Knowledge Checks

Bookmark this page

DAT236x-M5-07

1/1 point (graded)

Which of the following aspects of a simple recurrent network contribute to the vanishing gradient problem?

- ○ Large number of nodes in the hidden layers
- ○ Large number of sequences
- ◉ Long sequences ✔
- ○ Short sequences

Submit You have used 1 of 1 attempt

✔ Correct (1/1 point)

DAT236x-M5-08

1/1 point (graded)

Which two techniques can be used to reduce overfitting?

- ☑ Use L1/L2 regularization in the optimizer

- [x] Provide more data to the model
- [ ] Increase the number of hidden dimensions
- [ ] Decrease dropout probability where dropout probability is defined as the chance of dropping a node

✔

Submit You have used 1 of 1 attempt

✔ Correct (1/1 point)

DAT236x-M5-04

0/1 point (graded)

You have a network with two dense layers. The first layer has 10000 weights and 100 bias, the second layer has 1000 weights with 10 bias. If you add a dropout layer in-between the two layers with a dropout probability of 0.5, what would be the total number of parameters in the final trained model?

- (•) 10000 + 100 + 500 + 5 ✖
- ( ) 5000 + 50 + 500 + 5
- ( ) 10000 + 100 + 1000 + 10
- ( ) 10000 + 100 + 500 + 10

Submit You have used 1 of 1 attempt

✖ Incorrect (0/1 point)

DAT236x-M5-02

0/1 point (graded)
True or False. Multi-layer perceptron can be thought of as a special case of a basic Recurrent unit.

- (•) True ✖
- ( ) False

Submit You have used 1 of 1 attempt

✖ Incorrect (0/1 point)

DAT236x-M5-03

1/1 point (graded)
Which two of the following options are recurrent units?

- [x] Long short term memory unit
- [ ] Dropout unit
- [x] Gated recurrent unit
- [ ] Minibatch unit

✔

Submit You have used 1 of 1 attempt

✔ Correct (1/1 point)

DAT236x-M5-06

1/1 point (graded)
Which two statements are true in the context of classic LSTM cell introduced in the lecture?

- [x] Both forget and update gates must have sigmoid activation
- [ ] Either forget or update gates may have tanh activation
- [ ] Forget gate and update gate may have tanh activation if the input gate has tanh activation
- [x] Input gate can have ReLU activation function regardless of other gates' activation units

✔

Submit You have used 1 of 1 attempt

✔ Correct (1/1 point)

DAT236x-M5-01

1/1 point (graded)
What are the applications where recurrent networks can be used?

- ○ Time series forecasting
- ○ Classifying words in a web query that names a commercial product
- ○ Classifying emails into SPAM in an inbox
- ○ Predicting how to steer an autonomous vehicle
- ◉ All of the above ✔

Submit You have used 1 of 1 attempt

✔ Correct (1/1 point)

DAT236x-M5-05

1/1 point (graded)
What are two key motivations behind LSTM cell?

- [ ] Better model the spatial structure in input data
- [ ] Increase model capacity with more number of parameters
- [x] Increase memory to handle recurrences with long sequence
- [x] Tackle vanishing gradient in deep recurrences

✔

Submit You have used 1 of 1 attempt

✔ Correct (1/1 point)

Discussion

Hide Discussion

Topic: Mod5-2 Knowledge Checks / Knowledge Checks

Add a Post

Show all posts ▼ by recent activity ▼

English