

Quiz 7: Introduction to Sequence Models

Introduction to Supervised Learning

*Required

1. Email address *

2. Please enter your name: *

Based on some information of the past T data points, we want to predict one of the three following categories for the next return of FB: category 0 if the return is $< -1\%$, category 1 if the return is between -1% and $+1\%$ and category 2 if the return is $> 1\%$

Facebook, Inc. Common Stock

NASDAQ: FB

✓ Following

230.34 USD **-1.57 (0.68%)** ↓

Jun 2, 09:31 EDT · Disclaimer

1 day 5 days 1 month 6 months YTD 1 year **5 years** Max



Here is the description of the training data:

- At each time step t , we have a feature vector x_t of size D representing the information we have gathered about the FB stock at time t .
- The whole sequence of feature vectors is: x_1, \dots, x_F
- The corresponding sequence of targets is: y_1, \dots, y_F (where each $y_i \in \{0, 1, 2\}$)
- We have the following sequences of features and the corresponding targets:

Sequences	Targets
x_1, \dots, x_T	y_{T+1}
x_2, \dots, x_{T+1}	y_{T+2}
\vdots	\vdots
x_{F-T}, \dots, x_{F-1}	y_F

Preprocessing

3. How many sequences do we have in our training data?

1 point

Mark only one oval.

☐ F

☐ F - T

☐ F - T - 1

4. Let N be the number of sequences. What is the shape of our training tensor data?

1 point

Mark only one oval.

- ☐ (N, D)
- ☐ (N, T, D)
- ☐ (N, T)

5. What is the shape of our training target data after the one-hot encoding of the targets?

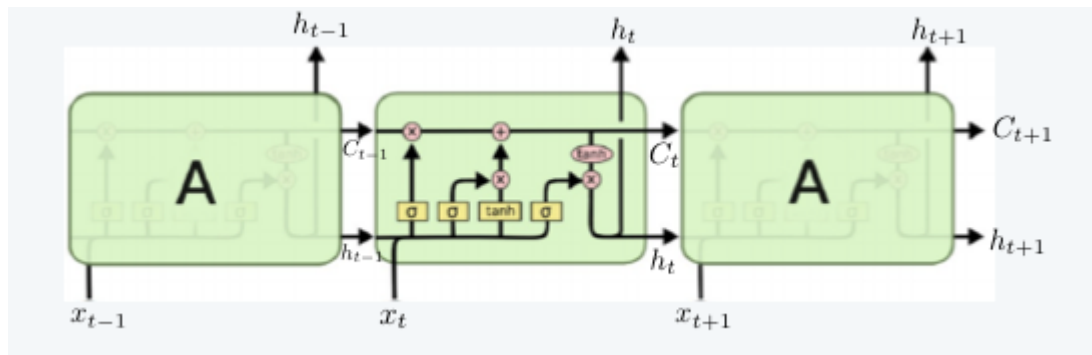
1 point

Mark only one oval.

- ☐ $(N, 3)$
- ☐ $(N,)$
- ☐ $(N, T, 3)$

The LSTM layer

We want to use an LSTM layer to process the sequences. Let d be the output vector size at each time step t .

