RNNs

Main application area: NLP & others

Ex. Not interested at this time

AUTO COMPLETION

Will let you know if it changes in the future

X → Y

Translation

English → French

NER
Named entity recognition

X: Rudolph Smith bought 1000 shares of tesla Inc. in March 2020

Y: [Rudolph Smith]$_{\text{Person}}$ bought 1000 shares of [tesla Inc.]$_{\text{Company}}$ in [March 2020]$_{\text{time}}$

Sentiment Analysis

X: Text

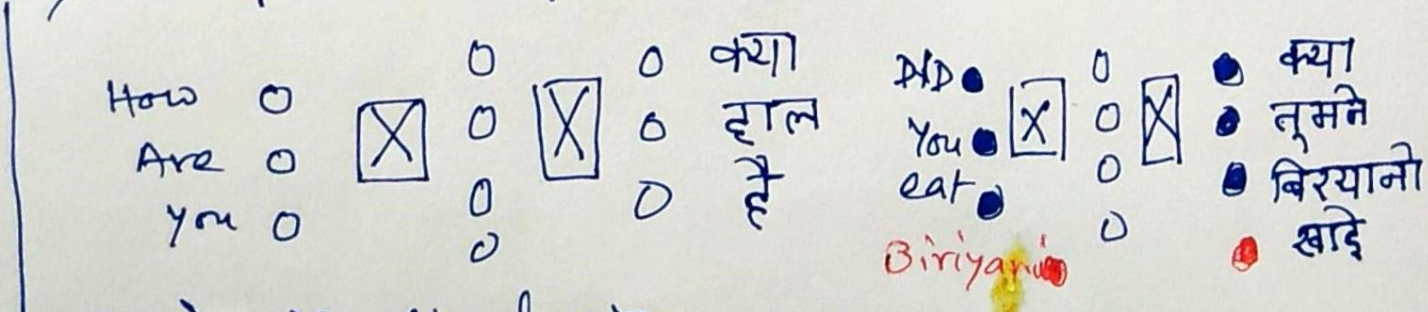
Y: ⭐ ★ ★ ★ ★

Why Not ordinary ANNs?

⇒ sequences are Important.

How are you? ≠ You how are?

⇒ Input & output size are varying



→ . No fixed size of Neurones can be decided upfront.

⇒ Too much computation in traditional settings.

Say, 25,000 word vocabulary we construct upfront

How → [0, 0, 0, 0, 1, ---- 0] ONE HOT Encoding

Are → [0, 0, 1, 0, ... 0]

You → [0, 0, 0, ... 0, 1, 0]

output also need the word vocabulary, say 42,000 words

क्या → [0, 1, 0, ... 0]

दाल → [0, 0, —— 1, 0, 0, 0]

है → [0, 0, 1, ... 0]

⇒ Not one to one mapping

Ex. On sunday I ate golgappa

I ate golgappa on sunday

→ ~~many~~ meaning is same but ANN dont share parameters.

Summary: 3 Major problems:

→ Variable size of input/output of neurons

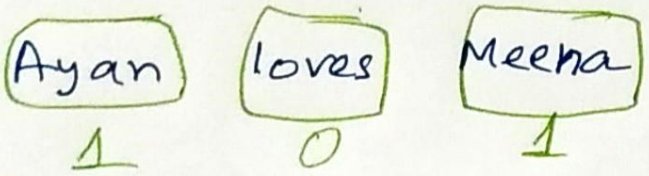
→ Too much computation

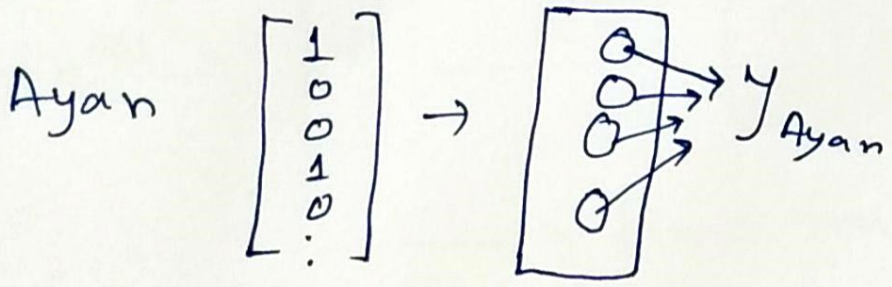
→ No parameter sharing.

