



$$P = \frac{P - M}{\sigma}$$

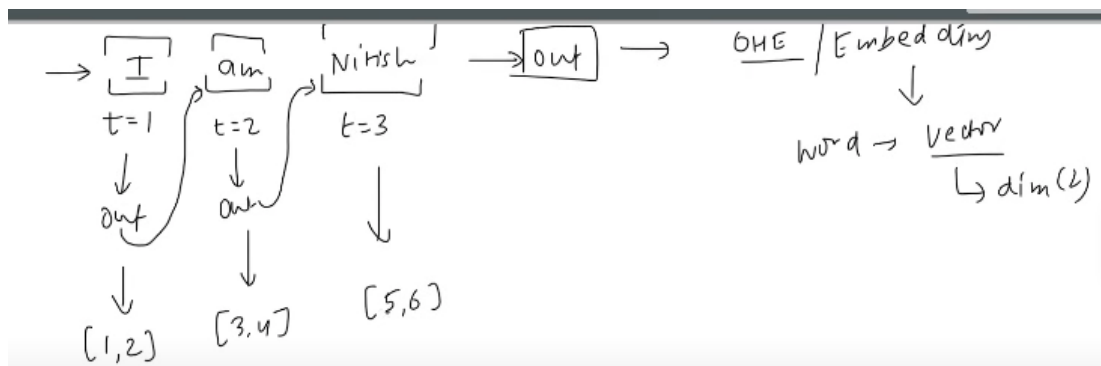
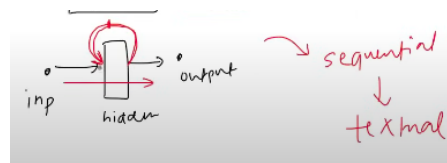
<https://colab.research.google.com/drive/1zmUGFtmyceXvAAjIDM6ncK40fwCDlqzZ>

RNN using pytorch

- Mainly textual , sequential data
- RNN arch - <https://youtu.be/4KpRP-YUw6c?si=ll5boxGJIZvt8w5G>

RNN is a type of NN designed for processing seq data. Unlike trad feedforward network which process inputs independently , RNNs maintain a memory of prev inputs by using loops in their arch.

This makes them well-suited for tasks wher context and order matter such as time series forecasting , speech recog , text .

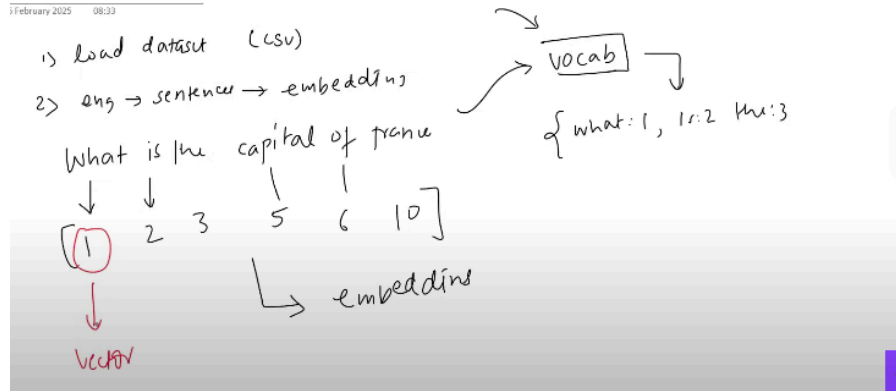


at t=1 "I" is going inside and RNN is doing calc and outputs O1 #To maintain consistency O0 is also sent to first pass(random numbers)

- "am" is sent along with o1 and get o2
- "nitish" + o2 → o3

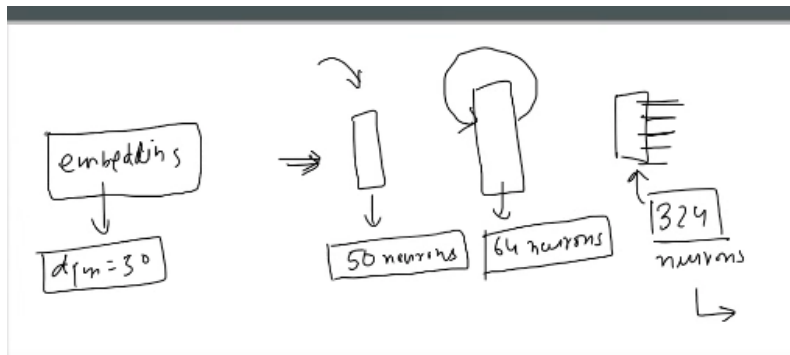
THIS IS CALLED UNFOLDING THRU TIME

- Strategy - load dataset ; ENg sentences → embeddings or OHE; BUild rnn arch ; Train eval



a vocab of all unique words is formed before embedding

- RNN arch:



- Take note of shapes of input and output via layers

<https://colab.research.google.com/drive/1JO7PGLkj8P7Jy8KM3rBHop6jaUruxvQB>

Predicting the next word

- mobile keyboards

Language Modelling is the task of predicting the next word in a sequence based on the context of previous words.

Handwritten notes on the left:

- unsupervised
- supervised
- input

Syllabus text on the right:

- What is the total duration of the course?
- The total duration of the course is 7 months. So
- What is the syllabus of the mentorship program?
- We will be covering the following modules:
- Python Fundamentals
- Python libraries for Data Science
- Data Analysis
- SQL for Data Science
- Maths for Machine Learning
- ML Algorithms
- Practical ML
- MLOPs
- Case studies

- But how do you train on a dataset of words? - unsupervised task ko supervised convert kro -
 - Iss data set mai seprate all sentences (Input : first word ; Output : second word) ; (input : first+ second word ; output - thrid word)
 - but LSTM cant understand english - create vocab like before.
- Will use nltk for tokenization (word_tokenize from nltk.tokenize)

LSTM Blog - <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>

https://colab.research.google.com/drive/1cC0O0YDObHfR_078B3ZPXSPtMu13ykGf