

## **Assignment (Answers)**

**Q1 ANS:** Collinearity

**Q2 ANS:** Random Forest

**Q3 ANS:** Decision Tree are prone to overfit

**Q4 ANS:** Training data

**Q5 ANS:** Anomaly detection

**Q6 ANS:** Case based

**Q7 ANS:** Both a and b

**Q8 ANS:** Both a and b

**Q9 ANS:** 2

**Q10 ANS:** KMeans

**Q11 ANS:** Neither feature nor number of groups is known

**Q12 ANS:** SVG

**Q13 ANS:** Underfitting

**Q14 ANS:** Reinforcement learning

**Q15 ANS:** Mean squared error

**Q16 ANS:** Linear, binary

**Q17 ANS:** supervised learning

**Q18 ANS:** euclidean distance

**Q19 ANS:** removing columns which have high variance in data

**Q20 ANS:** input attribute

**Q21 ANS:** SVM allows very low error in classification

**Q22 ANS:** Only 2

**Q23 ANS:**  $-(6/10 \log(6/10) + 4/10 \log(4/10))$

**Q24 ANS:** weights are regularized with the l1 norm

**Q25 ANS:** Support vector machine

**Q26 ANS:** Either 2 or 3

**Q27 ANS:** increase by 5 pound

**Q28 ANS:** Minimize the squared distance from the points

**Q29 ANS:** As the value of one attribute increases the value of the second attribute also increases

**Q30 ANS:** Convolutional Neural Network