0	Congratulations! You passed! Grade received 80% To pass 80% or higher	Go to next item	
Pı	ractical aspects of Deep Learning		
La	test Submission Grade 80%		
1.	If you have 20,000,000 examples, how would you split the train/dev/test set? Choose the best option. 99% train. 0.5% dev. 0.5% test.	1/1 point	
	90% train. 5% dev. 5% test.		
	O 60% train. 20% dev. 20% test.		
	Expand Constant		
	Correct Yes. Given the size of the dataset, 0.5% of the samples are enough to get a good estimate of how we model is doing.	ell the	
2.	The dev and test set should:	1/1 point	
	Be identical to each other (same (x,y) pairs)	1 / 1 point	
	O Have the same number of examples		
	Come from the same distribution Come from different distributions		
	∠ [⊅] Expand		
	⊘ Correct		
	If your Neural Network model seems to have high variance, what of the following would be promising thin try?	ngs to 1/1 point	
	☐ Get more test data ✓ Get more training data		
	✓ Correct		
	 Make the Neural Network deeper ✓ Add regularization 		
	✓ Correct		
	☐ Increase the number of units in each hidden layer		
	Expand		
	You are working on an automated check-out kiosk for a supermarket, and are building a classifier for appl bananas and oranges. Suppose your classifier obtains a training set error of 0.5%, and a deviset error of 7°		
	bananas and oranges. Suppose your classifier obtains a training set error of 0.5%, and a dev set error of 79 of the following are promising things to try to improve your classifier? (Check all that apply.) Increase the regularization parameter lambda		
	✓ Increase the regularization parameter lambda ✓ Correct		
	Decrease the regularization parameter lambda Cot more training date		
	✓ Get more training data ✓ Correct		
	Use a bigger neural network		
	∠ ⁷ Expand		
	 ✓ Correct Great, you got all the right answers. 		
5.	Which of the following are regularization techniques?	1/1 point	
	✓ Dropout.		
	Correct Correct. Using dropout layers is a regularization technique.		
	✓ Weight decay. ✓ Correct		
	Correct. Weight decay is a form of regularization. Gradient Checking.		
	Increase the number of layers of the network.		
	∠ [™] Expand		
6.	To reduce high variance, the regularization hyperparameter lambda must be increased. True/False?	1/1 point	
	● True		
	O False		
	∠ [™] Expand		
	Correct Correct. By increasing the regularization parameter the magnitude of the weight parameters is redu This helps reduce the variance.	ıced.	
	With the inverted dropout technique, at test time:		
í.	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	0 / 1 point	
	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) 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		
	the calculations used in training Expand		
	During training a deep neural network that uses the tanh activation function, the value of the gradients is practically zero. Which of the following is most likely to help the vanishing gradient problem?	1/1 point	
	O Increase the number of layers of the network. O Use a larger regularization parameter.		
	O Use Xavier initialization.		
	O Increase the number of cycles during the training.		
	∠ Expand ⊘ Correct		
	○ Correct Correct. A careful initialization can help reduce the vanishing gradient problem.		
9.	Which of the following actions increase the regularization of a model? (Check all that apply)	0 / 1 point	
	Increase the value of the hyperparameter lambda.		
	Correct Correct. When increasing the hyperparameter lambda we increase the effect of the L_2 penalization.		
	☐ Increase the value of keep_prob in dropout. ✓ Make use of data augmentation.		
	✓ Correct		
	Correct. Data augmentation has a way to generate "new" data at a relatively low cost. Thus making use of data augmentation can reduce the variance. Normalizing the data.		
	! This should not be selected Incorrect. Data normalization doesn't affect the variance of the model.		
	Incorrect. Data normalization doesn't affect the variance of the model. Decrease the value of the hyperparameter lambda.		
	∠ [™] Expand		
	Suppose that a model uses, as one feature, the total number of kilometers walked by a person during a ye		
	another feature is the height of the person in meters. What is the most likely effect of normalization of the data?		
	It won't have any positive or negative effects. It will make the data easier to visualize.		
	It will make the training faster.		
	O It will increase the variance of the model.		
	∠ Expand ⊘ Correct		
	Correct. Since the difference between the ranges of the features is very different, this will likely caus process of gradient descent to oscillate, making the optimization process longer.	se the	