|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Register Number |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |



**SRM Institute of Science and Technology**

**Set - C**

**College of Engineering and Technology**

**School of Computing**

SRM Nagar, Kattankulathur – 603203, Chengalpattu District, Tamil Nadu

**Academic Year: 2023-24 (EVEN)**

**Test: CLA-T3** **Date: 07.05.2024**

**Course Code & Title: 18CSC305J – Artificial Intelligence**  **Duration:** 50 minutes

**Year & Sem: III Year / VI Sem** **Max. Marks:** 25

**Course Articulation Matrix:**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| S. No | Course Outcome | PO1 | PO2 | PO3 | PO4 | PO5 | PO6 | PO7 | PO8 | PO9 | PO10 | PO11 | PO12 |
| 1 | CO1 | 3 | 2 | 3 | - | - | - | - | - | - | - | - | - |
| 2 | CO2 | 3 | 2 | 3 | - | - | - | - | - | - | - | - | - |
| 3 | CO3 | 2 | 3 | 3 | - | - | - | - | - | - | - | - | - |
| 4 | CO4 | 2 | 3 | 2 | - | - | - | - | - | - | - | - | - |
| 5 | CO5 | 2 | 3 | 3 | 2 | - | - | - | - | - | - | - | - |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Answer all the Questions**  **(10 Marks)** | | | | | | |
| Q. No | Question | Marks | BL | CO | PO | PI Code |
| 1. | ACROSS 3. The data that represents quantities can take on any value within a certain range **(CONTINUOUS)** 5. It is an area of computer science that deals with simulating human problem-solving and mental processing in a computerized model **(COGNITIVEMODELING)**  DOWN 1. Action Description Language allows **(NEGATIVE)** literals but not by STRIPS 2. Goal is to update each of the weights in the neural network **(BACKWARDPASS)** 4. It helps the model to shift the activation function towards the positive or negative side **(BIAS)** | 5 | 1 | 4 | 1 | 1.1.1 |
| 2. | Refer the following Document Term matrix     1. Identify the most common and the rarest terms of the corpus.   **t5 and t6**   1. **What is the term frequency of a term that is used a maximum number of times in that document? t3** | 2 | 2 | 5 | 1 | 1.2.1 |
| 3. | **Identify the below application with the right term Information Extraction or Information Retrieval**   1. **Consider a program that can identify all person names or locations from raw text. – Information extraction** 2. **Google search engine – Information retrieval** 3. **Recalling the memory regarding the submission of the mini-project report – Information retrieval** | 3 | 2 | 5 | 1 | 1.1.1 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Answer any Three Questions**  **(3 x 5 = 15 Marks)** | | | | | | |
| Q. No | Question | Marks | BL | CO | PO | PI Code |
| 1. | Do the following queries using STRIPS language for the below diagram     1. Initial and Goal State representation (1 mark)   **Initial State:**  on(A,table)^on(B,table)^clear(A)^clear(B)^empty\_hand  **Final State:**  on(A,table)^on(B,A)^clear(B)^empty\_hand   1. List of Actions and Their Representation (4 marks)   Unstack (B, A)  Stack (B, A)  Lift (B)  Place (B)  **Unstack (B, A):**  **Pre-condition: empty\_hand^on(B,A)^clear(B)^on(A,table)**  **Post-condition: Delete-list: empty\_hand^on(B,A)^clear(B)**  **Add-list: holding(B)^clear(A)^on(A,table)**  **Stack (B, A):**  **Pre-condition: holding(B)^clear(A)^on(A,table)**  **Post-condition: Delete-list: holding(B)^clear(A)**  **Add-list: on(B,A)^clear(B)^empty\_hand^on(A,   table)**  **Lift (B):**  **Pre-condition: on(B,table)^clear(B)^empty\_hand^on(A,table)^clear(A)**  **Post-condition: Delete-list: on(B,table)^clear(B)^empty\_hand**  **Add-list: holding(B)^on(A,table)^clear(A)**  **Place (B):**  **Pre-condition: holding(B)^on(A,table)^clear(A)**  **Post-condition: Delete-list: holding(B)**  **Add-list: on(B,table)^clear(B)^empty\_hand^on(A,table)^clear(A)** | 5 | 3 | 4 | 2 | 2.1.3 |
