Machine learning

1. A
2. A
3. B
4. C
5. C
6. B
7. D
8. A
9. A
10. B
11. B
12. C

13. Regularization is based on the principle Occam's razor, which simply states that when came across two similar hypotheses choose the simpler one. This allows choosing the simpler concept than a complex one. Regularization refers to the modifications that can be made to the learning algorithm that helps to reduce this generalization error and not the training error. It reduces by ignoring the less important features. It also helps prevent overfitting, making the model more robust and decreasing the complexity of a model. The regularization techniques are Lasso Regression, Ridge Regression, and Elastic Net Regression. Regularization can be used for feature reduction.

14. a. Ridge Regression (L2 Regularization)

Its goal is to solve problems of data overfitting and when the data suffers from multicollinearity. A standard linear or polynomial regression model will fail in the case where there is high collinearity among the feature variables. Ridge Regression adds a small squared bias factor to the variables. Such a squared bias factor pulls the feature variable coefficients away from this rigidness, introducing a small amount of bias into the model but greatly reducing the variance.

b. Least Absolute Shrinkage and Selection Operator (LASSO, L1 Regularization)

In opposite to Ridge Regression it only penalizes high coefficients. Lasso has the effect of forcing some coefficient estimates to be exactly zero when hyper parameter θ is sufficiently large. Therefore, one can say that Lasso performs variable selection producing models much easier to interpret than those produced by Ridge Regression. Basically, it is reducing the variability and improving the accuracy of linear regression models.

c. Elastic Net

Combines characteristics of both lasso and ridge. Elastic Net reduces the impact of different features while not eliminating all of the features. Lasso will eliminate many features, and reduce overfitting in your linear model. Ridge will reduce the impact of features that are not important in predicting your y values. Elastic Net combines feature elimination from Lasso and feature coefficient reduction from the Ridge model to improve your model's predictions.

15. The error is the linear regression is represented by E. The error is used to show the variability in the value of **y** in the graph (linear regression curve) showing the relationship between the value of **x** and **y**, and if the error is not represented it shows that the value of x which is provided is enough to determine the value of y. In machine learning the term error means that the model is incomplete and tat the end result will differ while real world application.