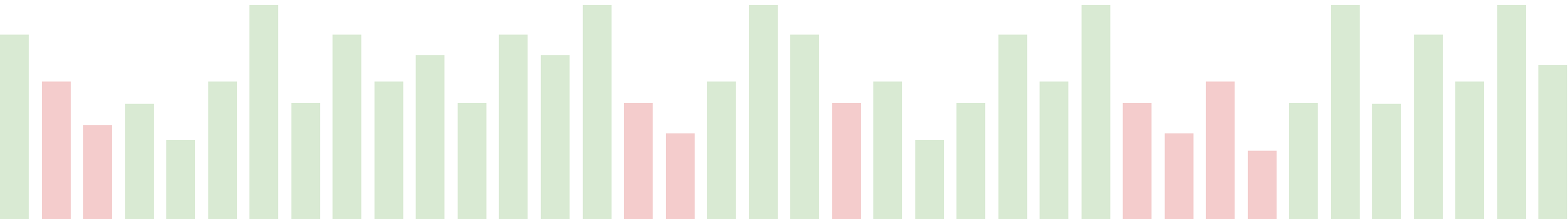


Predicting Stock Prices Using Deep Learning Sequential Models

Jiwon, Justin, Matt, Sean



Problem definition

Goal:

- **Train deep-learning regression models to predict stock prices from sequential price data**

Why?

- **Compare performance of traditional RNNs to LSTMs and Transformer models**
 - RNNs: vanishing gradient problem
 - LSTMs use gating mechanisms (input, forget, output)
 - Transformers use attention mechanism
 - **Can using multiple attention heads further improve performance?**

Dataset

Kaggle stock price dataset from Yahoo Finance

<https://www.kaggle.com/datasets/jacksoncrow/stock-market-dataset>

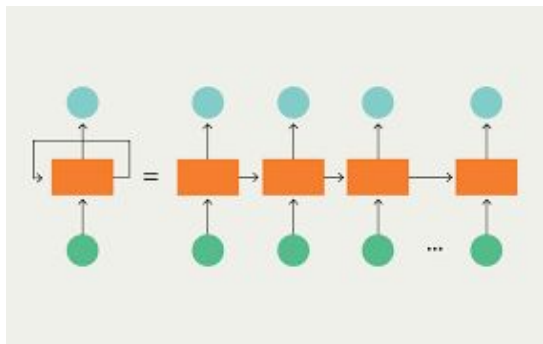
This dataset contains historical daily prices for all tickers currently trading on NASDAQ.

It contains prices from the creation of the symbol up to April 2020.

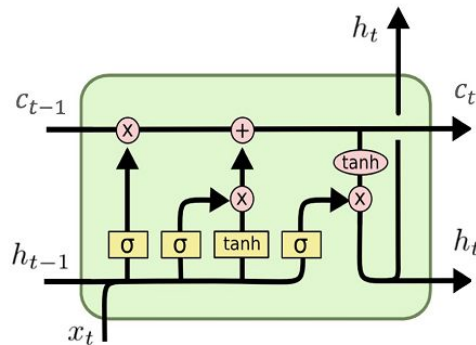
Date	# Open	# High	# Low	# Close	# Adj Close	# Volume
1999-03-10	51.125	51.15625	50.28125	51.0625	44.60090637207031	5232000
1999-03-11	51.4375	51.734375	50.3125	51.3125	44.81925964355469	9688600
1999-03-12	51.125	51.15625	49.65625	50.0625	43.72743606567383	8743600
1999-03-15	50.4375	51.5625	49.90625	51.5	44.9830322265625	6369000
1999-03-16	51.71875	52.15625	51.15625	51.9375	45.365169525146484	4905800

Algorithms

- Traditional RNN

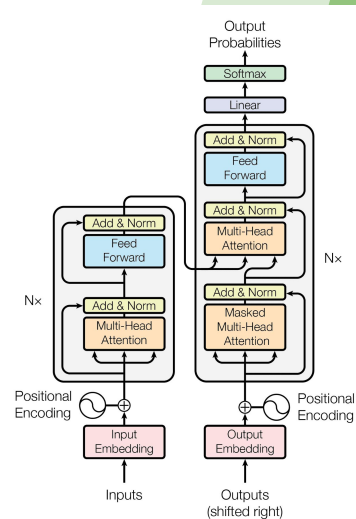


- LSTM



LSTM
(Long-Short Term Memory)

- Transformers



Experiments

Model Performance Measure:

- Test loss in predicting stock price

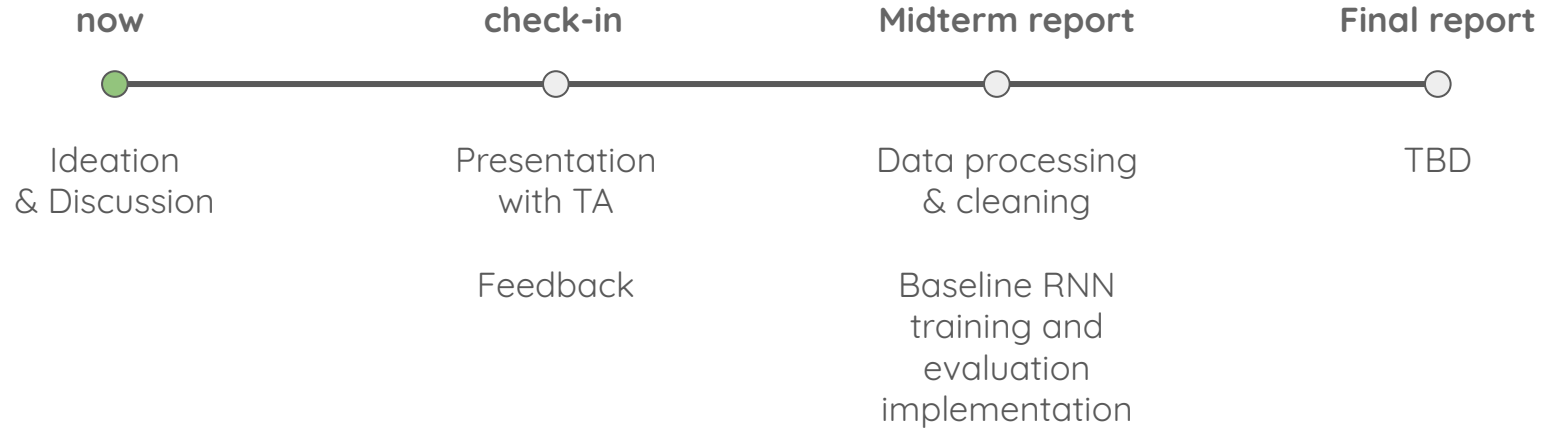
Analysis:

- Compare the relative performance of RNNs, LSTMs, and Transformers to each other

Scope:

- Open-ended:
 - Which stocks/ETFs should we conduct our experiment on?
 - ETFs
 - SPY (S&P 500)
 - QQQ (NASDAQ)
 - Some subset of individual stocks
 - Top 5 tech companies
 - How much training data to use?
 - All data since ticker symbol was created
 - Some shorter period of data

Milestone



Q&A

