Data Analytic - Tutorial IV
(Supervised Models)

1. Which of the following is a statistical-based classification method?
    a. Bayesian classifier
    b. Support vector machine
    c. k-Nearest neighbour classifier
    d. CART

2. You have a dataset of different flowers containing their petal lengths and color. Your model has to predict the type of flower for given petal lengths and color. Which type of task it is?
    a. Regression task
    b. Classification task
    c. Clustering task
    d. None of the above

3. Which of the following is NOT true in the case of Navie Bayes' classification technique?
    a. Predicts class membership probabilities.
    b. Based on Bayes' theorem on probability.
    c. All classes are mutually exclusive and exhaustive.
    d. All attributes are independent given a class.

4. Which of the following statements are limitations of a Naive Bayes Classifier
    a. It relies on categorical attributes
    b. Assumes all attributes are independent given a class.
    c. If data is less, it estimates poorly.
    d. None of the above

5. What is the number of parameters needed to represent a Naive Bayes classifier with n Boolean variables and a Boolean label?
   a. 2n+1
   b. N+1
   c. N
   d. 2n

6. Building an SVM for a given training data becomes
   a. an equality constraint convex optimization problem.
   b. an inequality constraint convex optimization problem.
   c. a top down, divide and conquer, recursive problem.
   d. a mapping problem.

7. Which of the following estimation strategies comes under the category of "Leave-one-out" cross-validation?
   a. Bootstrap method.
   b. Random subsampling.
   c. k-fold cross-validation
   d. N-fold cross-validation.

8. Model of which classifier(s) can be expressed in the form a graph structure?
   a. Bayesian classifier
   b. Support vector machine
   c. Decision tree
   d. Artificial neural network based classifier

9. Which of the following specifications is true for a perfect classifier?
   a. TPR=1, FPR=0, precision=1, F1_score=1
   b. TPR=0, FPR=1, precision=0, F1_score=1
   c. TPR=1, FPR=1, precision=0, F1_score=1
   d. TPR=0, FPR=0, precision=1, F1_score=0

10. Which of the following specifications is true for a perfect classifier?
    a. TPR=1, FPR=0, precision=1, F1_score=1
    b. TPR=0, FPR=1, precision=0, F1_score=1
    c. TPR=1, FPR=1, precision=0, F1_score=1
    d. TPR=0, FPR=0, precision=1, F1_score=0

11. Consider the student who has determined that he has 70% probability of acceptance and that only 70% of the accepted students will receive dormitory housing. Of the accepted students who receive dormitory housing, 80% will have at least one roommate. What is the probability of being accepted and receiving dormitory housing and having no roommates?

12. Suppose a voter poll is taken in three states. In state A, 50% of voters support the liberal candidate, in state B, 60% of the voters support the liberal candidate, and in state C, 35% of the voters support the liberal candidate. Of the total population of the three states, 40% live in state A, 25% live in state B, and 35% live in state C. Given that a voter supports the liberal candidate, what is the probability that she lives in state B?

13. Lets assume we have documents about movie reviews.We have two classes positive (+) and negative (-), and take the following miniature training and test documents simplified from actual movie reviews.

	Cat	Documents
Training	-	just plain boring
	-	entirely predictable and lacks energy
	-	no surprises and very few laughs
	+	very powerful
	+	the most fun film of the summer
Test	?	predictable with no fun

What are the predictive probabilities of positive (+) and negative (-) class of the Test documents

14. A Multi National Company XYZ decided to automate the process of assessing their employee performance and promoting them to the next level in the hierarchy. They have decided to build a Machine Learning classifier, which are pre-trained on four different features of the employees - Age, Marital Status, Gender, and Skin Color from the historical data, which will decide whether any particular employee will get promoted or not {0-Not promoted, 1- promoted}. All the feature values of the employees, ground truth, and the output by the ML models from the test set are given in the below table.

ID	Age	Marital Status	Gender	Skin Color	Truth	Prediction
1	35	Married	F	W	1	0
2	27	Unmarried	F	W	0	0
3	40	Married	M	B	1	1
4	24	Married	M	B	1	1
5	33	Unmarried	F	B	0	0
6	27	Unmarried	F	B	1	0
7	45	Married	M	W	0	1
8	26	Unmarried	M	W	0	1

Find out Accuracy, Precision, and Recall of the Model