Google Cloud

The Art and Science of Machine Learning

Quiz Question Answers

In each quiz, the questions and answers are randomized, so when you take the quiz, the order is different than what you see here.

Module 2: The Art of ML

Regularization

Question 1: Which regression model uses the L1 regularization technique?
A: Ridge Regression
Feedback: This answer is incorrect, please review the module again.

*B: Lasso Regression
Feedback: This answer is correct.

C: Both A and B
Feedback: This answer is incorrect, please review the module again.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 2: What would you use to solve the problem of overfitting?
A: Regularization technique
Feedback: This answer is partially correct, please review the module again.

B: Cross Validation
Feedback: This answer is partially correct, please review the module again.

C: Drop out

Feedback: This answer is partially correct, please review the module again.

*D: All of the above
Feedback: This answer is correct.

Question 3: Why would you use the square of the L2 norm?
A: To increase the calculation of derivatives
Feedback: This answer is incorrect, please review the module again.

B: To minimize the training error
Feedback: This answer is incorrect, please review the module again.

*C: To simplify the calculation of derivatives
Feedback: This answer is correct.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 4: Which regression model uses the L2 regularization technique?
*A: Ridge Regression
Feedback: This answer is correct.

B: Lasso Regression
Feedback: This answer is incorrect, please review the module again.

C: Both A & B
Feedback: This answer is incorrect, please review the module again.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 5: Why is regularization useful?
A: It makes models more linear
Feedback: This answer is incorrect, please review the module again.

B: It makes models more non-linear
Feedback: This answer is incorrect, please review the module again.

*C: It limits overfitting
Feedback: This answer is correct.

D: It makes models smaller
Feedback: This answer is incorrect, please review the module again.

Question 6: Which of the following is not true about the L1 regularization?
*A: L1 has one solution
Feedback: This answer is correct.

B: L1 has built in feature selection
Feedback: This answer is incorrect, please review the module again.

C: L1 has a sparse solution
Feedback: This answer is incorrect, please review the module again.

D: L1 penalizes the sum of absolute value of weights.
Feedback: This answer is incorrect, please review the module again.

Module 3: Hyperparameter Tuning

Hyperparameter Tuning

Question 1: Which of the following is an example of a hyper-parameter?
A: Learning rate
Feedback: This answer is partially correct, please review the module again.

B: Regularization rate
Feedback: This answer is partially correct, please review the module again.

C: Batch size
Feedback: This answer is partially correct, please review the module again.

*D: All the above
Feedback: This answer is correct.

Question 2: How would you use the Cloud AI Platform Training for hyperparameter tuning?

A: Make the parameter a command-line argument
Feedback: This answer is partially correct, please review the module again.

B: Make sure different iterations of training don't clobber each other
Feedback: This answer is partially correct, please review the module again.

C: Supply hyperparameters to the training job
Feedback: This answer is partially correct, please review the module again.

*D: All the above
Feedback: This answer is correct.

Question 3: How would you ensure the outputs of different trials don't clobber each other?

*A: By employing a good naming convention for output folders which make it unique.
Feedback: This answer is correct.

B: By employing a good naming convention for input folders.
Feedback: This answer is incorrect, please review the module again.

C: Both A and B
Feedback: This answer is incorrect, please review the module again.

D: None of the above.
Feedback: This answer is incorrect, please review the module again.

Question 4: How do you supply hyperparameters to the training job?

A: Create the yaml file
Feedback: This answer is partially correct, please review the module again.

B: Use command-line parameters in the Cloud AI Platform command to supply the path to the yaml file.
Feedback: This answer is partially correct, please review the module again.

*C: Both A & B
Feedback: This answer is correct.

Google Cloud

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 5: Model improvement depends on which of the following:
A: Batch size
Feedback: This answer is partially correct, please review the module again.

B: Learning rate
Feedback: This answer is partially correct, please review the module again.

*C: Both A & B
Feedback: This answer is correct.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 6: To improve the model, which of the following is recommended?
A: Low learning rate
Feedback: This answer is incorrect, please review the module again.

B: High learning rate
Feedback: This answer is incorrect, please review the module again.

C: Very High learning rate
Feedback: This answer is incorrect, please review the module again.

*D: Goldilocks learning rate
Feedback: This answer is correct.

Question 7
A hyper-parameter is a setting you set before training, and it doesn't change afterwards.

*A: True
Feedback: This answer is correct.

B: False
Feedback: This answer is incorrect, please review the module again.

Google Cloud

Module 4: A Pinch of Science

L1 Regularization

Question 1: What is L2 Regularization?
*A: Added to sum of the squared parameter weights terms to the last function
Feedback: This answer is correct.

B: Multiplied to sum of the squared parameter weights terms to the last function
Feedback: This answer is incorrect, please review the module again.

C: Subtracted to sum of the squared parameter weights terms to the last function
Feedback: This answer is incorrect, please review the module again.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 2: What is L1 regularization?
A: Added to sum of the squared parameter weights terms to the last function
Feedback: This answer is incorrect, please review the module again.

*B: Adds the sum of the absolute value the parameter weights to the last function
Feedback: This answer is correct.

