LSTM with attention for Stock Prediction

Revolutionizing the Stock Market, through A.i

Group 12:
RAINA Ritvik (55828826)
CHAMBERLIN Matthew Milton (55777489)

01 What are LSTMs?

02 What does the LSTM architecture look like?

03 Where are LSTMs used?

04 LSTMs for Stock Prediction

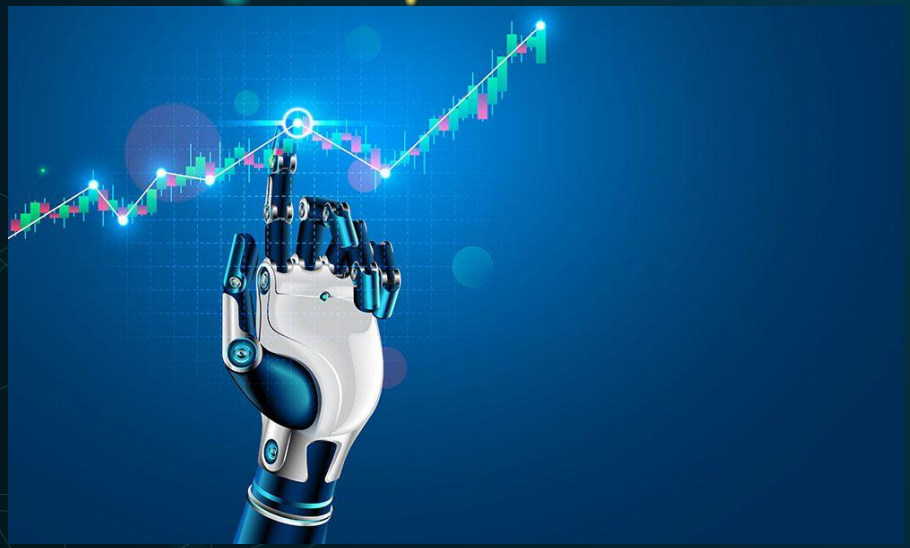
What are LSTMs?

01

Long short-term memory

- Form of Recurrent Neural Network
- Useful in time series prediction
- Takes into account contextual information as well as input to give output

LSTMs



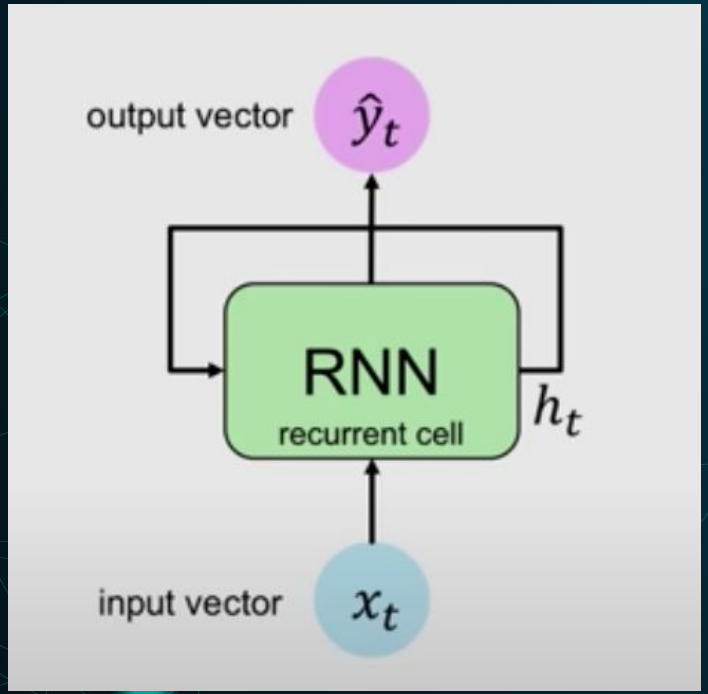
The ability for gradients to remain unchanged solves the vanishing gradient problem in traditional RNNs

