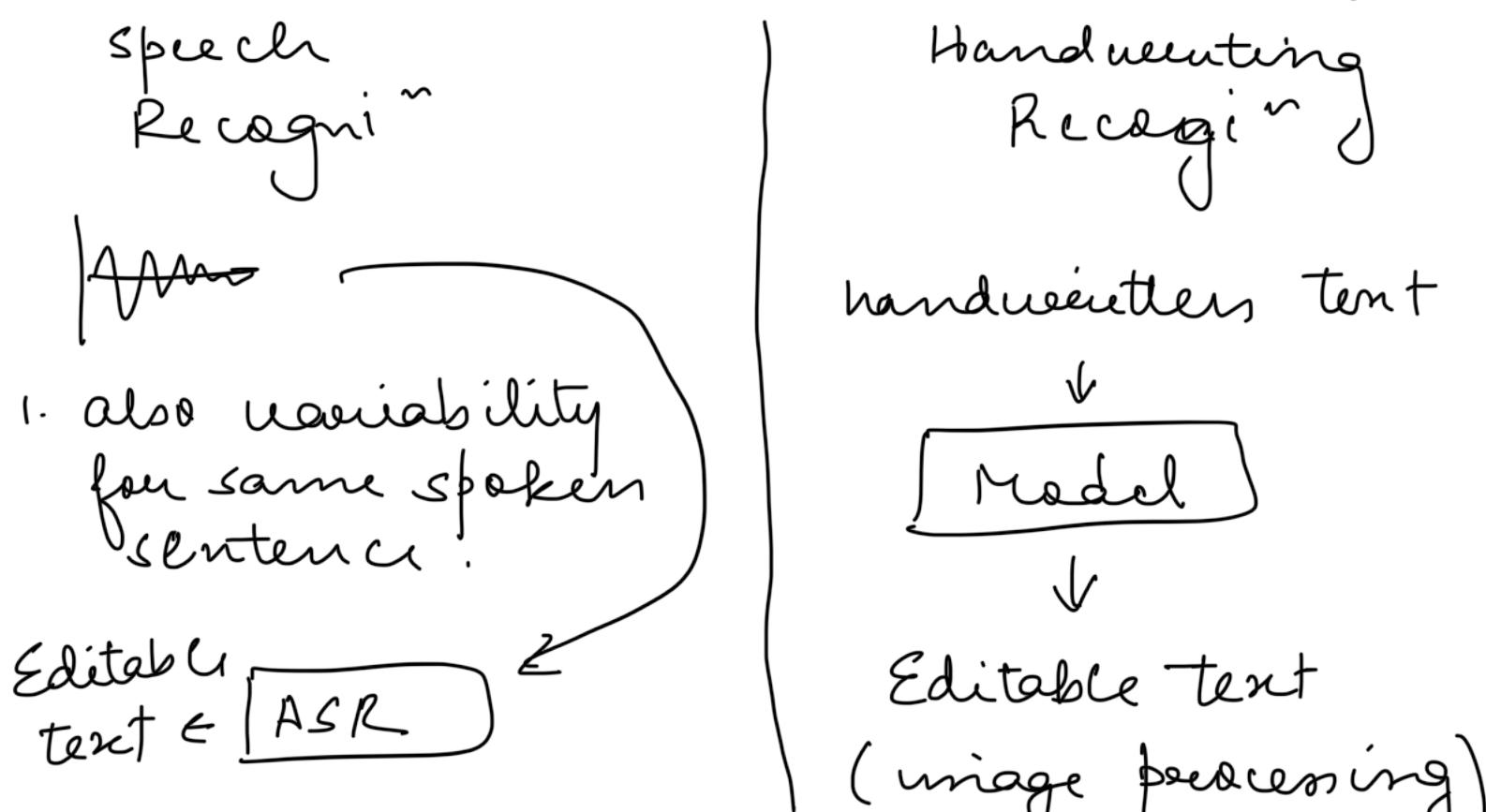


1 Introduction to ASR & Handwriting Recognition

- ASR → Automatic Speech recognition
(who said what)
or (speech to text)
- the goal,
~~text~~ → ^{digital conversion}
speech signal → **ASR** → 'tent'
transcripⁿ
- it is bit difficult, the signal keeps changing even if it is the same speaker, saying the same thing.



Process -

take speech signals

↪ take the phonemes

"and" → [ə] /ɪn/ /d/

↓
map it to the
text.

if mistakes,

again use the
language model
to rectify those.