RNN

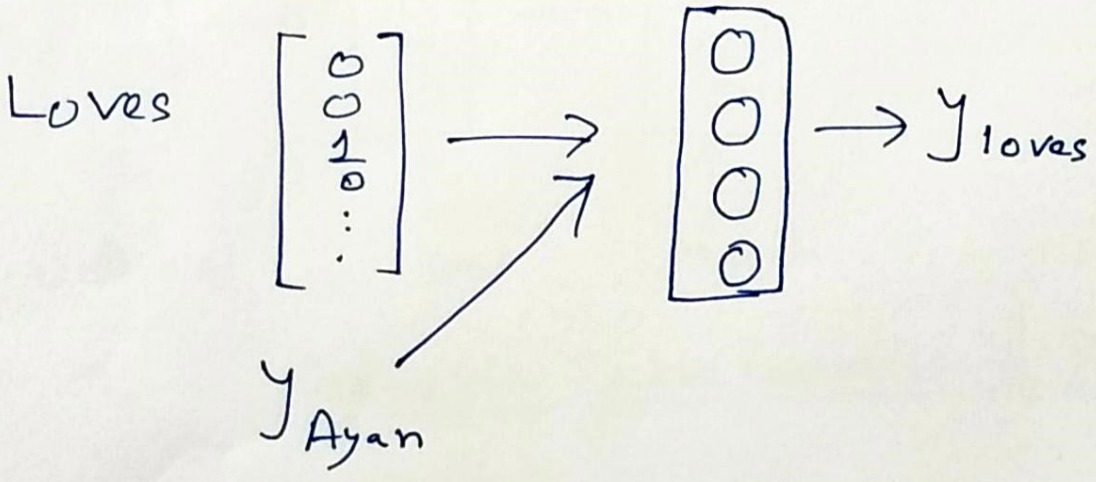
Named Entity Recognition

Ayan loves ~~xx~~ Meena

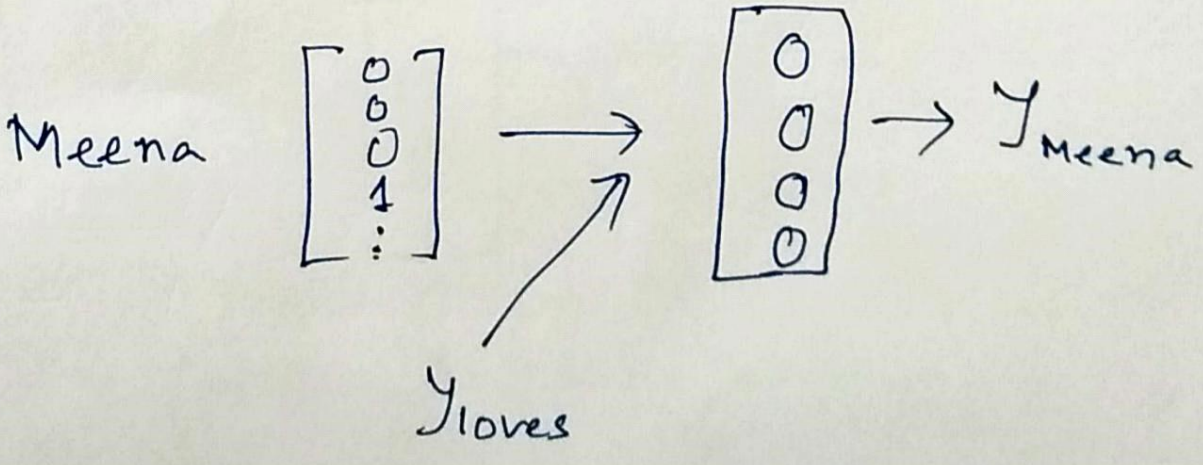




Next

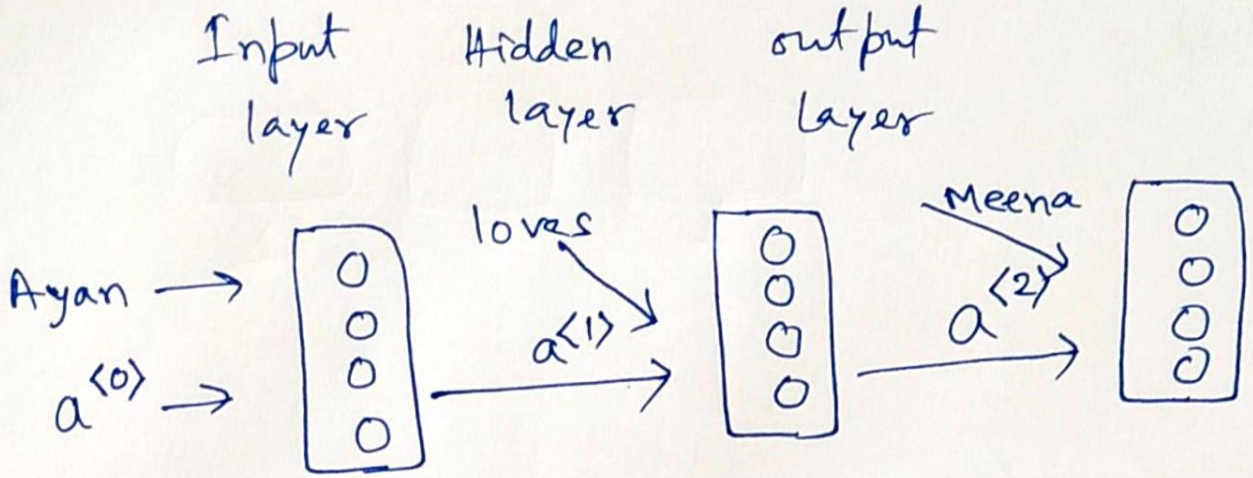


Next

