



Recommender Systems
Assignment- Week 6
TYPE OF QUESTION: MCQ/MSQ

Number of questions: 10

Total marks: 20

QUESTION 1:

In a decision tree _____ variable appear as internal node whereas _____ variables appear as t

- a) **Independent, dependent**
- b) Dependent, Independent
- c) Implicit, explicit
- d) Explicit, Implicit

Correct Answer: a

Explanation: Refer to Week 6 lecture 1 slide 3,4

QUESTION 2:

The input variable with highest information gain is used to _____ the decision tree

- a) Cut
- b) Prune
- c) **Branch**
- d) join

Correct Answer: c

Explanation: Refer to Week 6 lecture 1 slide 6,9-11

QUESTION 3:

Which of the following is the naïve Bayes assumption?



- a) The class variable is the independent variable
- b) All the independent variables are conditionally dependent on each other
- c) All the independent variables are dependent only on the class variable only and there is no dependency among themselves.**
- d) The class variable depends only on selected independent variable

Correct Answer: c

Explanation: Refer to Week 6 lecture 2 slide 7

QUESTION 4:

If a data set has n input attributes $X=(X_1, \dots, X_n)$ and a class variable C . Which of the following formula is based on the class conditional probability assumption.

- a) $P(C|X_1, \dots, X_n) = P(X_1|C) P(X_2|C) \dots P(X_n|C)$
- b) $P(X|C) = P(X_1|C) P(X_2|C) \dots P(X_n|C)$**
- c) $P(C|X_1, \dots, X_n) = P(X_1, \dots, X_n|C)$
- d) All of the above

Correct Answer: b

Explanation: Refer to Week 6 lecture 2 slide 7

QUESTION 5:

Which of the following should be performed while handling text data with k-NN classifier ?

- a) Only the stop word removal process be carried out
- b) Only the stop word removal and lematization be carried out
- c) The numerical features are extracted from the text documents and used as input attribute**
- d) Text data can be used directly without any pre-processing and feature extraction

Correct Answer: c

Explanation: Refer to Week 6 lecture 3 slide 3

QUESTION 6:

How does a large value of K affects performance of kNN classifier?

- a) Cross-validation is not possible



- b) The underlying clustering algorithm does not work
- c) It becomes sensitive to noise points
- d) Neighborhood may include points from other classes**

Correct Answer: d

Explanation: Refer to Week 6 lecture 3 slide 9

QUESTION 7:

The accuracy of a Rule-Based Classifier is

- a) Fraction of total records that satisfy the antecedent
- b) Fraction of total records that satisfy the consequent
- c) Fraction of records that satisfy both the antecedent and consequent of a rule, over those that satisfy the antecedent**
- d) Fraction of records that satisfy the antecedent and consequent of a rule

Correct Answer: c

Explanation: Refer to Week 6 lecture 4 slide 4

QUESTION 8:

Which of the following is not a step of sequential covering algorithm/

- a) Learn rules considering one class at a time.
- b) Learn one rule at a time.
- c) Remove all the tuple from the training set, that is covered by the rule.
- d) Calculate the accuracy of a rule**

Correct Answer: d

Explanation: Refer to Week 6 lecture 4 slide 7

QUESTION 9:

In a content based news recommender system is to be designed. The class variable is binary representing likes/dislikes of a person. The input variables are the numerical features extracted from text documents in terms of decimal numbers, type of the article represented as a categorical variable. Which of the following model is most appropriate to build the recommender system.



- a) k-NN Classifier
- b) Naïve Bayes Classifier
- c) Binary logistics regression**
- d) Depend solely on rule-based classification

Correct Answer: c

Explanation: Refer to Week 6 lecture 5 slide 3

QUESTION 10:

How does a non-linear fit detected in regression modeling?

- a) By examining the standard error of coefficients
- b) Through residual plots**
- c) By computing the F-statistic
- d) By inspecting the correlation matrix of predictors

Correct Answer: b

Explanation: Refer to Week 6 lecture 5 slide 15