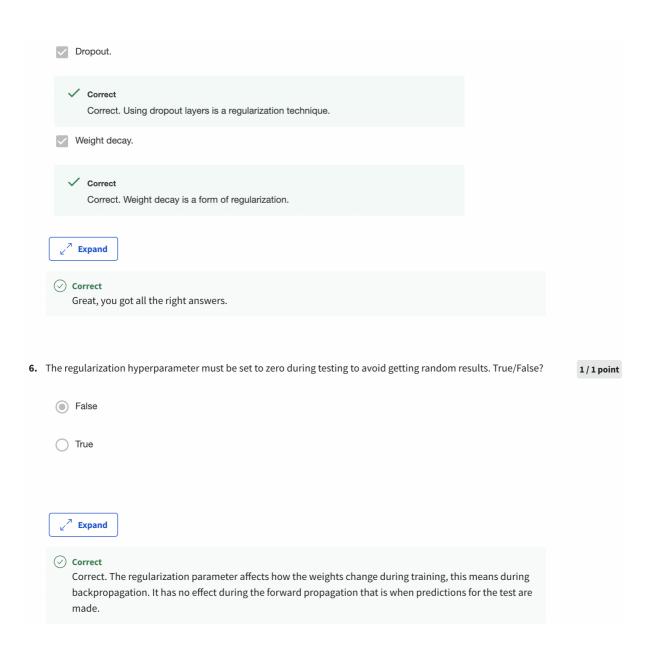
Practical aspects of Deep Learning

Congratulations! You passed! Grade Latest Submission To pass 80% or higher 1. If you have 20,000,000 examples, how would you split the train/dev/test set? Choose the best option. 1/1 point 99% train. 0.5% dev. 0.5% test. 60% train. 20% dev. 20% test. 90% train. 5% dev. 5% test. 2 Correct Yes. Given the size of the dataset, 0.5% of the samples are enough to get a good estimate of how well the model is doing.

2.	The dev and test set should:	1 / 1 point
	Come from different distributions	
	Be identical to each other (same (x,y) pairs)	
	Have the same number of examples	
	Come from the same distribution	
	$\mathcal{L}^{\mathcal{P}}$ Expand	
	⊘ Correct	
	If your Neural Network model seems to have high bias, what of the following would be promising things to try? (Check all that apply.)	1 / 1 point
	Add regularization	
	Get more training data	
	✓ Make the Neural Network deeper	
	✓ Correct	
	Increase the number of units in each hidden layer	
	✓ Correct	

	$\mathcal{L}^{\mathcal{P}}$ Expand	
	✓ CorrectGreat, you got all the right answers.	
4	You are working on an automated check-out kiosk for a supermarket and are building a classifier for apples,	1/1 waint
4.	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%)	1 / 1 point
	Get more training data.	
	Increase the regularization parameter lambda.	
	Use a bigger network.	
	∠ [¬] Expand	
	 Correct Yes. This can be helpful to reduce the bias of the model, and then we can start trying to reduce the high variance if this happens. 	
5.	Which of the following are regularization techniques?	1/1 point
	Increase the number of layers of the network.	
	Gradient Checking.	



7.	Which of the following are true about dropout?	1/1 point
	✓ It helps to reduce overfitting.	
	 Correct Correct. The dropout is a regularization technique and thus helps to reduce the overfit. 	
	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 probability that dropout doesn't eliminate a neuron is keep_prob. 	
	In practice, it eliminates units of each layer with a probability of keep_prob.	
	∠ [¬] Expand	
	✓ CorrectGreat, you got all the right answers.	
	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
	Increase the number of cycles during the training.	
	Increase the number of layers of the network.	
	Use a larger regularization parameter.	

Use Xavier initialization. ∠⁷ Expand **⊘** Correct Correct. A careful initialization can help reduce the vanishing gradient problem. **9.** Which of these techniques are useful for reducing variance (reducing overfitting)? (Check all that apply.) 1/1 point Dropout ✓ Correct L2 regularization ✓ Correct Xavier initialization Data augmentation ✓ Correct Vanishing gradient Gradient Checking Exploding gradient

∠ ⁷ Expand	
✓ CorrectGreat, you got all the right answers.	
10. Why do we normalize the inputs x ?	1 / 1 point
It makes the cost function faster to optimize	
It makes the parameter initialization faster	
It makes it easier to visualize the data	
Normalization is another word for regularizationIt helps to reduce variance	
∠ [™] Expand	
⊘ Correct	