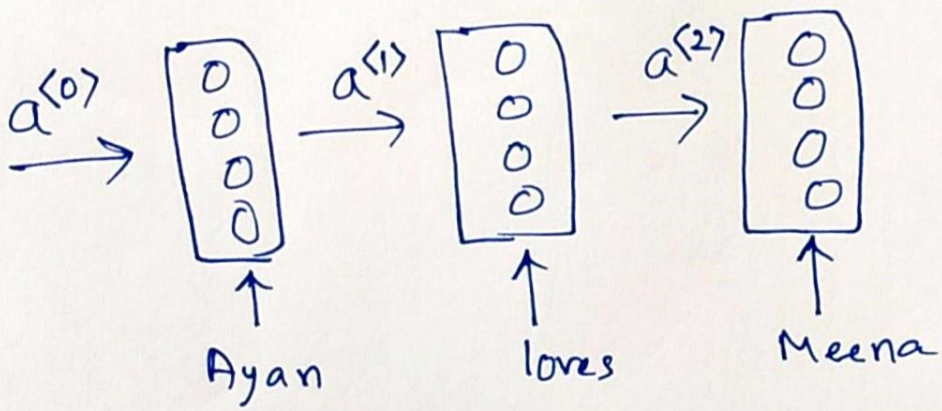


RNN

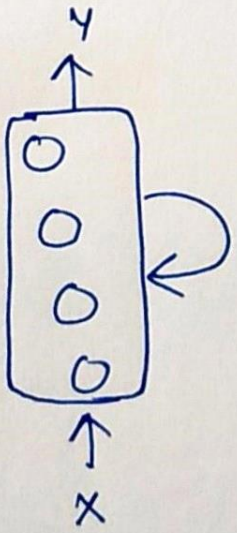
This network has only one layer



→ time travel in only one layer



Generic Representation of RNN

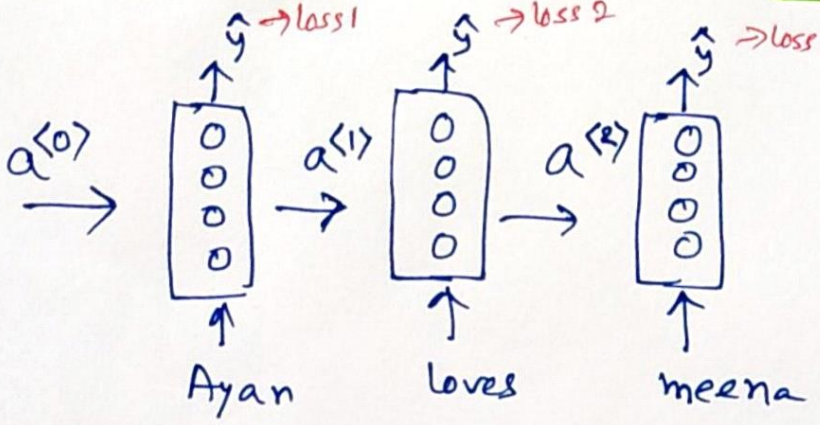


