|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Roll No |  |  |  |  |  |  |  |  |  |  |  |  |

****

**Presidency University**

**Bengaluru**

**Ph. D Course Work End Term Examinations – JAN-FEB 2025**

**Date**: 04-02-2025

**Time**: 9.30 AM TO 12.30 PM

**Max Marks**: 100

**Weightage**: 50%

**Semester**:

**Course Code**: CSE6005

**Course Name**: Intelligent Information Retrieval

**Department:** SOCSE

**Instructions:**

1. *Read the all questions carefully and answer accordingly.*
2. *Do not write any matter on the question paper other than roll number.*

**PART A**

**Answer all the Questions. Each question carries 10 marks. (6Qx 10M= 60M)**

|  |  |
| --- | --- |
| **1.** | Mention the difference between Information retrieval and data retrieval. |
| **2.** | Illustrate about search engine architecture with suitable sketch. |
| **3.** | Discuss about practical issues on the web and visualization in search interfaces. |
| **4.** | Discuss about neural network model of information retrieval. List the features of neural network model compare to other models? |
| **5.** | Let us consider a table which have features late payments, income/expenses and classifier. Predict classifier for I When late payments=6, income/expenses=1.15, J When late payments=22, income/expenses=0.45, and K When late payments=15, income/expenses=1.2 using nearest neighbors.   |  |  |  |  | | --- | --- | --- | --- | | Object | Late Payment | Income/Expenses | Classifier | | A | 0 | 1.2 | G | | B | 25 | 0.4 | P | | C | 5 | 0.7 | G | | D | 20 | 0.8 | P | | E | 30 | 0.85 | P | | F | 11 | 1.2 | G | | G | 7 | 1.15 | G | | H | 15 | 0.8 | P | |
| **6.** | Outline the drawback of content-based filtering. |

**PART B**

**Answer all the Questions. Each question carries 20 marks. (2Qx 20M= 40M)**

|  |  |
| --- | --- |
| **7.** | A “collection” consists of the following “documents”:  d1: Shipment of gold damaged in a fire.  d2: Delivery of silver arrived in a silver truck.  d3: Shipment of gold arrived in a truck.  Suppose that we use the term frequency as term weights and query weights. The following document indexing rules are also used:  • Stop words were not ignored  • Text was tokenized and lowercased  • No stemming was used  • Terms were sorted alphabetically  Use Latent Semantic Indexing (LSI) to rank these documents for the query “gold silver truck”. |
| **8.** | Consider the following table showing the purchase details of 5 customers in a supermarket:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Product | Customer1 | Customer2 | Customer3 | Customer4 | Customer5 | | Ice Cream | 1 | 0 | 1 | 1 | 0 | | Milk | 0 | 1 | 1 | 0 | 0 | | Cake | 1 | 0 | 1 | 1 | 1 | | Biscuit | 0 | 1 | 1 | 0 | 1 |  1. Find the customers who purchased only one product. 2. Identify the customers who purchased either Ice Cream or Cake, but not both. |