Congratulations! You passed!

Grade received 90% To pass 80% or higher

Go to next item

Practical aspects of deep learning

Latest Submission Grade 90%

1. If you have 10,000,000 examples, how would you split the train/dev/test set? **1 / 1 point**

- 60% train . 20% dev . 20% test
- 33% train . 33% dev . 33% test

Back
Practical aspects of deep learning
Graded Quiz • 30 min
Due Jan 25, 2:59 AM EST

Correct

2. The dev and test set should: **1 / 1 point**

- Come from the same distribution
- Come from different distributions
- Be identical to each other (same (x,y) pairs)
- **Have the same number of examples**

Correct

Back
Practical aspects of deep learning
Graded Quiz • 30 min
Due Jan 25, 2:59 AM EST

(Check all that apply.)

- [x] Add regularization

This should not be selected

- [ ] Make the Neural Network deeper
- [ ] Increase the number of units in each hidden layer
- [x] Get more training data

This should not be selected

- [ ] Get more test data

4. You are working on an automated check-out kiosk for a supermarket, and are building a classifier for apples, bananas and oranges. Suppose your classifier obtains a training set error of 0.5%, and a dev set error of 7%. Which **1 / 1 point**

Back
Practical aspects of deep learning
Graded Quiz • 30 min
Due Jan 25, 2:59 AM EST

Correct

- [ ] Decrease the regularization parameter lambda
- [x] Get more training data

Correct

- [ ] Use a bigger neural network

5. What is weight decay? **1 / 1 point**

- (•) A regularization technique (such as L2 regularization) that results in gradient descent shrinking the weights on every iteration.
- The process of gradually decreasing the learning rate during training.
- Gradual corruption of the weights in the neural network if it is trained on noisy data.

Correct

Back
Practical aspects of deep learning
Graded Quiz • 30 min
Due Jan 25, 2:59 AM EST

6. What happens when you increase the regularization hyperparameter lambda? **1 / 1 point**

- (•) Weights are pushed toward becoming smaller (closer to 0)
- Weights are pushed toward becoming bigger (further from 0)
- Doubling lambda should roughly result in doubling the weights
- Gradient descent taking bigger steps with each iteration (proportional to lambda)

Correct

Back
Practical aspects of deep learning
Graded Quiz • 30 min
Due Jan 25, 2:59 AM EST

- You apply dropout (randomly eliminating units) but keep the 1/keep_prob factor in the calculations used in training.
- You apply dropout (randomly eliminating units) and do not keep the 1/keep_prob factor in the calculations used in training
- (•) You do not apply dropout (do not randomly eliminate units) and do not keep the 1/keep_prob factor in the calculations used in training
- You do not apply dropout (do not randomly eliminate units), but keep the 1/keep_prob factor in the calculations used in training.

Correct

8. Increasing the parameter keep_prob from (say) 0.5 to 0.6 will likely cause the following: (Check the two that apply) **1 / 1 point**

- [ ] Increasing the regularization effect
- [x] Reducing the regularization effect

Back
Practical aspects of deep learning
Graded Quiz • 30 min
Due Jan 25, 2:59 AM EST

- [ ] Causing the neural network to end up with a higher training set error
- [x] Causing the neural network to end up with a lower training set error

Correct

9. Which of these techniques are useful for reducing variance (reducing overfitting)? (Check all that apply.) **1 / 1 point**

- [ ] Vanishing gradient

Back
Practical aspects of deep learning
Graded Quiz • 30 min
Due Jan 25, 2:59 AM EST

- [x] L2 regularization

Correct

- [x] Dropout

Correct

- [ ] Gradient Checking
- [x] Data augmentation

Correct

10. Why do we normalize the inputs x? **1 / 1 point**

- It makes it easier to visualize the data
- Normalization is another word for regularization--It helps to reduce variance
- It makes the parameter initialization faster
- (•) It makes the cost function faster to optimize