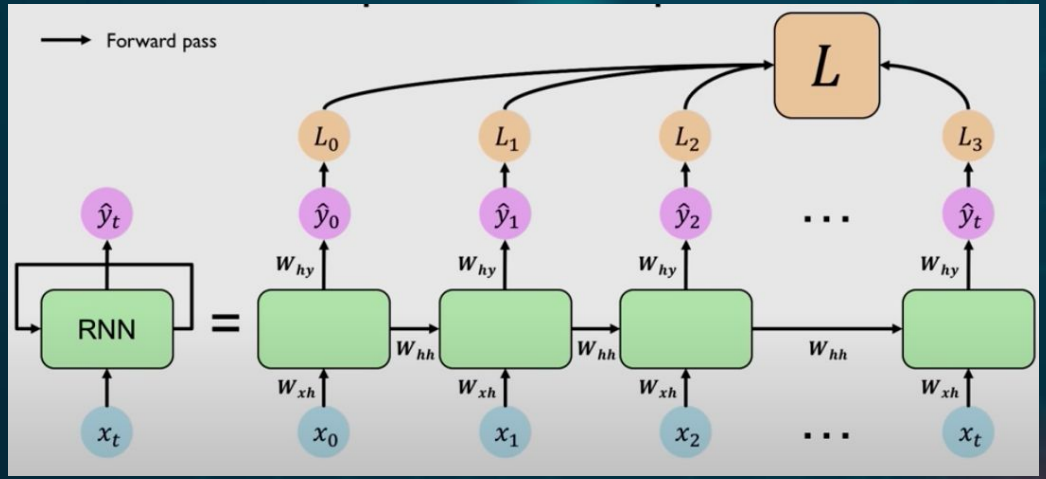
Input gate,
Output gate & Forget gate

What does the architecture look like?

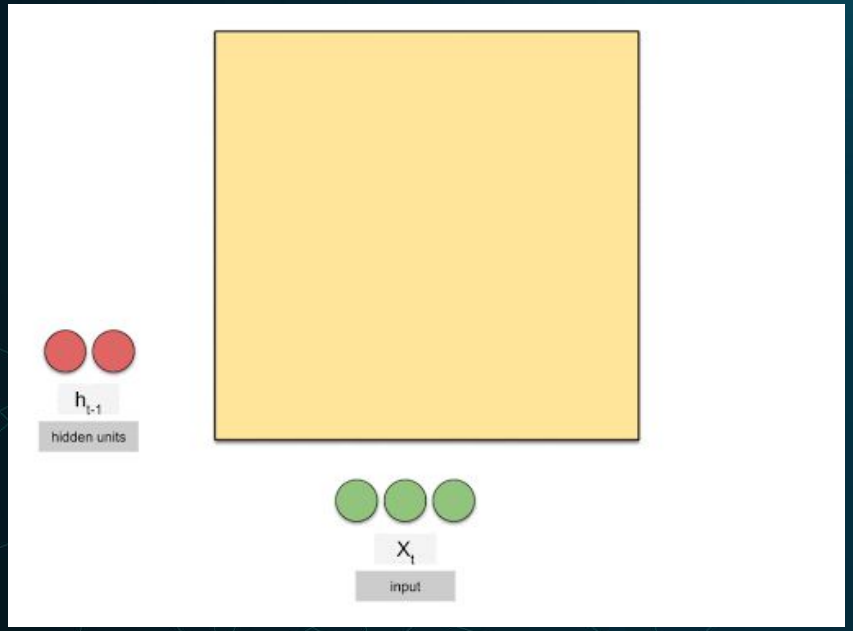
02

Recurrent Neural Networks (RNNs)





Source: MIT 6.S191 (2020): Recurrent Neural Networks



Update Hidden State

$$h_t = \tanh(\boldsymbol{W}_{\boldsymbol{hh}}^{\boldsymbol{T}} h_{t-1} + \boldsymbol{W}_{\boldsymbol{xh}}^{\boldsymbol{T}} x_t)$$

