

Machine learning worksheet

Answer1- Least square error

Answer2 –none of these

Answer3-Negative

Answer4- Both of them

Answer5- Low bias and high variance

Answer6- Predictive Model

Answer7- Regularization

Answer8- SMOTE

Answer9- Sensitivity and specificity

Answer10- True

Answer11- Apply PCA to project high dimensional data

Answer12- A,B and C are correct

Answer 13-The word regularize means to make things regular or acceptable. This is exactly why we use it for. Regularizations are techniques used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting

- Answer14- Ridge Regression
- LASSO (Least Absolute Shrinkage and Selection Operator) Regression
- Elastic-Net Regression

Ridge Regression

Ridge regression is a method for analyzing data that suffer from multi-collinearity.

$$Loss = \sum_{i=1}^n (y_i - (w_i x_i + c))^2 + \lambda \sum_{i=1}^n w_i^2$$

- LASSO is a regression analysis method that performs both feature selection and regularization in order to enhance the prediction accuracy of the model.

$$Loss = \sum_{i=1}^n (y_i - (w_i x_i + c))^2 + \lambda \sum_{i=1}^n |w_i|$$

Elastic-Net Regression

Elastic-Net is a regularized regression method that linearly combines the L1 and L2 penalties of the LASSO and Ridge methods respectively.

$$Loss = \sum_{i=0}^n (y_i - (w_i x_i + c))^2 + \lambda_1 \sum_{i=0}^n |w_i| + \lambda_2 \sum_{i=0}^n w_i^2$$

Answer 15- Within a linear regression model tracking a stock's price over time, the error term is the difference between the expected price at a particular time and the price that was actually observed. The error term is known as the residual, disturbance, or remainder term. An 'error' is a variable in a statistical or mathematical model, which is created when the model does not fully represent the actual relationship between the independent variables and the dependent variables.