Machine Learning

Assignment 5

Answer 1 A(Least Square Error)

Answer2 A (Linear Regression is Sensitive to outlier)

Answer3 B (Negative)

Answer4 C (both of them)

Answer5 A (high bias and high Variance)

Answer6 B (Predictive Model)

Answer7 B (Regularization)

Answer 8 A (cross Validation)

Answer 9 A(TPR and FPR)

Answer 10 A True(the better the model is at predicting 0 class as 0 and 1 class as 1 high theAUC. Distinguishing B/W patients with the disease and no disease.

Answer 11 B (Apply PCA to project high to choosethe learning rate.)

Answer 12 D (it does not make use of dependent variable)

Answer 13 Regularization means restricting a model to avoid overfitting by shrinking the coefficient estimates to zero. When a model suffers from overfitting, we should control the model's complexity. Technically, regularization avoids overfitting by adding a penalty to the model's loss function:

There are three commonly used regularization techniques to control the complexity of machine learning models:

1-L2 regularization

2-L1 regularization

3-E3-lastic Net

Formula = Regularization = Loss function + Penalty

Answer 14 The term ‘regularization’ refers to a set of techniques that regularizes learning from particular features for traditional algorithms or neurons in the case of neural network algorithms.It normalizes and moderates weights attached to a feature or a neuron so that algorithms do not rely on just a few features or neurons to predict the result. This technique helps to avoid the problem of overfitting to understand regularization, let’s consider a simple case of linear regression. Mathematically, linear regression is stated as below:

y = X0 + X1x1 + X2x2 + ….. + Xnxn

where y is the value to be predicted;

x1, x2, …., xn are features that decides the value of y;

w0 is the bias;

w1, w2, ….., wn are the weights attached to x1, x2, …., xn relatively.

Now to build a model that accurately predicts the y value, we need to optimize above mentioned bias and weights.To do so, we need to use a loss function and find optimized parameters using gradient descent algorithms and its variants.

To know more about building a machine learning application and the process, check out below blog:

There are three main regularization techniques:-

1-Ridge Regression (L2 Norm)

2-Lasso (L1 Norm)

3-Dropout

1) L2 Regularization

As seen above, the original loss function is modified by adding normalized weights. Here normalized weights are in the form of squares.

You may have noticed parameters λ along with normalized weights. λ is the parameter that needs to be tuned using a cross-validation dataset. When you use λ=0, it returns the residual sum of square as loss function which you chose initially. For a very high value of λ, loss will ignore core loss function and minimize weight’s square and will end up taking the parameters’ value as zero.Now the parameters are learned using a modified loss function. To minimize the above function, parameters need to be as small as possible. Thus, L2 norm prevents weights from rising too high.

2)Lasso Regression (L1 Regularization) Also called lasso regression and denoted as below:

This technique is different from ridge regression as it uses absolute weight values for normalization. λ is again a tuning parameter and behaves in the same as it does when using ridge regression.

3)Dropout

Dropout is a regularization technique used in neural networks. It prevents complex co-adaptations from other neurons. In neural nets, fully connected layers are more prone to overfit on training data. Using dropout, you can drop connections with 1-p probability for each of the specified layers. Where p is called keep probability parameter and which needs to be tuned. With dropout, you are left with a reduced network as dropped out neurons are left out during that training iteration.

Answer15 Linear regression most often uses mean-square error (MSE) to calculate the error of the model.

1.MSE is calculated by:measuring the distance of the observed y-values from the predicted y-values at each value of x;

2.squaring each of these distances;

3.calculating the mean of each of the squared distances.

4.Linear regression fits a line to the data by finding the regression coefficient that results in the smallest MSE.