Input Vector

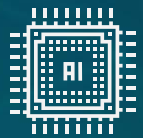
$$x_t$$

Output Vector

$$\hat{y}_t = \boldsymbol{W}_{\boldsymbol{hy}}^{\boldsymbol{T}} h_t$$

Source: MIT 6.S191 (2020): Recurrent Neural Networks and towardsdatascience.com

Long Short Term Memory (LSTM) Networks



FORGET

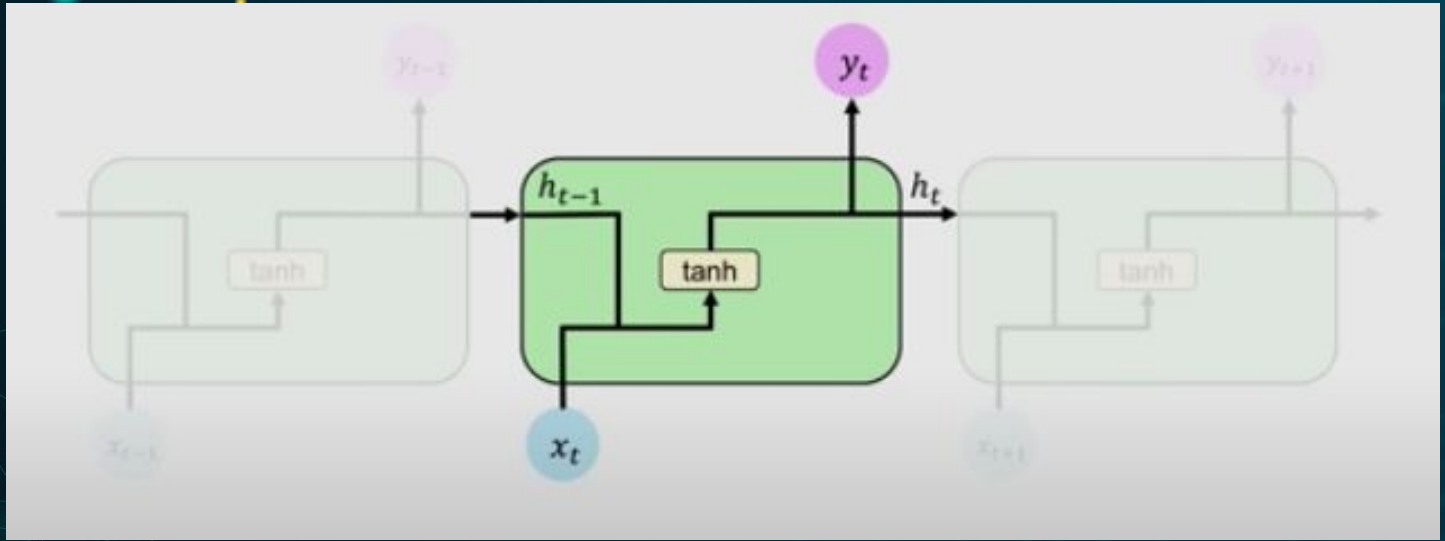
STORE

UPDATE

OUTPUT

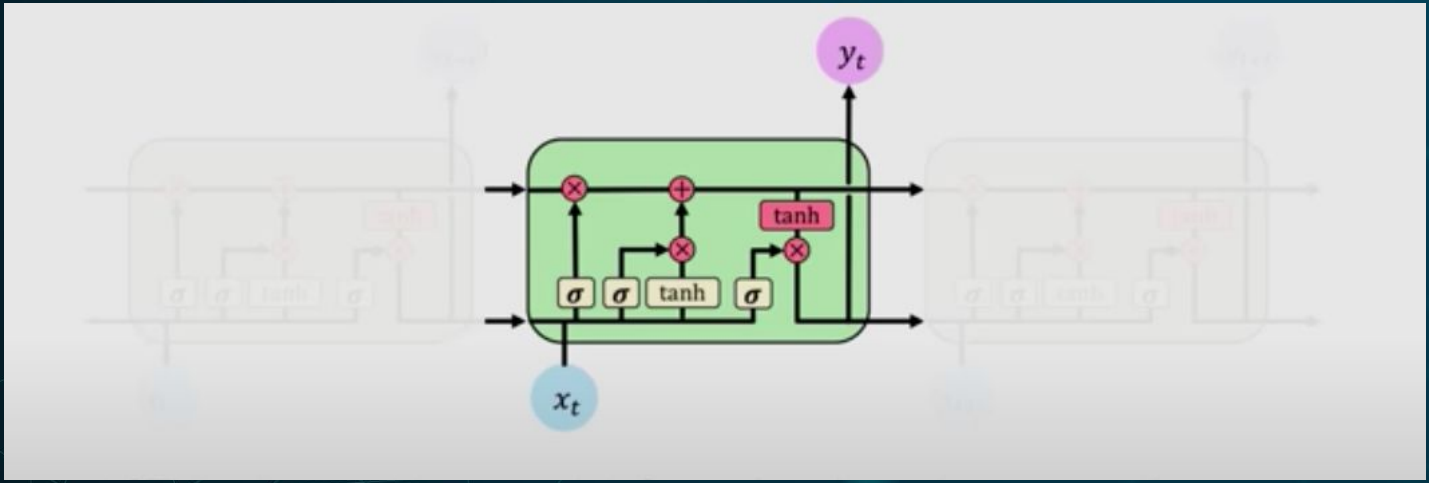
Source: MIT 6.S191 (2020): Recurrent Neural Networks

