# Image result for GLOBAL ACADEMY OF TECHNOLOGY LOGO GLOBAL ACADEMY OF TECHNOLOGY

**Department of Computer Science and Engineering**

# QUESTION BANK

**Subject Name:** Big Data Analytics

**Subject Code:** 18CS72

**Faculty:** Dr. MJK/RMS/HPC

|  |
| --- |
| 1. What is PIG in Big Data. Explain the applications and features of PIG 2. Explain the difference between: 3. PIG and Map Reduce 4. PIG and SQL 5. Pig and HIVE 6. Explain the PIG Architecture for scripts dataflow and processing. 7. Explain the Pig Data Model with suitable examples. 8. Explain linear and non-linear relationship with suitable graphs 9. Explain Analysis of Variance (ANOVA) and correlation indicators of linear relationship. 10. Write a note on the following w.rt.t Regression Analysis 11. Simple Linear Regression 12. Multiple Regression 13. Write a note on 14. Finding Similar Sets 15. Collaborative Filtering (CF) as similar sets or users 16. Explain the different measures for finding similar items or users 17. Briefly explain Frequent Itemset Mining and Association Rule. Explain the different applications of Association Rules. 18. Explain with an example Apriori algorithm to evaluate candidate key. 19. Explain the five phases in text mining process pipeline 20. Briefly explain the areas and applications of Text Mining. List the different challenges in the area of text mining 21. What is Web Mining. With a neat diagram explain the taxonomy of Web Mining 22. What is Web Usage Mining? With a neat figure explain the different phases of Web Usage Mining. 23. What is Web Content Mining? Explain the techniques involved in Web Content Mining. 24. Describe HITS algorithm to iterate and compute the hubs and authorities. 25. What is a PageRank. Explain the Computation of page rank and page iteration of PageRank algorithm 26. Using the in-degrees as conferring authority 27. Using the relative authority of the parents over linked children   19.Explain the parameters in social graph network topological analysis using centralities and PageRank. |
|  |
|  |

**Module 1**

**Ai techniques, problem and domains**

**Tic tac toe (all three versions)**

**Water jug problem**

**Bfs and dfs algorithm**

**Answering questions based on database**

**Hill climbing :**

**Simple hill climbing**

**Simulated annealing algorithm**

**A\* and AO\* algorithm**