

# Prashanth.S (19MID0020)

## Lab Quiz

ATTEMPT SCORE

7 / 10

### 01 Multiple Choice

1 / 1

Calculate the following distance measures between the two vectors,  $v1 = [0, 1, 1, 0, 0, 0, 1]$   $v2 = [1, 0, 1, 0, 1, 0, 0]$  use Cosine Distance measure

a.  $1/3$

☒

b.  $2/3$

☐

c.  $1/2$

☐

d.  $3/4$

☐

### 02 Multiple Choice

1 / 1

Cluster is \_\_\_\_\_

b. Operations on a database to transform or simplify data in order to prepare it for a machine-learning algorithm

☐

d. None of these

☐

c. Symbolic representation of facts or ideas from which information can potentially be extracted

☐

a. Group of similar objects that differ significantly from other objects

☒

### 03 Multiple Choice

1 / 1

Assume, you want to cluster 7 observations into 3 clusters using K-Means clustering algorithm. After first iteration clusters, C1, C2, C3 has following observations: C1: {(2,2), (4,4), (6,6)} C2: {(0,4), (4,0)} C3: {(5,5), (9,9)} What will be the cluster centroids if you want to proceed for second iteration?

c. C1: (2,2), C2: (0,0), C3: (5,5)

☐

b. C1: (6,6), C2: (4,4), C3: (9,9)

☐

a. C1: (4,4), C2: (2,2), C3: (7,7)

☒

**04 Multiple Choice****0 / 1**

Suppose in a classification problem, you are using a decision tree and you use the Gini index as the criterion for the algorithm to select the feature for the root node. The feature with the \_\_\_\_ Gini index will be selected.

- ☐ (B) highest
- ☒ (C) least
- ☐ (D) None of these
- ☐ (A) maximum

**05 Multiple Choice****0 / 1**

Agglomerative clustering falls under which type of clustering method?

- ☐ a. partition
- ☒ b. hierarchical

**06 Multiple Choice****1 / 1**

Choose the data mining tool

- ☒ Weka
- ☐ C
- ☐ C++
- ☐ Flash

**07 Multiple Choice****1 / 1**

what is the formula to calculate Confidence(A -> B)?

- ☐ d.  $\text{Support}(A \ B) / \text{Support}(B)$
- ☒ a.  $\text{Support}(A \ B) / \text{Support}(A)$
- ☐ b.  $\text{Support}(A \ B) / \text{Support}(B)$
- ☐ c.  $\text{Support}(A \ B) / \text{Support}(A, B)$

08 Multiple Choice

0 / 1

What is the minimum no. of variables/ features required to perform clustering?

- ☐ d. 3
- ☐ a. 0
- ☐
- ☒ c. 2
- ☐ b. 1

09 Multiple Choice

1 / 1

What does FP growth algorithm do?

- ☐ b. It mines all frequent patterns through pruning rules with higher support
- ☐ a. It mines all frequent patterns through pruning rules with lesser support
- ☒ It mines all frequent patterns by constructing a FP tree

10 Multiple Choice

1 / 1

ARFF stands for

- ☒ Attribute Relationship File Format
- ☐ Attribute Related File Format
- ☐ Active Relational formatted File
- ☐ Association Rule based File Format