RNN Model

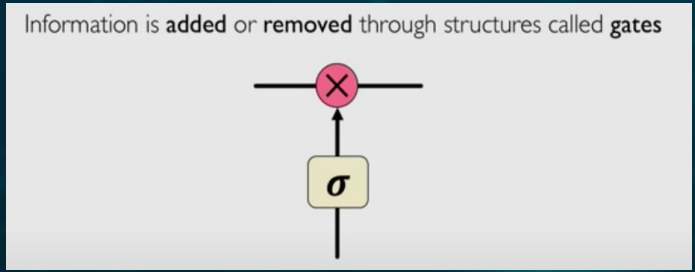


Simple RNN can be modeled

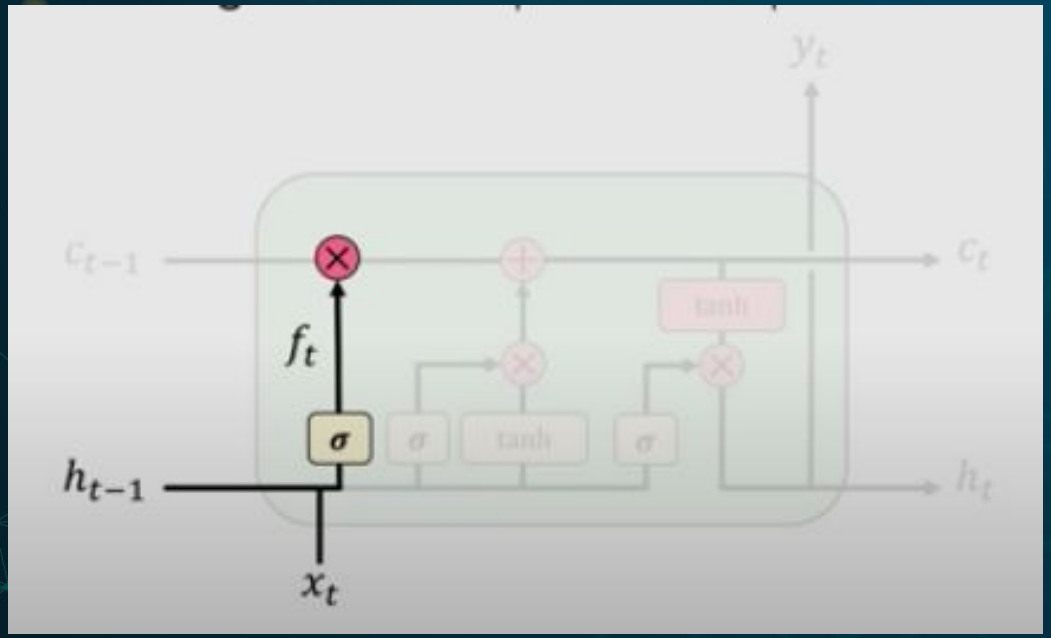
LSTM Model



What an LSTM can be modeled as

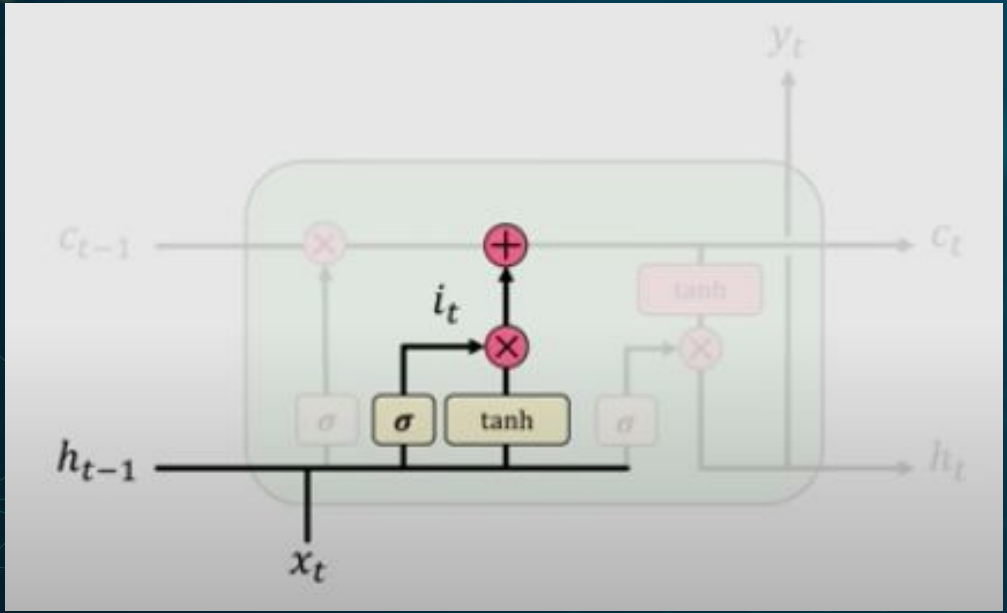


LSTMs



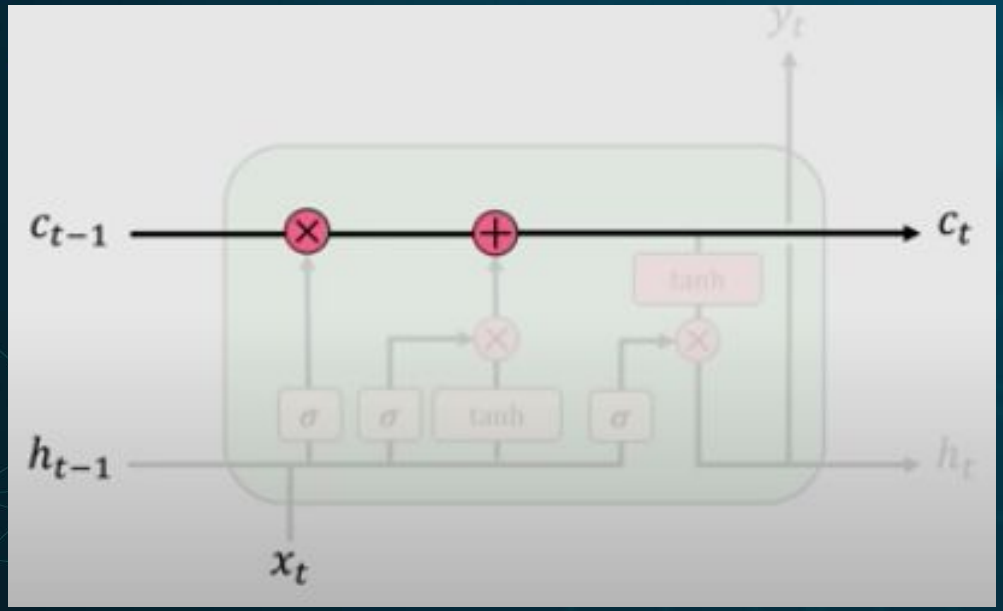
FORGET Gate - Forget

LSTMs



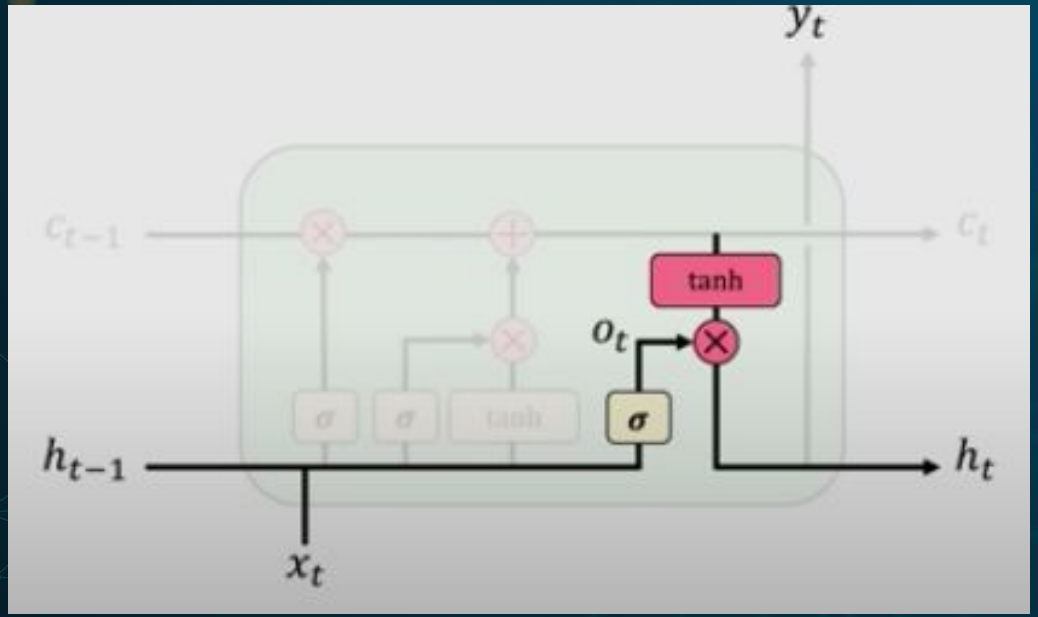
INPUT Gate - STORE

LSTMs



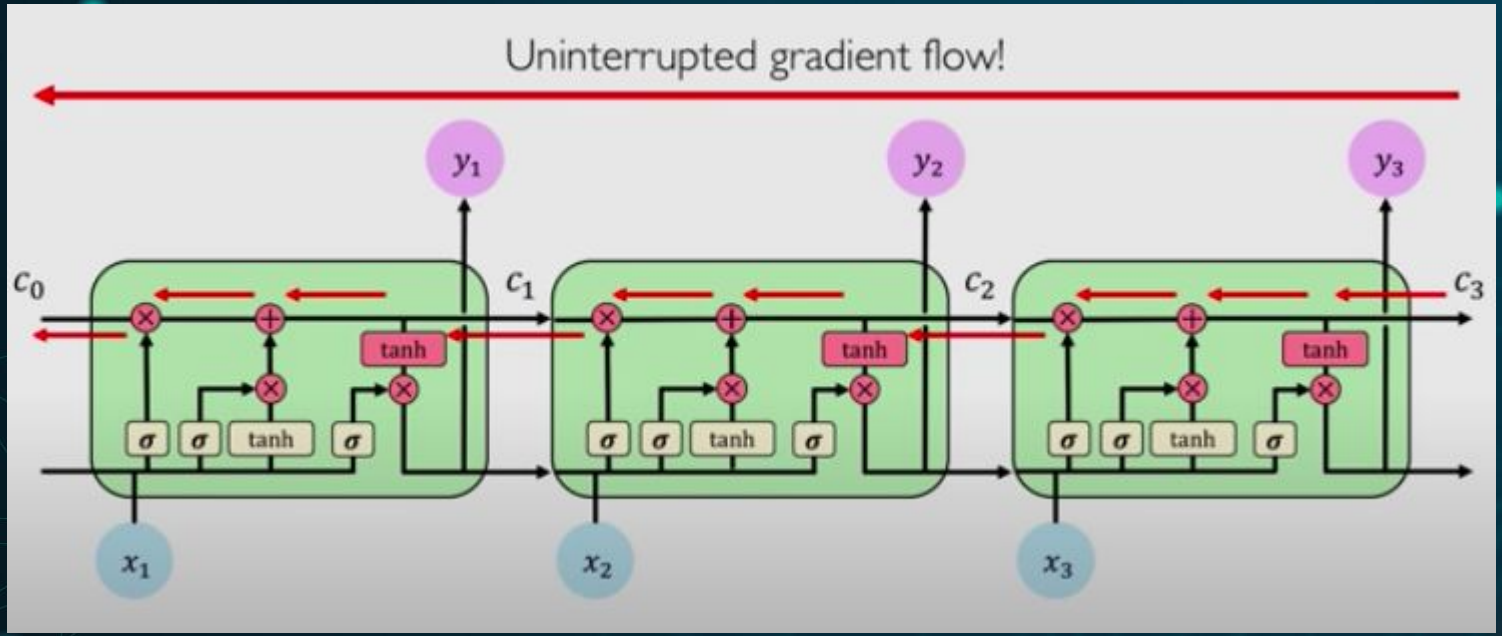
FORGET + INPUT - UPDATE

LSTMs



OUTPUT Gate - OUTPUT

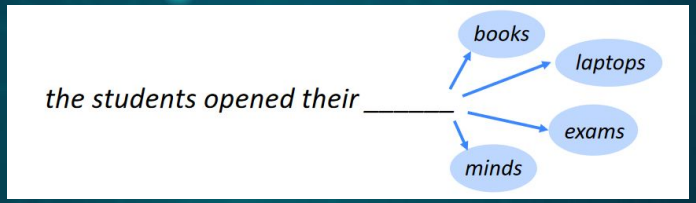
LSTMs


Uninterrupted gradient flow!
y_1
y_2
y_3
c_0
c_1
c_2
c_3
tanh
tanh
tanh
σ
σ
tanh
σ
σ
σ
tanh
σ
σ
σ
