**ANSWER KEY SUBMISSION**

|  |  |  |  |
| --- | --- | --- | --- |
| **Date of Exam & Session** | **14/10/2022 & AN** | **Category of Exam** | **CLA1** |
| **Course Name** | **Artificial Neural Networks** | **Course Code** | **18CSE388T** |
| **Name of the Faculty submitting** | **Ms.L.Sasikala** | **Date of submission of Answer Key** | **18/10/2022** |
| **Department to which the faculty belongs to** | **CSE** | **Total Marks** | **50** |

**PART - A (5x1 = 5)**

**ANSWER ALL THE QUESTIONS**

|  |  |  |
| --- | --- | --- |
| **Q.No** | **Questions** | **Marks** |
| 1 | In which network topology, loops are allowed?  a) Feed Forward network  b) Forward Feed network  c) Lateral recurrence network  **d) Direct recurrence network** | **1** |
| 2 | What is the simplest and also commonly used linear activation function?  **a) Binary function**  b) Quadrille function  c) Propagation function  d) Fermi function | **1** |
| 3 | Which function converts vector inputs to scalar network inputs?  a) Activation function  **b) Propagation function**  c) Output function  d) Input function | **1** |
| 4 | When the first artificial neural network is founded in the year \_\_\_\_\_\_\_?  a) 1957  **b) 1958**  c) 1959  d) 1960 | **1** |
| 5 | What is the condition to activate the neuron?  a) Network input < Threshold value  **b) Network input > Threshold value**  c) When it reaches highest point of sensation  d) when it reaches zero input | **1** |
| 6 | Which of the following is not example of unsupervised neural network?  a) Self organizing feature map  b) Hebb network  **c)** **Back Propagation network**  d) Self organizing feature map and Hebb network | **1** |
| 7 | What does RNN stands for?  **a) Recurrent Neural Network**  b) Recurring Neural Network.  c) Removable Neural Network  d) Restoring Neural Network | **1** |
| 8 | Which of the following option is not the disadvantage of Recurrent Neural Network?  a) Training an RNN is quite a challenging task  **b) Inputs of any length can be processed in this model.**  c) Exploding and gradient vanishing is common in this model.  d) It cannot process very long sequences if using 'tanh' or 'relu' as an activation function | **1** |
| 9 | Why are linearly separable problems of interest of neural network researchers?  **a) because they are the only class of problem that perceptron can solve successfully.**  b) because they are the only mathematical functions that are continue  c) because they are the only class of problem that network can solve successfully  d) because they are the only mathematical functions you can draw | **1** |
| 10 | What is back propagation?  a) it is the transmission of error back through the network to adjust the inputs  **b) it is the transmission of error back through the network to allow weights to be adjusted so that the network can learn.**  c) it is another name given to the curvy function in the perceptron.  d)it is the transmission of error back through the network to adjust the output. | **1** |

