Practical aspects of deep learning

10/10 points (100%)

Quiz, 10 questions

~	Congra	atulations! You passed!	Next Ite
	1.	1/1 points	
	If you l	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	
	0	98% train . 1% dev . 1% test	
	Corr	ect	
	~	1 / 1 points	

2.

The dev and test set should:

Come from the same distribution

Correct



Come from different distributions

D . • 1	O	1 .	
Practical asp	nacts at:daa:	ວ ໄດລການ ເນດ	/\(\dots\)!\(\dots\)
i iacticai asi	TICES BE INTENDED	al to gardo otoer	(same (x.v) pairs)

10/10 points (100%)

Quiz,	10	questions
-------	----	-----------

Have the same number of examples



1/1 points

3.

If your Neural Network model seems to have high bias, what of the following would be promising things to try? (Check all that apply.)

Get more training data

Un-selected is correct

Add regularization

Un-selected is correct

Increase the number of units in each hidden layer

Correct

Get more test data

Un-selected is correct

Make the Neural Network deeper

Correct



1/1 points

4.

Practical as	You are working on an automated check-out kiosk for a supermarket, and pectodion of the points (100%) 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 regularization parameter lambda				
	Correct				
	Decrease the regularization parameter lambda				
	Un-selected is correct				
	Get more training data Correct				
	Use a bigger neural network Un-selected is correct				
	1/1 points				
	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.				
	Correct				

Practical as		radual corruption of the weights in the neural network if it is a refer to the property data. The property data	10/10 points (100%)
Quiz, 10 questions			
	~	1/1 points	
	6. What hap lambda?	pens when you increase the regularization hyperparameter	
	O W	eights are pushed toward becoming smaller (closer to 0)	
	Correct		
	_ w	eights are pushed toward becoming bigger (further from 0)	
	O D	oubling lambda should roughly result in doubling the weights	
		radient descent taking bigger steps with each iteration proportional to lambda)	
	~	1 / 1 points	
	7. With the i	nverted dropout technique, at test time:	
	d	ou do not apply dropout (do not randomly eliminate units) and o not keep the 1/keep_prob factor in the calculations used in aining	
	Correct		
		ou do not apply dropout (do not randomly eliminate units), but eep the 1/keep_prob factor in the calculations used in training.	
		ou apply dropout (randomly eliminating units) and do not keep ne 1/keep_prob factor in the calculations used in training	

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

${\begin{tabular}{l} 1/keep_prob factor in the calculations used in training. \\ Practical aspects of deep learning \end{tabular}}$

10/10 points (100%)

Quiz, 10 questions

~	1/1 points
	sing the parameter keep_prob from (say) 0.5 to 0.6 will likely cause the ng: (Check the two that apply)
	Increasing the regularization effect
Un-s	elected is correct
	Reducing the regularization effect
Corre	ect
	Causing the neural network to end up with a higher training set error
Un-s	elected is correct
	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-selected is correct

Practical aspects of deep learning

10/10 points (100%)

Fractical asp	ects of deep learning	10
Quiz, 10 questions	Dropout	
	Correct	
	Gradient Checking	
	Un-selected is correct	
	Data augmentation	
	Correct	
	Xavier initialization	
	Un-selected is correct	
	Exploding gradient	
	Un-selected is correct	
	L2 regularization	
	Correct	
,		
	1/1 points	
	10. Why do we normalize the inputs x ?	
	Normalization is another word for regularizationIt helps to	

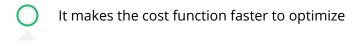
reduce variance

U It makes the parameter initialization faster

Practical aspects of deepelearning lize the data

10/10 points (100%)

Quiz, 10 questions



Correct

