Day 6

Agenda

- Stacking strategies
- Yahoo Finance data
- LSTM in theory
- Capstone 2

What is LSTM

- Long short-term memory (LSTM) is the artificial recurrent neural network (RNN)

- LSTM has "memory cells" that can remember information for long periods of time.

- It also has three gates that control the flow of information into and out of the memory cells: the input gate, the forget gate, and the output gate.

What's so special about it

- The key difference between LSTMs and other types of neural networks is the way that they deal with information over time.

- Traditional neural networks process information in a "feedforward" way

- LSTMs can process information in a "recurrent" way, meaning that they can take in input at one-time step and use it to influence their output at future time steps

4 main components

- The forget gate controls how much information from the previous time step is retained in the current time step

- The input gate controls how much new information from the current time step is added to the cell state

- The output gate controls how much information from the cell state is used to produce an output at the current time step

- The cell state is a vector that represents the "memory" of the LSTM network; it contains information from both the previous time step and the current time step.

GRU: Gated Recurring Unit

- "forget" gate, which allows them to selectively forget information from the previous timestep

- "update" gate, which allows them to control how much information from the current timestep is passed on to the next time step.

LSTM vs RNN (Pros)

- LSTM are much better at handling long-term dependencies

- LSTMs are much less susceptible to the vanishing gradient problem. This is because they use a different kind of activation function, known as an LSTM cell, which helps to preserve information over long sequences.

- LSTMs are very efficient at modeling complex sequential data. This is because they can learn high-level representations that capture the structure of the data.

- They are more complicated than traditional RNNs

LSTM vs RNN (Cons)

- They are more complicated than traditional RNNs

- They are not well-suited for online learning tasks, such as prediction or classification tasks where the input data is not a sequence

- LSTMs can be slow to train on large datasets. This is due to the fact that they must learn the parameters of the LSTM cells, which can be computationally intensive

- LSTMs may not be appropriate for all types of data. For example, they may not work well with highly nonlinear data or data with a lot of noise.

Limitation

- Inability to handle temporal dependencies that are longer than a few steps

- Limited context window size (2-4)

- LSTMs can be slow to train on large datasets. This is due to the fact that they must learn the parameters of the LSTM cells, which can be computationally intensive

- LSTMs may not be appropriate for all types of data. For example, they may not work well with highly nonlinear data or data with a lot of noise.