Practical aspects of Deep Learning Graded Quiz • 50 min

Congratulations! You passed!

Grade Latest Submission

received 80% Grade 80%

Z Expand

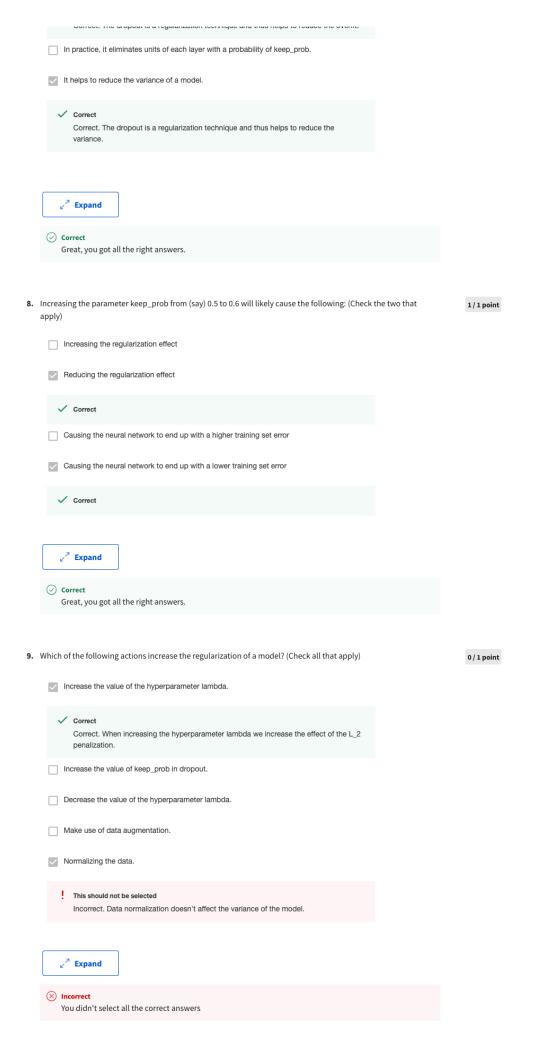
To pass 80% or higher

Go to next item

1.	If you have 10,000,000 examples, how would you split the train/dev/test set?	1/1 point
	33% train. 33% dev. 33% test	
	60% train. 20% dev. 20% test	
	98% train. 1% dev. 1% test	
	∠ ⁷ Expand	
	⊘ Correct	
2.	In a personal experiment, an M.L. student decides to not use a test set, only train-dev sets. In this case which of the following is true?	1/1 point
	He might be overfitting to the dev set.	
	He won't be able to measure the bias of the model.	
	Not having a test set is unacceptable under any circumstance.	
	He won't be able to measure the variance of the model.	
	_∠ [¬] Expand	
	✓ Correct Yes. Although not recommended, if a more accurate measure of the performance is not necessary it is ok to not use a test set. However, this might cause an overfit to the dev set.	
3.	If your Neural Network model seems to have high variance, what of the following would be promising things to try?	1/1 point
	Increase the number of units in each hidden layer	
	Get more test data	
	Get more training data	
	✓ Correct	
	Make the Neural Network deeper	
	Add regularization	
	✓ Correct	

	Great, you got all the right answers.	
	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 19% and a dev set error of 21%. Which of the following are promising things to try to improve your classifier? (Check all that apply, suppose the human error is approximately 0%)	0/1 point
	Get more training data.	
	Increase the regularization parameter lambda.	
	Use a bigger network.	
	∠ ⁷ Expand	
	No. First, the high bias problem must be addressed.	
5.	What is weight decay?	1/1 point
	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.	
	Gradual corruption of the weights in the neural network if it is trained on noisy data.	
	 A regularization technique (such as L2 regularization) that results in gradient descent shrinking the weights on every iteration. 	
	∠ [™] Expand	
	⊘ Correct	
6.	What happens when you increase the regularization hyperparameter lambda?	1/1 point
	Weights are pushed toward becoming bigger (further from 0)	
	Gradient descent taking bigger steps with each iteration (proportional to lambda)	
	Weights are pushed toward becoming smaller (closer to 0)	
	Doubling lambda should roughly result in doubling the weights	
	2 possed	
	∠ Expand	
7.	Which of the following are true about dropout?	1/1 point
	It helps to reduce the bias of a model.	
	In practice, it eliminates units of each layer with a probability of 1- keep_prob.	
	✓ Correct Correct: The dropout is a regularization technique and thus helps to reduce the overfit.	

⊘ Correct



10. Why do we normalize the inputs x ?	
It makes it easier to visualize the data	
It makes the cost function faster to optimize	
Normalization is another word for regularizationit helps to reduce variance	
It makes the parameter initialization faster	
_∠ ^ス Expand	
⊘ Correct	