C: Both A and B
Feedback: This answer is incorrect, please review the module again.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 3: What are the benefits of sparse model?
A: Reduction in storage and memory
Feedback: This answer is partially correct, please review the module again.

B: Increased training speed
Feedback: This answer is partially correct, please review the module again.

Google Cloud

C: Increase prediction speed
Feedback: This answer is partially correct, please review the module again.

*D: All of the above
Feedback: This answer is correct.

Question 4: What are the differences between L1 regularization and L2 regularization?
A: L2 regularization reduces the number of features, while L1 regularization reduces the amount of data analyzed at a time .
Feedback: This answer is incorrect, please review the module.

B: L1 regularization is more likely to result in overfitting than L2 regularization.
Feedback: This answer is incorrect, please review the module.

*C: L1 regularization tends to shrink coefficients to zero whereas L2 regularization tends to shrink coefficients evenly.
Feedback: This answer is correct.

D: L1 regularization results in a non-sparse solution and penalizes the sum of square weights, whereas L2 regularization results in a sparse solution and may have many solutions.
Feedback: This answer is incorrect, please review the module.

Question 5: Why is it recommended that a high-dimensional sparse vector weight drop to exactly 0?
A: This essentially removes the corresponding feature from the model
Feedback: This answer is partially correct, please review the module again.

B: Zeroing out features will save RAM
Feedback: This answer is partially correct, please review the module again.

C: May reduce noise in the model
Feedback: This answer is partially correct, please review the module again.

*D: All of the above
Feedback: This answer is correct.

Logistic Regression

Question 1: What are the constraints used in the sigmoid activation function?
A: Weighted sum
Feedback: This answer is partially correct, please review the module again.

B: W transpose X
Feedback: This answer is partially correct, please review the module again.

C: Plus B from a linear regression
Feedback: This answer is partially correct, please review the module again.

*D: All of the above
Feedback: This answer is correct.

Question 2: What does the output of a sigmoid represent?
A: Some function with a range of zero to one
Feedback: This answer is incorrect, please review the module again.

*B: Calibrated probability estimate
Feedback: This answer is correct.

C: Cumulative distribution function
Feedback: This answer is incorrect, please review the module again.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 3: A cross-entropy of zero indicates:
A: Your model's prediction was perfect.
Feedback: Correct! The cross-entropy compares the model's prediction with the actual results. The cross-entropy goes down as the prediction gets more accurate. A cross-entropy of zero indicates a perfect prediction.

B: Your model's prediction was bad (inaccurate).
Feedback: This answer is not correct. The cross-entropy compares the model's prediction with the actual results. A low cross-entropy value means a good

Google Cloud

predictive model. A high cross-entropy indicates a large difference and therefore a bad predictive model and more training is needed.

C: You do not have enough data to make predictions; you need more data.
Feedback: This answer is not correct. The cross-entropy compares the model's prediction with the actual results. The cross-entropy goes down as the prediction gets more accurate. A cross-entropy of zero indicates a perfect prediction.

D: The model's predictions were exactly 50% right and 50% wrong.
Feedback: This answer is incorrect, please review the module again. The cross-entropy compares the model's prediction with the actual results. The cross-entropy goes down as the prediction gets more accurate. A cross-entropy of zero indicates a perfect prediction.

Question 4: Why is regularization important in logistic regression?
*A: Avoids overfitting.
Feedback: This answer is correct.

B: Keeps training time down by regulating the time allowed.
Feedback: This answer is incorrect, please review the module again.

C: Finds errors in the algorithm.
Feedback: This answer is incorrect, please review the module again.

D: Encourages the use of large weights.
Feedback: This answer is incorrect, please review the module again.

Question 5: Why is it important for the gradient not to become zero?
A: Training will stop
Feedback: This answer is partially correct, please review the module again.

B: Leads to a vanishing gradient problem
Feedback: This answer is partially correct, please review the module again.

*C: Both A and B
Feedback: This answer is correct.

D: None of the above.
Feedback: This answer is incorrect, please review the module again.

Module 5: The Science of Neural Networks

Training Neural Networks

Question 1: Which of the following statements are correct about Batch Normalization?
A: Solves the problem called internal co-variance shift
Feedback: This answer is partially correct, please review the module again. .

B: Helps with exploding gradients
Feedback: This answer is partially correct, please review the module again.

C: Helps the intermediate inputs at each layer stay within a tighter range
Feedback: This answer is partially correct, please review the module again.

*D: All the above
Feedback: This answer is correct.

Question 2: Which of the following gives non-linearity to a neural network?
A: Stochastic Gradient Descent
Feedback: This answer is incorrect, please review the module again.

*B: Rectified Linear Unit
Feedback: This answer is correct.

C: Convolution function
Feedback: This answer is incorrect, please review the module again.

D: None of the above
Feedback: This answer is incorrect, please review the module again

Question 3: During the training process for deep networks, gradients can vanish, and each additional layer in your network can successively reduce signal vs noise. How can you fix this?
*A: Use non-saturating, nonlinear activation functions such as ReLUs.
Feedback: This answer is correct.

