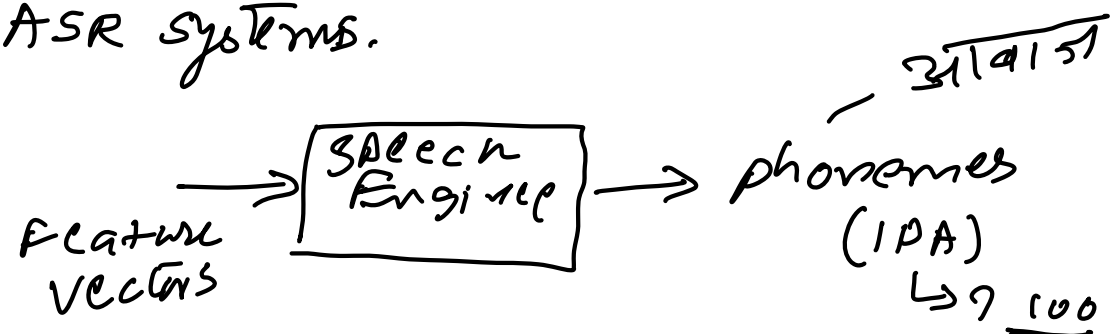
Speech Recognition Models:

- significance of speech recognition
- there is a speech Engine, as part of all ASR systems.



Is speech Engine language dependent or independent of language? - Independent

ASR requires domain information:- vocabulary - (Language domain)

set, pronunciation model, word usage model (we require all these database in a ASR System).

we see only the feature vector (given)
and we need to discover words spoken
(phoneme sequence).

we have feature vector & we need to
find out the words spoken

wordspoken ⟶ feature vectors
A (cause) B (Effect)

is it A ?

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

find probability of A given

Bayes theorem

whether A_1 has caused B, then

$$P(A_1|B) = \frac{P(B|A_1) * P(A_1)}{P(B)}$$

· Whether A_2 has caused B, then

(2) ✓ $$P(A_2|B) = \frac{P(B|A_2) * P(A_2)}{P(B)}$$

$$\vdots$$

$$P(A_n|B) = \frac{P(B|A_n) * P(A_n)}{P(B)}$$

If A ⇒ book, look, took, cook, ---
A_1 A_2 .... A_n

then select the word of maximum probability, and that is corresponding to this

given feature vector.

Since $p(B)$ denominator is common in all, it can be dropped to save computation and it will not affect the result

∴ given feature vector, we can find probability of a word "A" by

$$P(A|B) = p(B|A) * P(A)$$ ← Naïve Bayes theorem

If no. of words are
are V (vocabulary), and length of words

is N (sentence length), then no. of possible sequences are: ?

Vocabulary is (V) 2 → (0 & 1)

and length of sentence (string) = (N) 3

then possible no. of sentences = 2^3

∴ possible no. of sentences = V^N (EXPONENTIAL)

100^{10} = 100.....0

no. of

But many sentences can be ignored due to language →

give me a call. ✓

a call me give. ✗ give call me ✗

Hence, all sentences probability is not required to be calculated.