**20EC705/JO**

**Hall Ticket Number:**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **IV/IV B.Tech (Regular) DEGREE EXAMINATION** | | | | |
| **January, 2024** | | **Electronics & Communication Engineering** | | |
| **Seventh Semester** | **Deep Learning** | | | |
| **Time:** Three Hours | | | **Maximum:** 70 Marks | |
| ***Answer question 1 compulsory.*** | | | | **(14X1 = 14Marks)** |
| ***Answer one question from each unit.*** | | | | **(4X14=56 Marks)** |
|  | | | |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | CO | BL | M |
| 1 | a) | What is neural network? List the basic components of ANN. | CO1 | L1 | 1M |
|  | b) | Draw the structure of Mc-culloch pitts model. | CO1 | L1 | 1M |
|  | c) | Define the curse of dimensionality. | CO1 | L1 | 1M |
|  | d) | List different types of unsupervised learning models. | CO1 | L1 | 1M |
|  | e) | What type of activation function can be used to get the probability values corresponding to model predictions? | CO2 | L1 | 1M |
|  | f) | What is the relationship between width and depth of neural network models with that of model performance metrics bias and variance? | CO2 | L1 | 1M |
|  | g) | Suppose it is required to learn hierarchical representations of input distributions to reduce the variance in learning, which one do you do, increasing the width OR increasing the depth of the model? | CO2 | L1 | 1M |
|  | h) | What is the purpose of activation functions? | CO2 | L1 | 1M |
|  | i) | Draw a two flattened layer neural network diagram illustrating the weight sharing feature of convolution neural networks. | CO3 | L1 | 1M |
|  | j) | What is the purpose of using padding and stride in convolution layers? | CO3 | L1 | 1M |
|  | k) | Define the term receptive field in convolution layers. | CO3 | L1 | 1M |
|  | l) | Which will capture the long range dependencies better: RNN or LSTM? | CO4 | L1 | 1M |
|  | m) | What is the impact of Forgetting gate in LSTM? | CO4 | L1 | 1M |
|  | n) | In what way vanishing gradient problem effect the learning rate? Learning becomes slow or fastens? How? | CO4 | L1 | 1M |
| Unit-I | | | | | |
| 2 | a) | Define machine learning and distinguish between supervised and unsupervised learning. Provide an example of each type of learning algorithm. | CO1 | L1 | 7M |
|  | b) | Explain briefly maximum likelihood estimation algorithm and list its properties. | CO1 | L3 | 7M |
| (OR) | | | | | |
| 3 | a) | Explain briefly about multilayer perceptron model. | CO1 | L1 | 7M |
|  | b) | What propagates backward in back propagation algorithm? How that process helps in optimizing the training processes? | CO1 | L1 | 7M |
| **Unit-II** | | | | | |
| 4 | a) | List and explain the various activation functions used in modelling of artificial neuron. Also explain their suitability with respect to applications | CO2 | L1 | 7M |
|  | b) | Differentiate Machine learning Vs Deep learning. | CO2 | L2 | 7M |
| **(OR)** | | | | | |
| 5 | a) | What is representation learning? Explain different types of representation learning models in deep learning. | CO2 | L1 | 7M |
|  | b) | Explain the training process of Restricted Boltzmann Machines? | CO2 | L4 | 7M |
| **Unit-III** | | | | | |
| 6 | a) | Explain about ResNet architecture and list its features. | CO3 | L3 | 7M |
|  | b) | How CNNs are different from traditional neural networks with all fully connected layers? Explain with the help of various connection properties of convolution layers. | CO3 | L2 | 7M |
| **(OR)** | | | | | |
| 7 | a) | What is the context between the depth and filter size in building the CNN models? | CO3 | L1 | 7M |
| **P.T.O**  **20EC705/JO** | | | | | |
|  | b) | The input consists of 7 channels of size 16x16. Use convolution later. Find the number of weights in each of the configurations below. (a) The output of the first layer consists of 8 feature maps, and the filters are of size 5x5. The convolution is done with stride 2 and zero padding is used. (b) Now suppose we made this a fully connected layer, but the number of input and output units are kept the same as in the network described in the Part (a) above. | CO3 | L1 | 7M |
| **Unit-IV** | | | | | |
| 8 | a) | In the LSTM model, explain exactly how the cell state is updated from Ct-1 to Ct, using the previous state ht-1 and the current input xt. | CO4 | L2 | 7M |
|  | b) | List the architectures capable of learning from sequential data. What makes them different from others? | CO4 | L1 | 7M |
| **(OR)** | | | | | |
| 9 | a) | Vanishing gradient problem makes it difficult to train RNNs. How the LSTM dealt with that problem? | CO4 | L2 | 7M |
|  | b) | What are the challenges and limitations of using BPTT for training RNN’S. | CO4 | L1 | 7M |