B: Sigmoid or tanh activation functions
Feedback: This answer is incorrect, please review the module again

C: Both A and B
Feedback: This answer is incorrect, please review the module again

D: None of the above.
Feedback: This answer is incorrect, please review the module again.

Question 4: How can you solve the problem where gradients explode?
A: Grading and clipping
Feedback: This answer is partially correct, please review the module again.

B: Batch normalization
Feedback: This answer is partially correct, please review the module again.

*C: Both A and B
Feedback: This answer is correct.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 5: In a neural network, which of the following techniques is used to deal with overfitting?
A: Dropout
Feedback: This answer is partially correct, please review the module again.

B: Regularization
Feedback: This answer is partially correct, please review the module again.

C: Batch Normalization
Feedback: This answer is partially correct, please review the module again.

*D: All of these
Feedback: This answer is correct.

Google Cloud

Multi-class Neural Networks

Question 1: Which of the following trainer modes should you use if you're not sure of the best parameters and want to use a parameter sweep?
A: Single Parameter
Feedback: This answer is incorrect, please review the module again.

*B: Parameter Range
Feedback: This answer is correct.

C: Both A and B
Feedback: This answer is incorrect, please review the module again.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 2: Which of the following normalization methods supports using feature normalization?
A: Binning normalizer
Feedback: This answer is partially correct, please review the module again.

B: Gaussian normalizer
Feedback: This answer is partially correct, please review the module again.

C: Min-max normalizer
Feedback: This answer is partially correct, please review the module again.

*D: All of the above
Feedback: This answer is correct.

Question 3: What is a Multiclass Neural Network?
A: Creates a multiclass logistic regression classification model
Feedback: This answer is incorrect, please review the module again.

B: Creates a multiclass classification model using the decision jungle algorithm
Feedback: This answer is incorrect, please review the module again.

Google Cloud

*C: Creates a multiclass classification model using a neural network algorithm
Feedback: This answer is correct.

D: Creates a multiclass classification model using the decision forest algorithm
Feedback: This answer is incorrect, please review the module again.

Question 4: Which tensorFlow function can you use to calculate the softmax?
*A: Sparse softmax cross entropy with logits
Feedback: This answer is correct.

B: Sparse categorical entropy
Feedback: This answer is incorrect, please review the module again.

C: Sparse categorical cross entropy
Feedback: This answer is incorrect, please review the module again.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 5: Which function you use If labels aren't mutually exclusive?
A: Sparse softmax cross entropy with logits
Feedback: This answer is incorrect, please review the module again.

B: Sigmoid cross entropy
Feedback: This answer is incorrect, please review the module again.

*C: Sigmoid cross entropy with logits
Feedback: This answer is correct.

D: Sparse softmax cross entropy
Feedback: This answer is incorrect, please review the module again.

Google Cloud

Module 6: Embeddings

Embeddings

Question 1: Which of the following is an efficient way to create embedding?
*A: Use the embedding column method in the tf.feature column and pass in the categorical column that you want to embed.
Feedback: This answer is correct.

B: Use the embedding column method in the tf.add column and pass in the feature column that you want to embed.
Feedback: This answer is incorrect, please review the module again.

C: Both A and B
Feedback: This answer is incorrect, please review the module again.

D: None of the above.
Feedback: This answer is incorrect, please review the module again.

Question 2: Which of the following is a disadvantage of dense representation?
A:It is inefficient for storage.
Feedback: This answer is partially correct, please review the module again.

B: It is inefficient for computation.
Feedback: This answer is partially correct, please review the module again.

*C: Both A and B
Feedback: This answer is correct.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 3: Which of the following is true about embedding?
A: An embedding is a weighted sum of the feature crossed values.
Feedback: This answer is partially correct, please review the module again.

Google Cloud

B: Embedding is a handy adapter that allows a network to incorporate spores or categorical data.
Feedback: This answer is partially correct, please review the module again.

C: The number of embeddings is the hyperparameter to your machine learning model.
Feedback: This answer is partially correct, please review the module again.

*D: All of the above
Feedback: This answer is correct

Question 4: Which of the following statements is true about the logit layer?
*A: For a logit layer, as opposed to a single logit, there is no guarantee that the total probability of all the digits equals one.
Feedback: This answer is correct.

B: For a logit layer, similar to a single logit, there is no guarantee that the total probability of all the digits will equal one.
Feedback: This answer is incorrect, please review the module again.

C: For a logit layer, as opposed to a single logit, there is no guarantee that the total probability of all the digits will equal zero.
Feedback: This answer is incorrect, please review the module again.

D: For a logit layer, similar to a single logit, there is no guarantee that the total probability of all the digits will equal zero.
Feedback: This answer is incorrect, please review the module again.

Question 5: Which of the following statement is incorrect?
A: The Categorical columns are represented by tensorflow as sparse tensors.
Feedback: This answer is incorrect, please review the module again.

B: Tensorflow can do math operations on sparse tensors without having to convert them into dense.
Feedback: This answer is incorrect, please review the module again.

Google Cloud

C: The more dimensions you have, the greater chance of overfitting.
Feedback: This answer is incorrect, please review the module again.

*D: None of the above
Feedback: This answer is correct.