**MARWADI UNIVERSITY**



**Faculty of Technology**

**Information and Communication Technology B.TECH**

**SEM: III MU FINAL EXAM (open book) DECEMBER:2022**

**Subject: - Machine Learning (01CT0519) Date:-16/12/2022**

**Total Marks:-30 Time: 90 min**

**Instructions:**

1. **All Questions are Compulsory.**
2. **Make suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**

# Attempt any four (CO4,CO5) 12

* + 1. You are designing a deep learning system to detect driver fatigue in cars. It is crucial that your model detects fatigue, to prevent any accidents. Design the system that achieves the given problem statement. Also state, which of the following is the most appropriate evaluation metric : Accuracy, Precision,

Recall, Loss Value. Explain your choice

* + 1. You wanted to solve a classification task. You first train your network on 20 samples. Training converges, but the training loss is very high. You then decide to train this network on 10,000 examples. Is your approach to fixing the problem correct ? If yes, explain the most likely result of training with 10,000 examples. If not, give a solution to this problem.
    2. You are solving the binary classification task of classifying images as cat vs. non-cat. You design a CNN with a single output neuron. Let the output of this neuron be z. The final output of your network, y' is given by: y' = σ(ReLU(z)). You classify all inputs with a final value y' ≥ 0.5 as cat images. What problem are you going to encounter?
    3. After visually inspecting the dataset for classifying the historical places, you realize that the training set only contains pictures taken during the day, whereas the dev set only has pictures taken at night. Explain what is the issue and how you would correct it.
    4. Google search is the most favourable option for all of us. Before some decades, google faced the issue of searching the questions having more than 32 words, asked by the user. What could be the reason behind that? Give two techniques for solving this issue.

Q -2 Attempt any one (CO2) 6

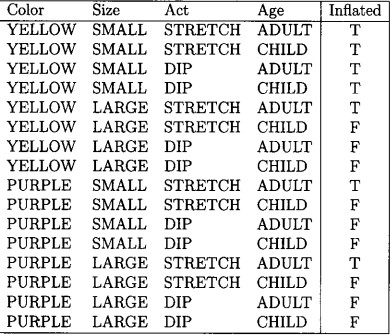
1. Show the final result of hierarchical clustering with complete link by drawing dendogram

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F |
| A | 0 |  |  |  |  |  |
| B | 0.12 | 0 |  |  |  |  |
| C | 0.51 | 0.25 | 0 |  |  |  |
| D | 0.84 | 0.16 | 0.14 | 0 |  |  |
| E | 0.28 | 0.77 | 0.70 | 0.45 | 0 |  |
| F | 0.34 | 0.61 | 0.93 | 0.20 | 0.67 | 0 |

1. Compare DBSCAN, Hierarchical, K-means in detail

# Attempt any two (CO1, CO3) 12

* + 1. Calculate the parameters of Naïve Bayes classifier for predicting inflated; and the training set error for the given data set



* + 1. Identify the first splitting attribute for decision tree by using ID3 algorithm With the following dataset.

|  |  |  |  |
| --- | --- | --- | --- |
| Major | Experience | Tie | Hired ? |
| B.Tech -ICT | Programming | Pretty | NO |
| B.Tech –ICT | Programming | Pretty | NO |
| B.Tech –ICT | Management | Pretty | YES |
| B.Tech -ICT | Management | Ugly | YES |
| Business Administration | Programming | Pretty | YES |
| Business Administration | Programming | Ugly | YES |
| Business Administration | Management | Pretty | NO |
| Business Administration | Management | Pretty | NO |

* + 1. Consider table which consist of land prices based on the area and proximity to the city center. You need to build linear regression based on this data. Use mean square error function
       1. Write hypothesis function, and equations for updating the different parameters to this estimators
       2. Use the training data (first four rows of table) to calculate parameter after two iterations gradient descent (use initial parameters with 0.5, learning rate 0.001).
       3. With prediction model found as per above steps predict the value for the given input in line 5 (table)

|  |  |  |  |
| --- | --- | --- | --- |
| Sr. No. | Area of land (meter square) | Distance from city  center (in Km) | Price |
| 1 | 25 | 12 | 13 |
| 2 | 30 | 10 | 18 |
| 3 | 21 | 25 | 9 |
| 4 | 28 | 2 | 17 |
| 5 | 35 | 15 | ? |