

Machine Learning

1. A. Least Square Error
2. A. Linear regression is sensitive to outliers
3. B. Negative
4. B. Correlation
5. C. Low bias and high variance
6. B. Predictive model
7. D. Regularization
8. D. SMOTE
9. A. TPR and FPR
10. B. False
11. A. Construction bag of words from an email
B. Apply PCA to project high dimensional data
12. A. We don't have to choose the learning rate
B. It becomes slow when the number of features is very large
13. Regularization is a technique used to prevent overfitting in machine learning models by adding a penalty to the loss function. This penalty discourages the model from becoming too complex and helps to generalize unseen data. Common forms of regularization include L1 and L2 regularization, which add penalties proportional to the absolute value and square of the coefficients.
14. a. **Lasso Regression (L1 regularization):** Adds a penalty equal to the absolute value of the magnitude of coefficients.
b. **Ridge Regression (L2 regularization):** Adds a penalty equal to the square of the magnitude of coefficients.

15. In a linear regression equation, the error term, also known as the residual, represents the difference between the observed values and the values predicted by the model.

The equation for a simple linear regression is $y = \beta_0 + \beta_1 x + \epsilon_1$

where y is the dependent variable,

β_0 is the intercept,

β_1 is the slope,

x is the independent variable,

ϵ_1 is the error term.