| 2. | Suppose we have a perceptron having weights corresponding to the three inputs that have the following values:  w1 = 3; w2 = −2; and w3 = 1  and the activation of the unit is given by the step function:  φ(v) = 1 if v≥0 otherwise 0  Calculate the output value y of the given perceptron for each of the following input patterns:    **Ans:**  To calculate the output value y for each of the given patterns we have to follow below two steps:  a) Calculate the weighted sum: v = Σi (wi xi)= w1 ·x1 +w2 ·x2 +w3 ·x3  b) Apply the activation function to v.  The calculations for each input pattern are:  **P1 :** v = 2·1−4·0+1·0=2, (2>0),  y=φ(2)=1  **P2 :** v = 2·0−4·1+1·1=−3, (−3<0),  y=φ(−3)=0  **P3 :** v = 2·1−4·0+1·1=3, (3>0),  y=φ(3)=1  **P4 :** v = 2·1−4·1+1·1=−1, (−1<0),  y=φ(−1)=0 | 5 | 3 | 4 | 2 | 2.1.3 |
| 3. | i. Suppose you are trying to cluster documents using a Bag of Words method. Typically, words like or, and, a, an, if, of, is and so on are not having great features. How do you make sure you are leveraging the more informative words better during the feature Engineering? (3 marks)  Words like if, of, … are called stop words. Typical pre-processing in standard NLP pipeline involves identifying and removing stop-words (except in some cases where context/ word adjacency information is important). Common techniques to remove stop words include :   1. TF-IDF – Term frequency inverse document frequency 2. Leveraging manually curated stop word lists and eliminating these words 3. We also reduce words to their roots – this is called lemmatization. This ensures a word that occurs several time receives more weightage even if the occurrences have different endings example: teach, teaching, teaches..   ii. What would be the scenario if you do not convert all characters to a single case (either lower or upper) during the pre-processing step of an NLP algorithm? (2 marks)  When all words are not converted  to a single case, the vocabulary size will increase drastically as words like Up/up or Fast/fast or This/this will be treated differently which isn’t a correct behaviour for the NLP task.  Sparsity is higher when building the language model since the cat is  treated differently from The cat. Suppose we are building an ngram model, we might end up with many ngrams in test set that never appeared in the training set. | 5 | 4 | 5 | 2 | 2.3.2 |
| 4. | i. Draw the architecture of a typical Convolutional Neural Network and label them. (2 marks)  Basic CNN Architecture: Explaining 5 Layers of Convolutional Neural Network  | upGrad blog  ii. Give a brief description on Convolutional Neural Network and Multi-layer Perceptron (3 marks)  **Convolutional Neural Network (CNN)**: the incumbent, current favorite of computer vision algorithms, winner of multiple ImageNet competitions. Can account for *local connectivity* (each filter is panned around the entire image according to certain size and stride, allows the filter to find and match patterns no matter where the pattern is located in a given image). The weights are smaller, and shared — less wasteful, easier to train than MLP. More effective too. Can also go deeper. Layers are sparsely connected rather than fully connected. It takes matrices as well as vectors as inputs. The layers are *sparsely connected* or partially connected rather than *fully connected*. Every node does not connect to every other node.  **Multilayer Perceptron (MLP)**: used to apply in computer vision, now succeeded by Convolutional Neural Network (CNN). MLP is now deemed insufficient for modern advanced computer vision tasks. Has the characteristic of *fully connected layers*, where each perceptron is connected with every other perceptron. Disadvantage is that the number of total parameters can grow to very high (number of perceptron in layer 1 multiplied by # of p in layer 2 multiplied by # of p in layer 3…). This is inefficient because there is redundancy in such high dimensions. Another disadvantage is that it disregards spatial information. It takes flattened vectors as inputs. A light weight MLP (2–3 layers) can easily achieve high accuracy with MNIST dataset. | 5 | 2 | 5 | 2 | 2.1.3 |