

## All The best For Exams - Rejinpaul Team

**Anna University Exams Nov / Dec 2016 – Regulation 2013**  
**Rejinpaul.com Unique Important Questions – 7th Semester BE/BTECH**

### **IT6006 Data Analytics**

#### **UNIT I**

1. What are analysis and reporting? How are they different from each other?
2. List any 2 possible web data from which effective analysis can be carried out. Justify your answer with an example.
3. Explain the different components of Big Data.
4. Explain in detail about any 3 resampling techniques.
5. Explain about prediction and prediction error in terms of regression and classification.
6. Discuss in detail about the role of sampling distribution with a relevant example.

#### **UNIT II**

1. Explain Bayesian model in detail
2. Discuss Neural network in detail
3. Discuss Linear Systems Analysis and Nonlinear Dynamics Basics
4. Explain Fuzzy logic

#### **UNIT III**

1. Explain the architecture for processing stream data.
2. What is filtering streams? What does a bloom filter consist of? Explain filtering streams with examples.
3. Describe about any one Real time Analytics Platform (RTAP) Application.(6)
4. Explain in detail with an algorithm about counting the distinct elements in a stream.
5. Explain how to count ones in a window using DGIM algorithm. Also explain how query answering is done using DGIM algorithm. Use the stream 1 0 1 0 1 1 0 0 0 1 0 1 1 1 0 1 1 0 0 1 0 1 1 0.
6. Explain how to estimate moments using AMS algorithm. Suppose the stream is c, e, c, f, a, e, g, f, f, b, b, c, g, b, a, a, f, d, a, e. The length of the stream is  $n = 20$  and consider the random positions 3, 7, 14. Illustrate the working of the AMS algorithm for this stream.
7. Write short notes on finding most popular elements using decaying window.

#### **UNIT IV**

1. Explain Apriori algorithm. Explain how the efficiency of Apriori is improved. Consider the following baskets and apply the Apriori algorithm to find the frequent itemsets with the support threshold = 3. The baskets are {I1, I2, I5}, {I2, I4}, {I2, I3}, {I1, I2, I4}, {I1, I3}, {I2, I3}, {I1, I3}, {I1, I2, I3, I5}, {I1, I2, I3}.
2. A database has five transactions. Let min sup = 60% and min conf=80%. T100=Milk, Onion, Nuts, Kiwi, Egg, Yoghurt; T200=Dhal, Onion, Nuts, Kiwi, Egg, Yoghurt; T300=Milk, Apple, Kiwi, Egg; T400=Milk, Curd, Kiwi, Yoghurt; T500=Curd, Onion, Kiwi, Ice cream, Egg Find all frequent itemsets using Apriori method.
3. What are limited pass algorithms? Explain in brief about each of them with necessary diagrams.
4. Explain in detail about any 2 approaches for counting Frequent Items in a Stream.
5. Describe in detail about Hierarchical methods and efficiency of Hierarchical Clustering.
6. Describe K-means clustering with an example (OR) Explain about K-means algorithm and how BFR algorithm can be used to cluster large scale data. .
7. Describe CLIQUE Algorithm with an example.

#### **UNIT V**

1. Discuss in detail Sharding
2. Explain (a) Hadoop – components and DFS (b)No SQL
3. Explain Mapreduce
4. Explain (a)Interaction techniques (b)Visualisation techniques

## **All The best For Exams - Rejinpaul Team**

Questions Are Expected for University Exams This May or may Not Be Asked for Exams

Please do not Copy (or) Republish This Questions, Students if You Find the Same Questions in Other Sources, Kindly report us to [rejinpaulteam@gmail.com](mailto:rejinpaulteam@gmail.com)