10/10 points (100%)

Quiz, 10 questions

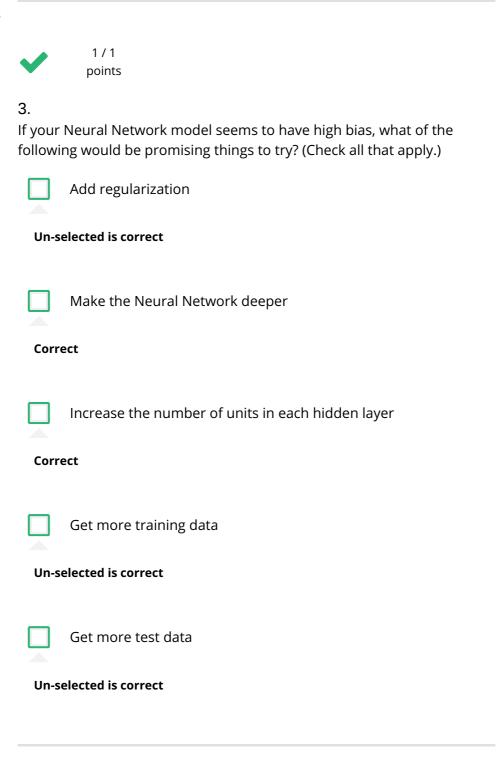
~	Congra	atulations! You passed!	Next Item
	~	1 / 1 points	
	1. If you set?	have 10,000,000 examples, how would you split the tra	in/dev/test
	Corr	98% train . 1% dev . 1% test	
		33% train . 33% dev . 33% test	
		60% train . 20% dev . 20% test	
	2 .	1 / 1 points	
	The de	ev and test set should:	
	Corr	Come from the same distribution	

Come from different distributions

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

10/10 points (100%)

Quiz, 10 questions





1/1 points

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.)

Increase the	rogula	rization	narameter	lamba	ا
increase the	regula	arization	Darameter	Idilibu	J

10/10 points (100%)

Quiz, 10 questions

COIT							
	Decrease the regularization parameter lambda						
Un-s	elected is correct						
	Get more training data						
Corre	ect						
	Use a bigger neural network						
Un-se	elected is correct						
5. Vhat is	1 / 1 points s weight decay?						
O .							
	The process of gradually decreasing the learning rate during training.						
0	A regularization technique (such as L2 regularization) that results in gradient descent shrinking the weights on every iteration.						
Corre	ect						
	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.						

/

1/1 points 6.

What happens when you increase the regularization hyperparameter $Practical\ asparets$ of $deep\ learning$

10/10 points (100%)

~ .			
()1117	1()	questio	ns
Ouiz.	10	uuesuu	HI.

0	Weights are pushed toward becoming smaller (closer to 0)
Corr	ect
	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 points

7.

With the inverted dropout technique, at test time:

- You apply dropout (randomly eliminating units) but 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

Correct

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), but keep the 1/keep_prob factor in the calculations used in training.



1/1 points

8.

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

	Increasing	tho	rogu	lariza	tion	offoc
l	Increasing	une	regu	IdHZd	uon	enec

10/10 points (100%)

Quiz, 10 questions

Un-s	elected is correct
Corre	Reducing the regularization effect
Un-se	Causing the neural network to end up with a higher training set error
	Causing the neural network to end up with a lower training set error
Corre	ect
	1 / 1 points of these techniques are useful for reducing variance (reducing ting)? (Check all that apply.)
	Vanishing gradient
Un-s	elected is correct
	L2 regularization
Corre	ect
Corre	Data augmentation
	Dropout

Correct

Practical aspects of deep learning

10/10 points (100%)

	_	_	
uiz, 10 questions		Kavier initialization	
	Un-sel	ected is correct	
		Gradient Checking	
	Un-sel	ected is correct	
		Exploding gradient	
	Un-sel	ected is correct	
			_
	~	1/1 points	
	10.		
	Why do \	ve normalize the inputs x ?	
		t makes the parameter initialization faster	
		Normalization is another word for regularizationIt helps to reduce variance	
		t makes it easier to visualize the data	
	O 1	t makes the cost function faster to optimize	
	Correc	t .	







10/10 points (100%)

Quiz, 10 questions