tanh
σ
x_1
x_2
x_3

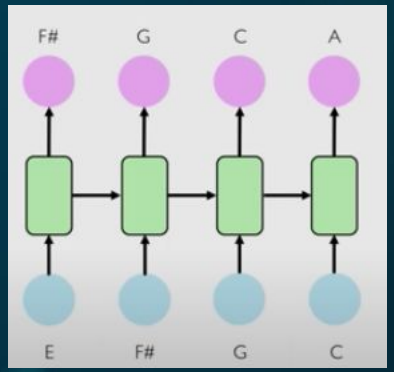
03

Where are LSTMs used?

Where are they used?



Language Modeling



Music Generation

Stock Price Predictions

WILL BE DISCUSSED AHEAD!!

LSTMs for Stock Prediction

04

LSTMs for Stock price prediction

Stock prices are perfect examples of time series data

Use time series analysis to gain useful patterns & stats from the data

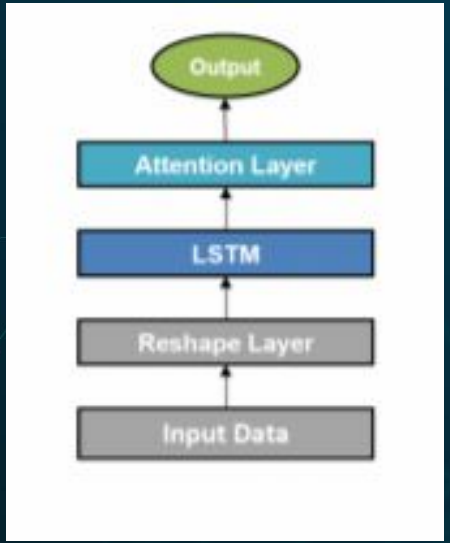
Time series forecasting to predict the future prices of stocks

LSTMs with Attention

$$e_t = tanh(W_a [x_1, x_2, ..., x_T] + b) \qquad \alpha_t = \frac{exp(e_t)}{\sum_{k=1}^{T} exp(e_k)}$$

