MACHINE LEARNING ASSINGMENTS_WORKSHEET1

Question No.	Answers
1.	A
2.	A
3.	B
4.	B
5.	C
6.	B
7.	D
8.	D
9.	A
10.	B
11.	B
12.	A, B

Answer 13.

Regularization: It is a technique to avoid overfitting and underfitting problem. Let's first see the diagram below, we can say that there can be a situation when model starts predicting values nearby the target or the model starts predicting values away from the target or it can predict values far away into high variance zone. So if our model is too easy and takes very less parameters and predict nearby the target then it may have high bias and low variance (underfitting) and if there are more parameters / predicting away from the target it will have high variance and low bias (overfitting) or in other words we can also say that in underfitting, the model is not learning points/making self-assumption whereas in overfitting , the model is learning everything and unable to distinguish between noise and the actual data, so this is the case of variance-bias trade off or overfitting/underfitting problem. To overcome this situation, REGULARISATION comes into picture. In the regularization, the model will itself shrink /adjust coefficient's value to zero or near to zero. These adjustments are done by three methods either LASSO, RIDGE AND ELASTIC NET REGRESSION.

Answer 14.

The Following Algorithms are used for regularization are as follows:

1.LASSO OR L1 REGRESSION 2.RIDGE OR L2 REGRESSION 3.ELASTIC NET REGRESSION

LASSO (L1) : It stands for Least Absolute Shrinkage and Selection Operator. It tunes the co-efficient value of the variable to zero i.e some of the features are completely neglected for evaluating output. So it not only helps in reducing overfitting but also helps in feature selection. It is well suited for the models showing high levels of multicollinearity or variable selection/parameter elimination.

RIDGE(L2): It only reduces the co-efficient close to zero but not zero. It is also used to analyse any data that suffers from multicollinearity.

Under LASSO/RIGDE regression, Alpha parameter are used as penalty which is used to adjust the co-efficient value. So, by changing the values of alpha, we are controlling the penalty term. The higher the values of alpha, the bigger is the penalty and therefore the magnitude of coefficients is reduce and thereby improving the model.

ELASTIC NET: It is the combination of both LASSO AND RIDGE regression. It incorporates penalties from both L1 and L2 regularization. It allows us to tune the alpha parameter where $\alpha = 0$ corresponds to ridge and $\alpha = 1$ to lasso. If we set 0 for alpha, the penalty function reduces to the L1 (ridge) term and if we set alpha to 1 we get the L2 (lasso) term. Therefore, we can choose an alpha value between 0 and 1 to optimize the elastic net. Thus it will shrink some coefficients and set some to 0 for sparse selection and thereby improving the model.

Answers 15:

Before understanding error,we need to know about Linear Regression. Linear regression is a statistical tool to define the relationship between one or more independent variable and dependent variable. It describes the relationship between two or more variable.It is expressed as y=B0+B1x+e where

Y: is the dependent variable, the variable you want to predict

X:is the independent variable, the variable we are using to predict y

B0:is the intercept

B1:is the slope/co-efficient

e: the regression residual error

This error is defined as the how much is the difference between the actual value and the predicted value. It is represented as black line in the diagram below, which is the vertical distance from the data points to the regression curve. If the errors are large,it will impact the dependent variable. So in order to find the best fit line, we need to minimise sum of squared errors(difference between the actual and predicted value and square of that)