(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 a email . b) apply pca to project high dimensional data –both options involve feature extraction.

(12) A) we don't have to choose the learning rate b) It becomes slow when number of feature is very large.

(13) Regularization is a technique used in machine learning to prevent overfitting by adding a penalty term to the loss function . It discourages overly complex models by penalizing large coefficients,thereby promoting simpler models that generalise better to unseen data.

(14) Common algorithms used for regularization:

1) lasso regression (L1 regularization): adds an absolute value penally.

2) ridge regression (L2 regularization): adds a squared value penalty.

3) elastic net : combines both l1 and l2 penalties.

(15) Error refers to difference between the observed values and predictive values generated by the model. This error can be attributed to: Bias, Variance,Irreducible error.