Practical aspects of deep learning

测验, 10 个问题

10/10 分 (100%)

恭喜！您通过了！

下一项

1 / 1 分

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

- 33% train . 33% dev . 33% test
- 60% train . 20% dev . 20% test
- 98% train . 1% dev . 1% test

正确

---

1 / 1 分

2。
The dev and test set should:

- Come from the same distribution

正确

-

Come from different distributions

Be identical to each other (same (x,y) pairs)

Have the same number of examples

---

1 / 1 分

3。
If your Neural Network model seems to have high variance, what of the following would be promising things to try?

Add regularization

正确

Make the Neural Network deeper

未选择的是正确的

Get more training data

正确

Get more test data

未选择的是正确的

Increase the number of units in each hidden layer

未选择的是正确的

---

1 / 1 分

10/10 分 (100%)

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 of the following are promising things to try to improve your classifier? (Check all that apply.)

- [x] Increase the regularization parameter lambda

正确

- [ ] Decrease the regularization parameter lambda

未选择的是正确的

- [x] Get more training data

正确

- [ ] Use a bigger neural network

未选择的是正确的

✔ 1 / 1 分

5。
What is weight decay?

- ○ The process of gradually decreasing the learning rate during training.
- ○ A technique to avoid vanishing gradient by imposing a ceiling on the values of the weights.
- ◉ A regularization technique (such as L2 regularization) that results in gradient descent shrinking the weights on every iteration.

正确

Gradual corruption of the weights in the neural network if it is trained on noisy data.

---

1 / 1 分

6。
What happens when you increase the regularization hyperparameter lambda?

Weights are pushed toward becoming smaller (closer to 0)

正确

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)

---

1 / 1 分

7。
With the inverted dropout technique, at test time:

You do not apply dropout (do not randomly eliminate 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

正确

You apply dropout (randomly eliminating units) but keep the 1/keep_prob factor in the calculations used in training.

---

1 / 1 分

8。
Increasing the parameter keep_prob from (say) 0.5 to 0.6 will likely cause the following: (Check the two that apply)

Increasing the regularization effect

未选择的是正确的

Reducing the regularization effect

正确

Causing the neural network to end up with a higher training set error

未选择的是正确的

Causing the neural network to end up with a lower training set error

正确

Practical aspects of deep learning

测验, 10 个问题

10/10 分 (100%)

1 / 1 分

9。

Which of these techniques are useful for reducing variance (reducing overfitting)? (Check all that apply.)

- [x] L2 regularization

正确

- [x] Dropout

正确

- [ ] Vanishing gradient

未选择的是正确的

- [ ] Gradient Checking

未选择的是正确的

- [x] Data augmentation

正确

- [ ] Exploding gradient

未选择的是正确的

- [ ] Xavier initialization

未选择的是正确的

Practical aspects of deep learning

测验, 10 个问题

10/10 分 (100%)

1 / 1 分

10。

Why do we normalize the inputs x ?

- It makes the parameter initialization faster
- It makes the cost function faster to optimize

正确

- It makes it easier to visualize the data
- Normalization is another word for regularization--It helps to reduce variance