Agenda:
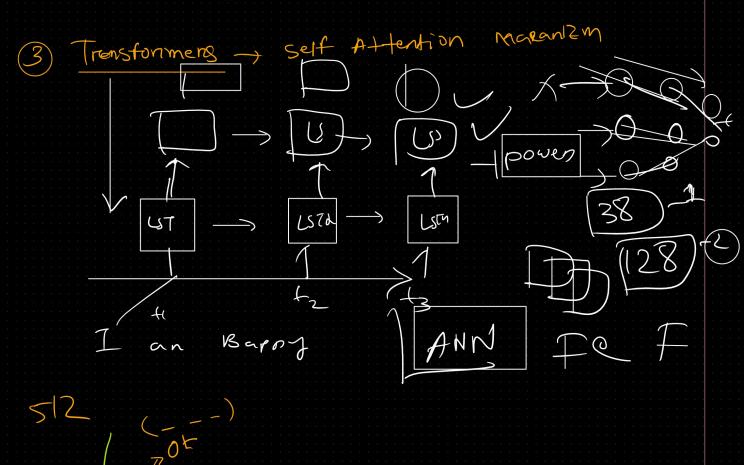
① Transformers Architecture

① Encoder and Decoder
   → Small contant

② Attention model → give Attention to Important words
   → Encoder and Decoder modified

["very bigger!"]

③ Transformers → self Attention Mechanism

I am Bappy

512   (- - -)
        → ot

$$0.8\ell \quad | \quad 0.5 \quad | \quad \cdot \quad)$$

"My name is Barry"