1. variability across diff. humans.
 2. diff. lang. needs diff. models of H.R.S. It checks for strokes if mistakes
- ↓
- Language model
- calculates the probab. for each word.
and rectifies-those mistakes!

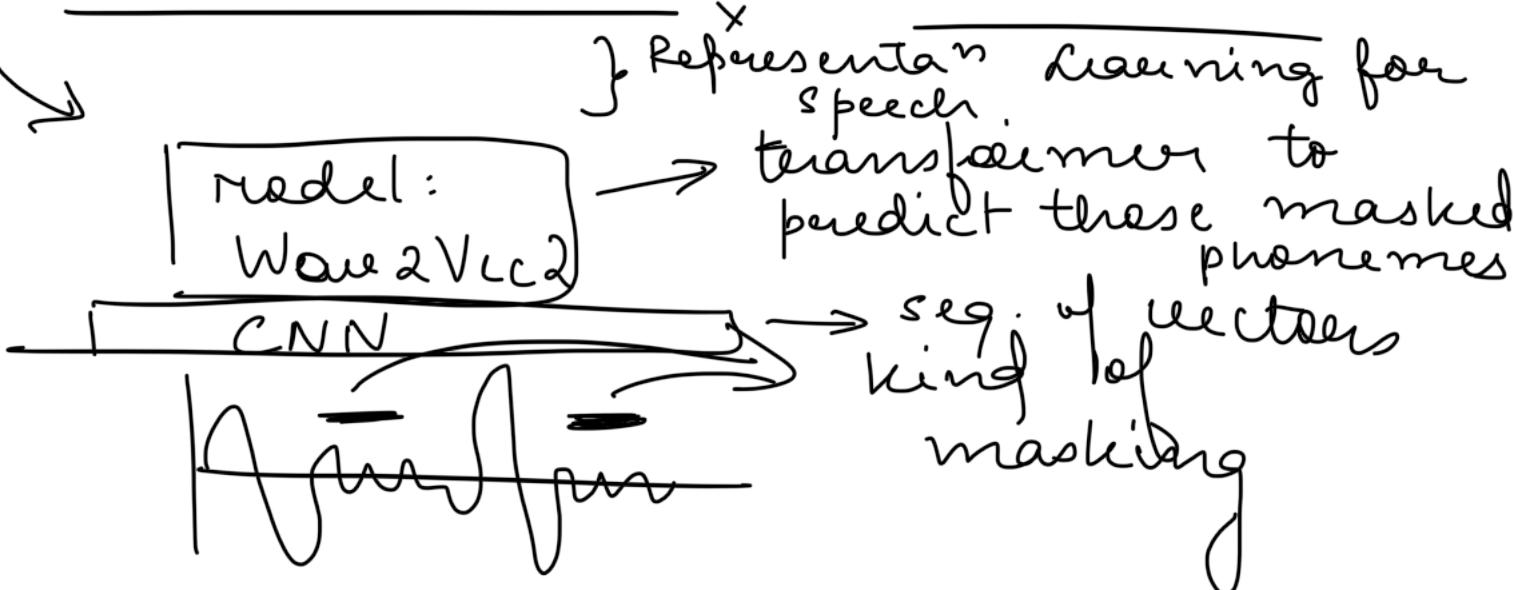
L2 ASR Workflow & Components

- ① first use DL.
- ② use the idea of a pre-trained model.
- ③ Fine-tunes the model for ASR based tasks

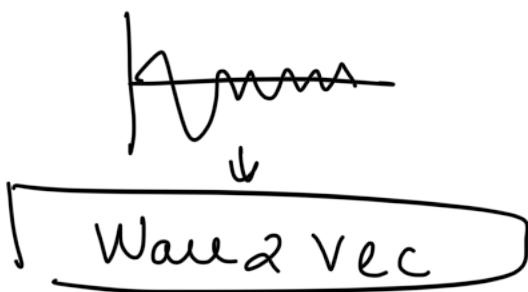
— × —
NLP → BERT (idea of masking)

the idea of self-supervised learning

get the embedding for each word



- use the idea of a similarity index (mostly cosine index)



- fine tune on labelled ASR data.



get the words here

- for evaluaⁿ of ASR →

errors ① substitution e.g. going
= s +
koring

② deletion → delete a word.
= D

③ insertion → inserting sth.

= 1 extra .

$$\text{Weed Eraser Rate} = \frac{S + D + I}{N} \times 100$$

↳ in term of %.

L³ Dens