MCA, Semester III

Department of Computer Science

MCAO 302: DATA SCIENCE USING PYTHON

Class Test-1

Max marks: 15   **Time: 30 mins**

1. Linear regression is mainly used for Regression.   [True/False]   [1]
2. Decision Tree can only be used for classification.   [True/False]   [1]
3. Write formula for Jaccard Coefficients measure and compute the value of Jaccard coefficients between the samples x and y: **x** = 1 0 0 0 0 0 0 0 0 0; **y** = 0 0 0 0 0 0 1 0 0 1   [1+1]

4. Define Pearson's correlation coefficient between two discrete random variable x and y.   [2]

5. What are the basic properties of probability mass function.   [2]

6. Write the gradient descent algorithm for linear regression (assuming three features only). Explain all variables.   [3]
7. Write the decision rule for Naïve Bayes Classifier. Consider the following training data, create conditional probability table for Age attribute.   [2+2]

Age	Income	Student	Credit rating	Buys compter ?
≤ 30	high	no	fair	no
≤ 30	high	no	excellent	no
31…40	high	no	fair	yes
> 40	medium	no	fair	yes
> 40	low	yes	fair	yes
> 40	low	yes	excellent	no
31…40	low	yes	excellent	yes
≤ 30	medium	no	fair	no
≤ 30	low	yes	fair	yes
> 40	medium	yes	fair	yes
≤ 30	medium	yes	excellent	yes
31…40	medium	no	excellent	yes
31…40	high	yes	fair	yes
> 40	medium	no	excellent	no

MCA, Semester III

Department of Computer Science

MCAO 302: DATA SCIENCE USING PYTHON

Class Test-1

Max marks: 15   **Time: 20 mins**

1. Linear regression is mainly used for Classification.   [True/False]   [1]
2. Decision Tree can be used for classification and regression problems.   [True/False]   [1]
3. Write formula for Cosine Similarity measure and compute the value of Cosine Similarity between the samples x and y: **x** = 1 0 0 0 0 0 0 0 0 0; **y** = 0 0 0 0 0 0 1 0 0 1   [1+1]

4. Define Pearson's correlation coefficient between two discrete random variable x and y.   [2]

5. What are the basic properties of probability mass function.   [2]

6. Write the gradient descent algorithm for linear regression (assuming four features only). Explain all variables.   [3]

7. Write the decision rule for Naïve Bayes Classifier. Consider the following training data, create conditional probability table for Income attribute.   [2+2]

Age	Income	Student	Credit rating	Buys compter ?
≤ 30	high	no	fair	no
≤ 30	high	no	excellent	no
31...40	high	no	fair	yes
> 40	medium	no	fair	yes
> 40	low	yes	fair	yes
> 40	low	yes	excellent	no
31...40	low	yes	excellent	yes
≤ 30	medium	no	fair	no
≤ 30	low	yes	fair	yes
> 40	medium	yes	fair	yes
≤ 30	medium	yes	excellent	yes
31...40	medium	no	excellent	yes
31...40	high	yes	fair	yes
> 40	medium	no	excellent	no