RNN

NER Training

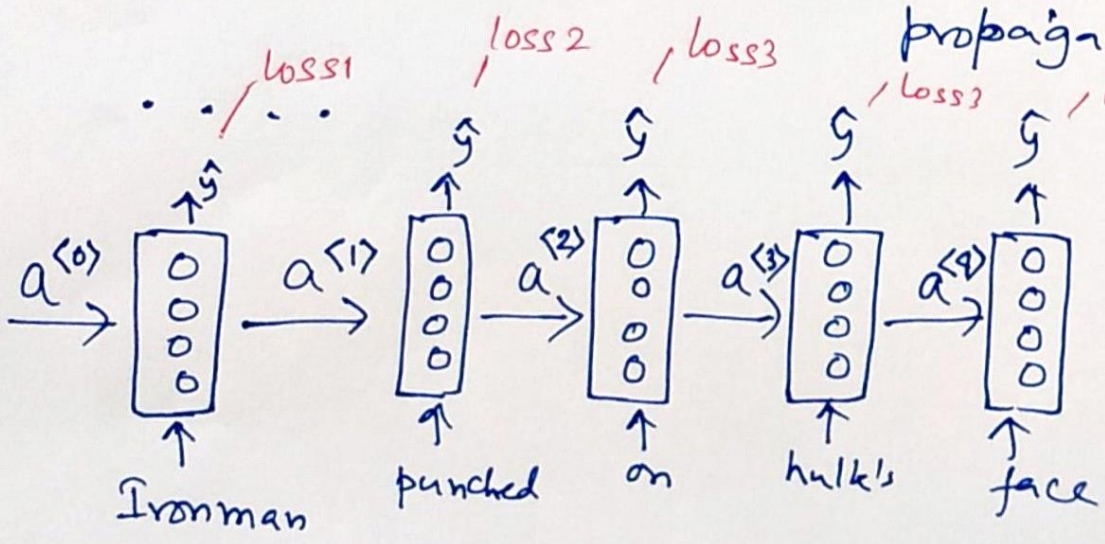
Training data

X	Y
Ayan loves Meena	1 0 1
Bob told Ahmed that pizza is delivered	1 0 1 0 0 0 0
Ironman punched on ~~hulk~~ hulk's face	1 0 0 1 0



