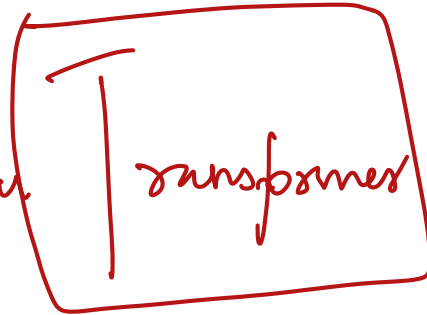


Transformer

i/p

Hello, How are you

wrote Story, Once



o/p
doing
upon

Predict the next word

Encoder

↑ ↑ ↑ ↑

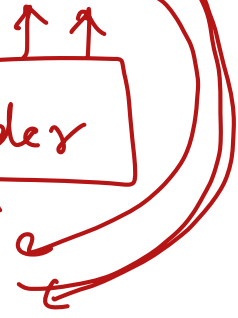
all i/p at
once

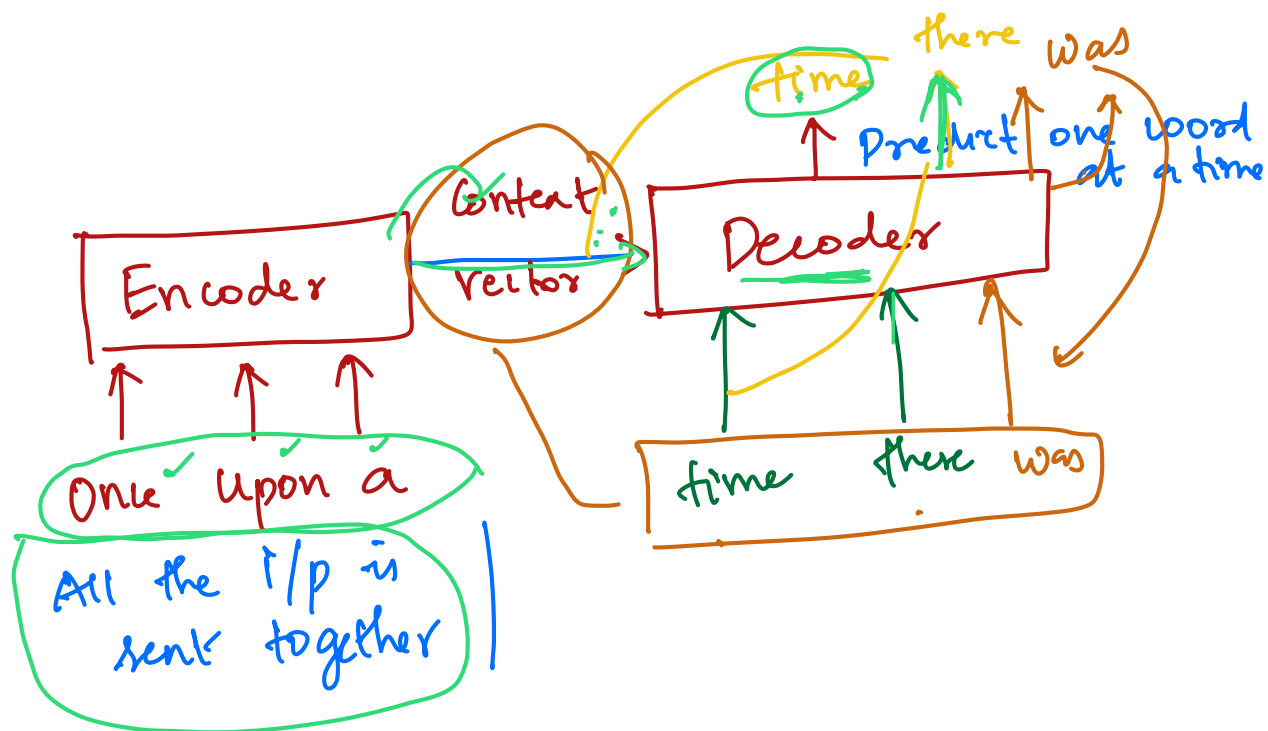


Decoder

↑ ↑

o/p (one
at a time)





Decoder at time t_1

Context vector \Rightarrow o/p 1

t_2

Context vector + (o/p 1) \Rightarrow o/p 2

t_3

Context vector + o/p 1 + o/p 2 \Rightarrow o/p 3

o/p 4 < End > it will stop

Chat GPT

→ next o/p

token

→ <End seq>