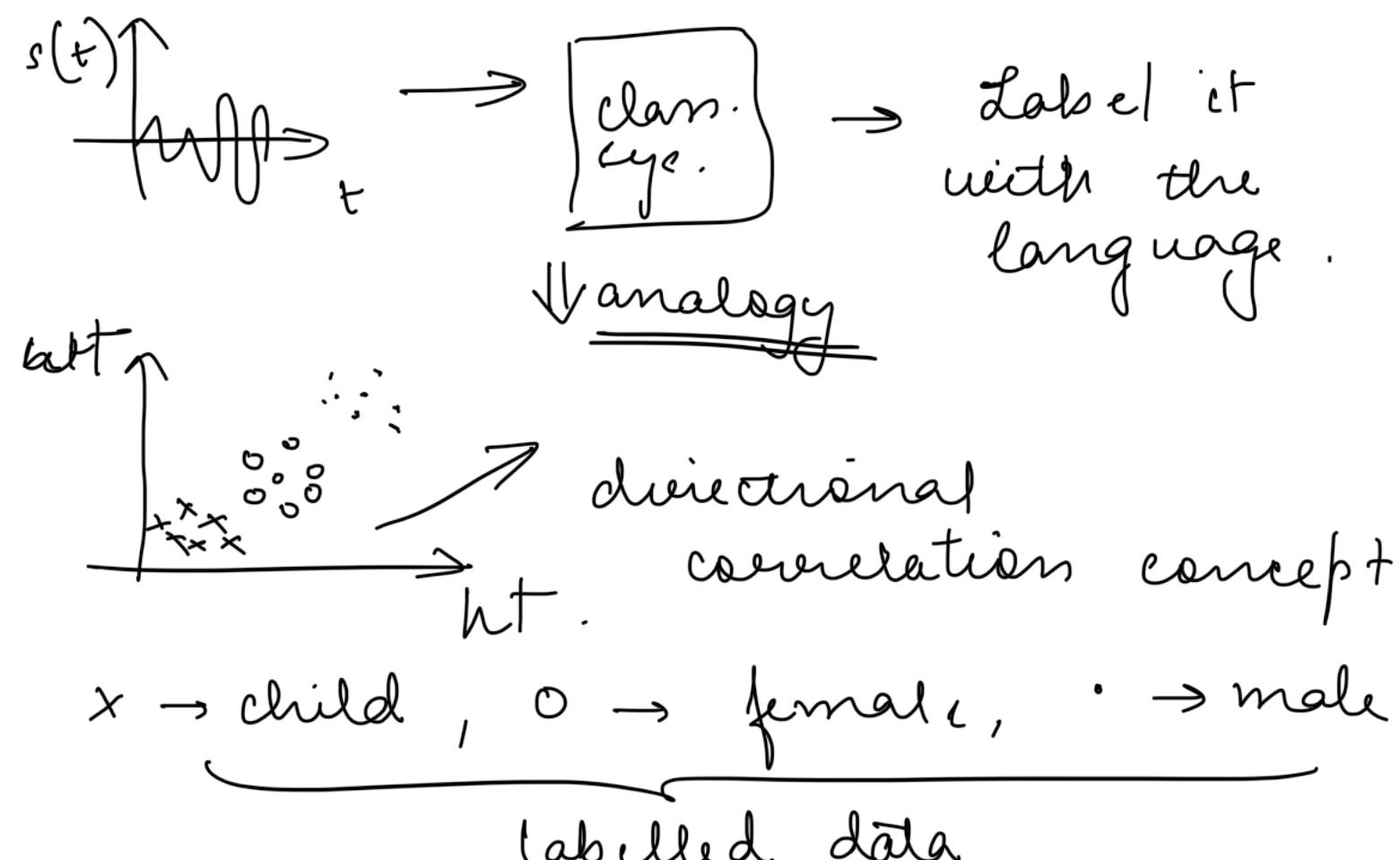
- each lang. has 45-60 phonemes are present.
- speech has some additional features like → understand gender, differentiate

btw man, women, child, estimate an age, underlying emotions or resp. issues, identify a known person

- there 4 experiments in week 5-8.

L2 Language Identification - Intro

- given a audio clip \rightarrow identify the language
- automation system in support.
 \hookrightarrow pattern classifier



1. training step, then
2. inferencing step / testing

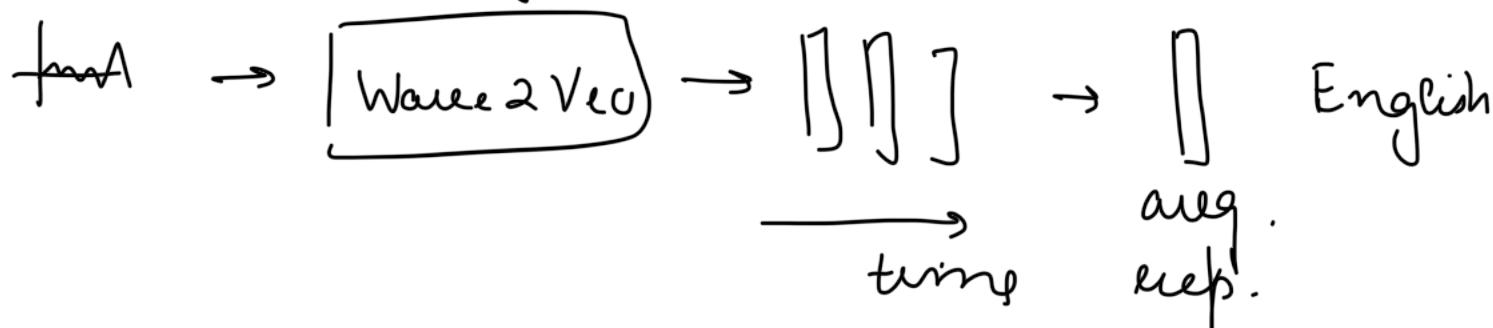
- we have to use some feature to give a representative vector for each English. \Rightarrow help to distinguish language.
- sep. dataset cf. the validation dataset to generalize the model.
- build a classifier to identify 3 languages.
 - dataset from HF \rightarrow audio clips
 - take 50' from eng., spanish & romanian.

training labelled data.

- representative features are built here.
 \hookrightarrow representation learning
- in text, we use BERT.

Speech \rightarrow Pre-trained model

Wave2Vec 2.0 model \Rightarrow for extracting speech representations



we will use neural networks

Classifer

L3 Demo