Trainable Weight Matrix

Resulting weighting



Demonstration

GO TO COLAB!!

(Click)

THANK YOU!

References:

181 - multivariate time SERIES forecasting USING LSTM. (2020, December 08). Retrieved April 15, 2021, from https://www.youtube.com/watch?v=tepxdcepTbY

Bansal, S. (2018, March 27). Language modelling and text generation using LSTMs - deep learning for NLP. Retrieved April 15, 2021, from https://medium.com/@shivambansal36/language-modelling-text-generation-using-lstms-deep-learning-for-nlp-ed36b224b275

Express, A. (2020, November 28). Attention mechanism in deep learning: Attention model keras. Retrieved April 15, 2021, from https://www.analyticsvidhya.com/blog/2019/11/comprehensive-guide-attention-mechanism-deep-learning/

Karim, R. (2020, July 04). Animated RNN, LSTM and GRU. Retrieved April 15, 2021, from https://towardsdatascience.com/animated-rnn-lstm-and-gru-ef124d06cf45

Sklearn.preprocessing.StandardScaler¶. (n.d.). Retrieved April 15, 2021, from https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html

Zan560. (2020, February 14). MIT 6.S191 (2020): Recurrent neural networks. Retrieved April 15, 2021, from https://www.youtube.com/watch?v=SEnXr6v2ifU