|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Reg. No. |  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |  |
| --- | --- | --- |
|  | Continuous Assessment Test- II [CAT - II] | |
| Year | : IV |
| Semester | : 08 |
| Branch / Section | : B.E. CSE - A |
| Subject Code | : CS8080 |
| Subject Name | : Information Retrieval Techniques |
| QP CODE | : |

*[Regulations 2021]*

**Date: Time: 90 Minutes Marks: 50**

**Answer ALL Questions**

**Part A [7 x 2 = 14 Marks]**

|  |  |  |  |
| --- | --- | --- | --- |
| Q.NO | **QUESTION** | **BT LEVEL** | **CO** |
| 3.1 | Define Text Classification and briefly explain its significance in natural language processing. | A1 | **CO3** |
| 3.2 | Compare and contrast Unsupervised and Supervised Algorithms in the context of text classification. | A2 | **CO3** |
| 3.3 | Describe the basic principles of Clustering and its application in text classification. | A1 | **CO3** |
| 3.4 | Brief the concept of Accuracy and Error in text classification metrics. | A1 | **CO3** |
| 3.5 | Discuss the importance of Feature Selection in text classification and provide two methods for achieving it. | A2 | **CO3** |
| 3.6 | Describe the construction and working principle of Inverted Indexes in text retrieval systems. | B1 | **CO3** |
| 3.7 | Differentiate between Sequential Searching and Multi-dimensional Indexing techniques in information retrieval. | A2 | **CO3** |

**Part B [3x12=36 Marks]**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 3.8 | a | Examine the concept of locally weighted learning in k-NN classifiers and its implications for handling non-uniform data distributions and varying class densities. Discuss how adaptive weighting schemes can improve the model's ability to discriminate between different text categories. Provide examples to illustrate your analysis. | B2 | **CO3** |
|  | **[OR]** | |  |  |
|  | b | Critically assess the importance of Feature Selection or Dimensionality Reduction techniques in improving the efficiency and effectiveness of text classification models. Provide real-world examples to support your argument. | B2 | **CO3** |
| 3.9 | a | Examine the role of support vectors in SVM classifiers and their significance in defining the decision boundary between classes. Discuss how the number and distribution of support vectors affect the model's complexity, training time, and prediction accuracy. Provide examples to illustrate your analysis. | B2 | **CO3** |
|  | **[OR]** | |  |  |
|  | b | Predicting a class label using naïve Bayesian classification. The training data set is given below. The data tuples are described by the attributes Owns Home?, Married, Gender and Employed. The class label attribute Risk Class has three distinct values. Let C1  corresponds to the class A, and C2 corresponds to the class B and C3 corresponds to the class C.  The tuple is to classify is,  X = (Owns Home = Yes, Married = No, Gender = Female, Employed = Yes)   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Owns** | **Married** | **Gender** | **Employed** | **Class** | | Yes | Yes | Male | Yes | B | | No | No | Female | Yes | A | | Yes | Yes | Female | Yes | C | | Yes | No | Male | No | B | | No | Yes | Female | Yes | C | | No | No | Female | Yes | A | | No | No | Male | No | B | | Yes | No | Female | Yes | A | | No | Yes | Female | Yes | C | | Yes | Yes | Female | Yes | C | | B2 | **CO3** |
| 3.10 | a | Develop a comprehensive plan for building a text search engine using Inverted Indexes. Include details on indexing process, query processing, and relevance ranking algorithms for efficient retrieval of relevant documents. | B2 | **CO3** |
|  | **[OR]** | |  |  |
|  | b | Let us take the last 10 days weather dataset with attributes weather, temperature, wind, and humidity. The outcome variable will be play? cricket or not. Apply the decision tree induction algorithm to build the decision tree.   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Day** | **Weather** | **Temperature** | **Humidity** | **Wind** | **Play?** | | 1 | Sunny | Hot | High | Weak | No | | 2 | Cloudy | Hot | High | Weak | Yes | | 3 | Sunny | Mild | Normal | Strong | Yes | | 4 | Cloudy | Mild | High | Strong | Yes | | 5 | Rainy | Mild | High | Strong | No | | 6 | Rainy | Cool | Normal | Strong | No | | 7 | Rainy | Mild | High | Weak | Yes | | 8 | Sunny | Hot | High | Strong | No | | 9 | Cloudy | Hot | Normal | Weak | Yes | | 10 | Rainy | Mild | High | Strong | No | | B2 | **CO3** |

COURSE-FACULTY COURSE-COORDINATOR COURSE-EXPERT HOD