

Answers to the questions:

- 1) d) Collinearity
- 2) b) Random Forest
- 3) c) Decision Trees are prone to overfitting
- 4) c) Training data
- 5) c) Anomaly detection
- 6) c) Case based
- 7) d) Both a) and b)
- 8) c) Both a) and b)
- 9) b) 2
- 10) d) KMeans
- 11) c) Neither feature nor number of groups is known
- 12) b) SVG
- 13) b) Underfitting
- 14) a) Reinforcement learning
- 15) b) Mean squared error
- 16) a) Linear, binary
- 17) a) Supervised learning
- 18) c) both a) and b)
- 19) d) none of these
- 20) c) input attribute
- 21) a) SVM allows very low error in classification
- 22) b) Only 2
- 23) (A) $-(6/10 \log(6/10) + 4/10 \log(4/10))$
- 24) (A) weights are regularized with the l1 norm
- 25) b) Logistic regression and Gaussian discriminant analysis
- 26) d) Either 2 or 3

27) (B) increase by 5 pounds

28) (D) Minimize the squared distance from the points

29) (B) As the value of one attribute increases, the value of the second attribute also increases

30) b) Convolutional Neural Network