Total loss $= \sum_{i=1}^{3} loss_i$

use gradient descent & back-propagate the loss



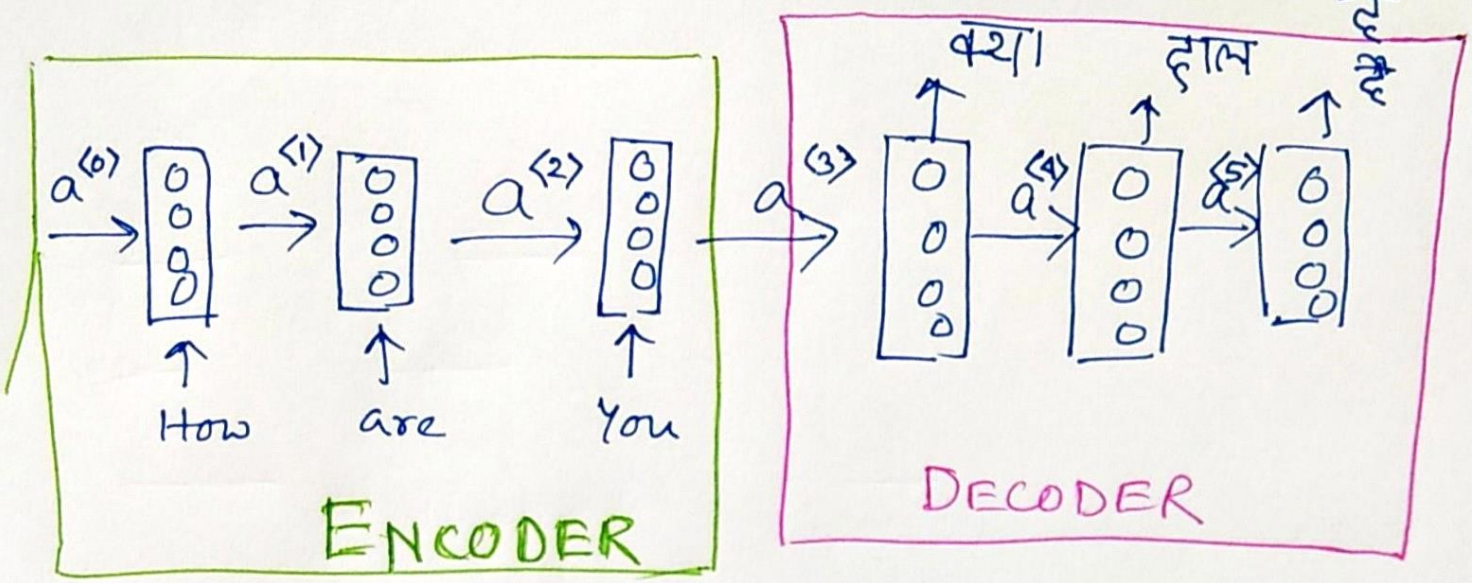
Total loss $= \sum_{i=0}^{4} loss_i$

RNN

Pg 6

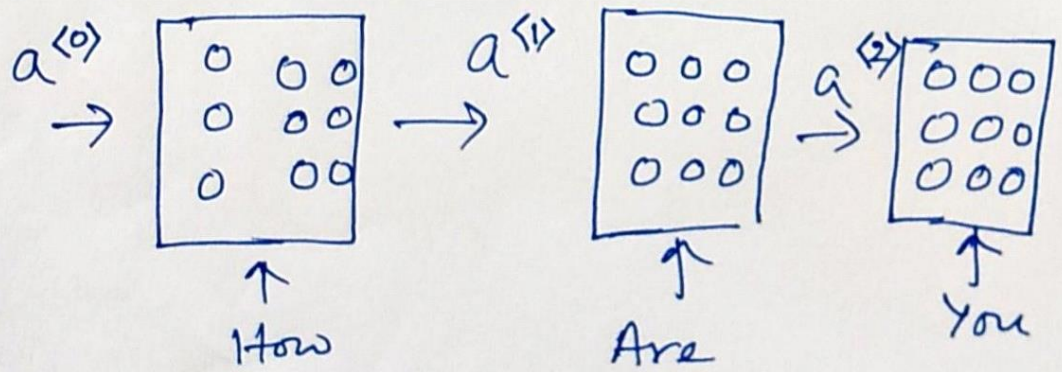
Language Translation

How are you → क्या हाल है



THE LAYER DOES NOT NEED TO BE A SINGLE LAYER ! IT CAN BE DEEP

DEEP RNN



END