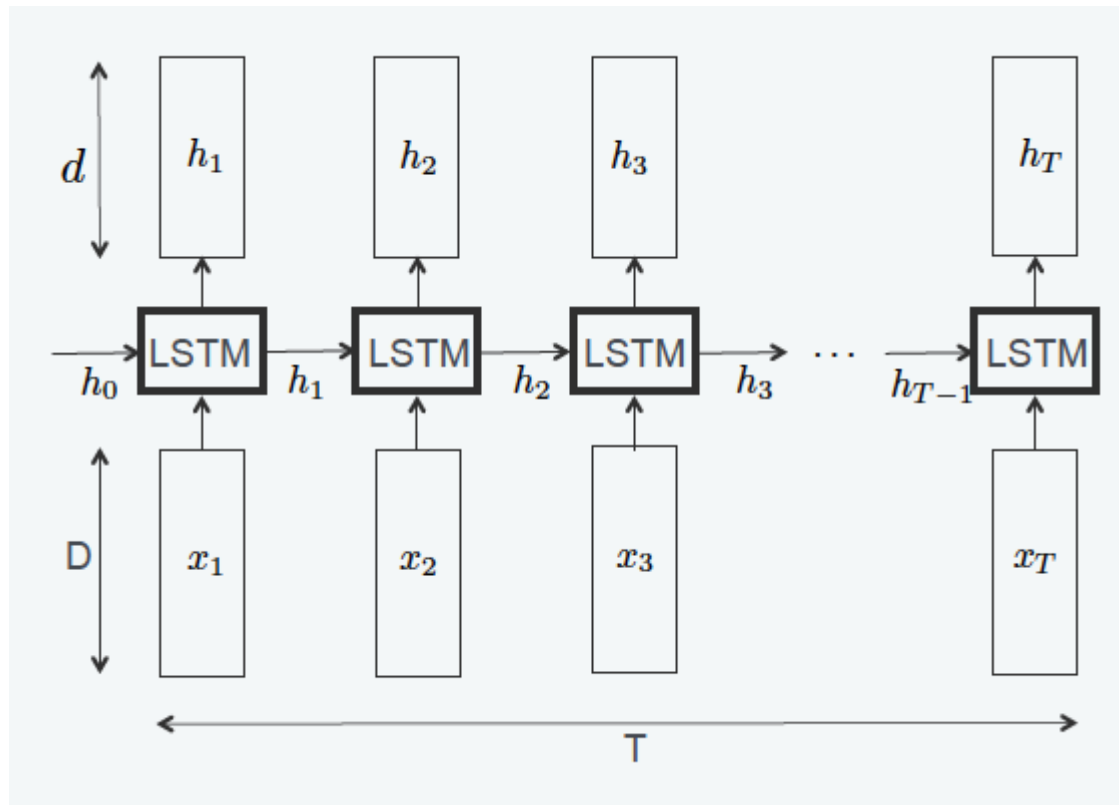


6. Why choosing an LSTM layer over a standard RNN layer? 1 point

7. How does the sigmoid activation function protect the cell state? 1 point

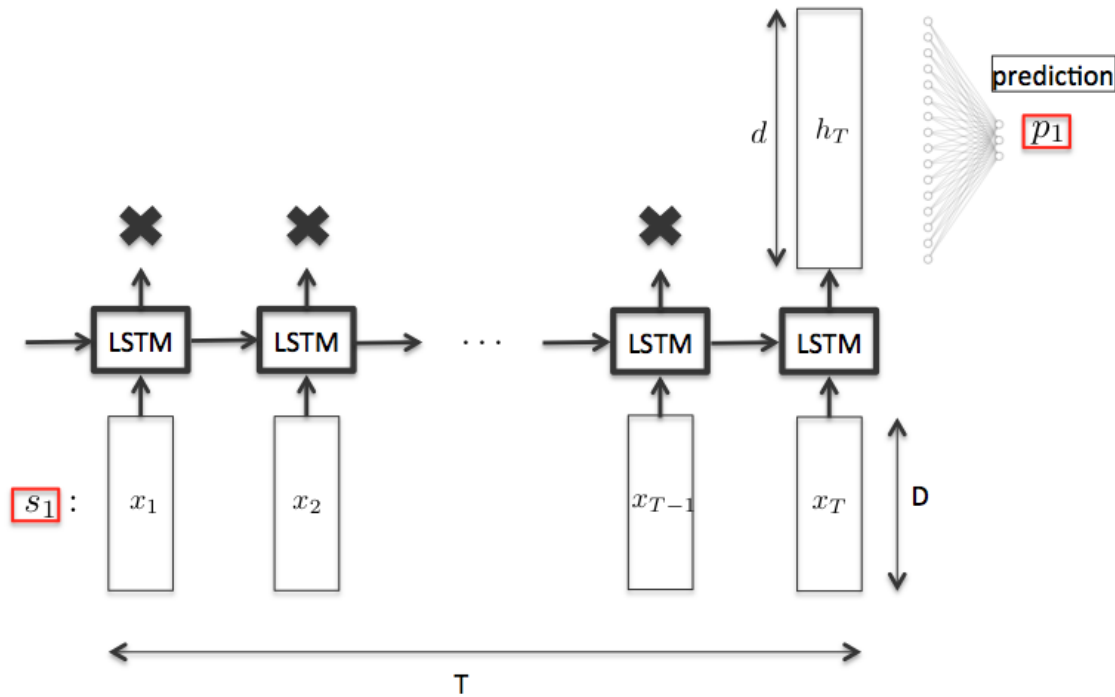
8. List all the parameters of the LSTM layer that should be learned using Gradient Descent. 1 point

9. For each sequence x_1, \dots, x_T , let h_1, \dots, h_T represent the output vectors. What information is represented by the vector h_t for each t in $\{1, \dots, T\}$? 1 point



The Supervised Model

Let's describe the forward propagation for the first sequence $s_1 = x_1, \dots, x_T$. The sequence is fed into an LSTM layer. We only keep the last output vector h_T of size d . The vector h_T is then fed into a Dense layer to output a vector of size 3.



10. Describe the evolution of the shape of data after each layer transformation: The LSTM layer and the Dense layer.

1 point

11. What activation function should be used in the Dense layer?

1 point

12. What loss function should be used?

1 point

Programming Session

13. Did you understand the problem?

Mark only one oval.

☐ Yes

☐ No

14. Do you have any questions about the Coursework?

Feel free to send us an email if you need more support.

15. Any comment?

This content is neither created nor endorsed by Google.

Google Forms