~~Machine Learning~~

LSTM Session

Vanishing & Exploding gradient.

function < 1 ⇒ Vanishing

> 1 ⇒ Exploding

LSTM ⟶ 8 weight matrices.

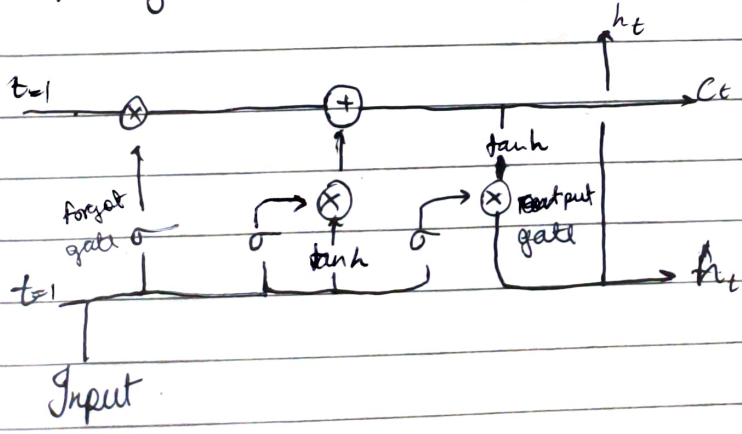
Improved ~~vanilla~~ version of RNN which is able to resolves the problems ~~the~~ of the vanishing and exploding gradient of RNNs.

- Output ⟶ Memory
  If function

- Forget Gate : ~~the~~ Hidden + Input
  ↓
  Sigmoid

~~Output Gate :~~

Output Gate : ~~tanh~~ ⟶ (X)

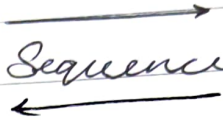


LSTM in Simple:

* Basically, while going through the sequence it uses it complex function & decides whether or not to keep a token in its memory.

* Doing this it can build context for the sentence & give better output.

* Bidirectional LSTM :



Is able to build context on the sequence by going start to end & end to start.

Drop out : In a layer as per specified percentage it deactivates the some percentage of the nodes
- It prevents over fitting

Epoch: Iterations through the NN (neural network), and improves its weights on every epoch.

Batch size :
- Scales down the batch input size.
- Due to this it ~~does not~~ remains light on the processing media part

Page No.
Date

Simplified LSTM: GRU.

- Instead of 4 sigmoid functions, it only has 3 sigmoid functions

---

OVA-RIZZ

1> A company
Challenges to outperform NVIDIA's chips.
Best struggling in the market because of domina
of NVIDIA.

2> Data Provinence Explorer: Researches work on improving model's accuracy by working on the data fitting