**PART - B (4x4= 16)**

**ANSWER ANY FOUR QUESTIONS**

|  |  |  |
| --- | --- | --- |
| **Q.No** | **Questions** | **Marks** |
| **11** | **Describe the order in which neuron activations are calculated.**  It is very important in which order the individual neurons receive and process the input and output the results.  **(i) Synchronous activation -** All neurons of a network calculate network inputs at the same time by means of the propagation function, activation by means of the activation function and output by means of the output function **.**  **(ii) Asynchronous activation -** The neurons do not change their values simultaneously but at different points of time. There exist different orders in **Asynchronous activation.**   1. **Random order -** With random order of activation,a neuron i is randomly chosen and its neti, ai and oi are updated. 2. **Random permutation -** Each neuron is chosen exactly once, but in random order, during one cycle. 3. **Topological order -** neurons are updated during one cycle and according to a fixed order.  This order is defined by network topology 4. **Fixed orders of activation during implementation -** When implementing feedforward networks, for example, it is common practice to establish the activation order once based on the topology and then apply that order without further verification during runtime. | **2**  **2** |
| **12** | **Distinguish between a feed forward network and a recurrent network.**  feed-forward neural network :   * + Cannot handle sequential data.   + Depends on the current input.   + Cannot memorize previous inputs.   + Bad at predicting what’s coming next.   Recurrent neural networks:   * + An algorithm for sequential data, used by Apple's Siri and Google's voice search.   + The first algorithm that remembers its input, due to an internal memory.   + When it makes a decision, it considers the current input and what it has learned from the inputs it received previously. | **2**  **2** |
| **13** | **Discuss the role of activation function in artificial neuron and list down names of some popular activation functions.**  The reason for using activation functions in Neural Networks are as follows:  1. The idea behind the activation function is to introduce nonlinearity into the neural network so that it can learn more complex functions.  2. Without the Activation function, the neural network behaves as a linear classifier, learning the function which is a linear combination of its input data.  3. The activation function converts the inputs into outputs.  4. The activation function is responsible for deciding whether a neuron should be activated i.e, fired or not.  5. To make the decision, firstly it calculates the weighted sum and further adds bias with it.  6. So, the basic purpose of the activation function is to introduce non-linearity into the output of a neuron.  Some of the popular activation functions that are used while building the deep learning models are as  follows:   * Sigmoid function * Hyperbolic tangent function * Rectified linear unit (RELU) function * Leaky RELU function * Maxout function | **2**  **2** |
| **14** | **What is Rprop? Explain the term “Rprop is an extension to back propagation of error”.**  Resilient backpropagation is an extension to backpropagation of error. We have two backpropagation specific properties that can occasionally be a problem:   * 1. Users of backpropagation can choose a bad learning rate η.   2. The further the weights are from the output layer; the slower backpropagation learns.   Martin Riedmiller et al. enhanced backpropagation and called their version resilient backpropagation (short Rprop).  **Resilient back propagation (Rprop),** an algorithm that can be used to train a neural network, is similar to the more common (regular) back-propagation. But it has two main advantages over back propagation:  First, training with Rprop is often faster than training with back propagation.  Second, Rprop doesn't require you to specify any free parameter values, as opposed to back propagation which needs values for the learning rate (and usually an optional momentum term).  The main disadvantage of Rprop is that it's a more complex algorithm to implement than back propagation. | **2**  **2** |
| **15** | **What do you mean by Gradient descent? Explain the problems associated with Gradient procedure?**  Gradient Descent is an optimization algorithm that aims to minimize the cost function or to minimize an error. Its main objective is to find the local or global minima of a function based on its convexity. This determines in which direction the model should go to reduce the error.  Problems associated with Gradient procedure are,  important general paradigm when  continuously parameterized hypothesis  the error can be differentiated with respect to the hypothesis parameters  **The key practical problems are:**   * converging to a local minimum can be quite slow * if there are multiple local minima, then there is no guarantee that the procedure will find the global minimum | **2**  **2** |
| **16** | **Explain Hebbian learning rule.**  Hebbian rule stated as "If neuron j receives an input from neuron i and if both neurons are strongly active at the same time, then increase the weight wi,j”.  **https://lh5.googleusercontent.com/Al4f1f8qNiNe8QPCdEXMSb6LXYnSrLiiFXe90tULR6_KlW_rEkpZKx6BZBjqJ3sHSfQgBrm9QtvrGRpCA2VORn7xXyNTYDkRoLc_16HxCjxd0_jgkx0Yeh8GsDxM3yGUFpdgG6SXjJJgj5TO5efK85vmfhtmGPkTP5cP7oK3Mwxd_6enYJqDMU7tY7hOxLjYn0m08A=nw**  with       https://lh5.googleusercontent.com/GOpUOdvHoZOl_--70MMOBUEShs0wo9at3SLqqMaeFnMX9dzAROXi6hKEdlXynE7JNVXc8QxewxTCyPS2tHXsZcvYogYxKmtuYFaC0Luhh2068Bv_zK9ZxLiSiu6PNf279k6WP_btbrRjjH6w-kNT_NAggYP5FqWVBX1L10u5bsXzUeHfrlR2yTsGlvNYYKA1EEUCQw=nw   being the change in weight from i to j, which is proportional to the following factors:   * + The output oiof the predecessor neuron i, as well as,   + The activation ajof the successor neuron j   + A constant η, i.e. the learning rate   + The changes in weight are simply added to the weight wi,j.   **The generalized form of the Hebbian Rule** :specifies the proportionality of the change in weight to the product of two undefined functions, but with defined input values. Thus, the product of the functions.  https://lh3.googleusercontent.com/mJc8eZ7Hj4K0E877sHVAd2CtnUe74Ncr4yT_2xTX_3a1EcF7oSC41rEVCfQRo3luJRnNUzXVJnMykKMVaLKcbhB9oFFdvn1kVyroOaKLa70aC_K3Qivd7LADBx31Nn-wWXz9jY_F6WmBIGAh31hJutslgsSeAlfaCqcN5rSU2hCJLbMAwzHzbLLiygvF13LU0jDBGQ=nw  Changes in weight = Learning rate . Pre synaptic signal . Post synaptic signal | **2**  **2** |

**PART - C (2 x12 = 24)**

**ANSWER ALL THE QUESTIONS**

|  |  |  |
| --- | --- | --- |
| **Q.No** | **Questions** | **Marks** |
| **17a** | **With neat sketch explain network topologies in ANN.**  Network consists of following 3 layers,  **Input Layer:**  The number of neurons in the input layer should be equal to the attributes or features in the dataset.  It provides information from the outside world to the network.  No computation is performed at this layer  Nodes just pass on the information to the next layer.  **N-Hidden Layer:**  Nodes of this layer are not exposed to the outer world.  Hidden layer performs all sort of computation on the features entered through the input layer.  Transfer the result to the output layer.  **Output Layer:**  This layer brings up the information learned by the network to the outer world.  **1.Feedforward network**  Each neuron in one layer has only directed connections to the neurons of the next layer.  Every neuron i is connected to all neurons of the next layer.  The decision making are based on the current input.  It doesn’t memorize the past data, and there’s no future scope.      **2. Recurrent Neural Network (RNN)**   * An RNN remembers each and every information through time. * It is useful in time series prediction only because of the feature to remember previous inputs. * Let’s take a character level RNN where we have a word “Hello”. * We provide the first 4 letters i.e., h,e,l,l and ask the network to predict the last letter i.e., ’o’. * The vocabulary of the task is just 4 letters {h,e,l,o}. * In real case scenarios involving natural language processing, the vocabularies include the words in entire Wikipedia database, or all the words in a language. * This is called Long Short-Term Memory   **(i) Direct recurrence neural network:**  Some networks allow neurons to be connected to themselves, which is called direct recurrence or self- recurrence, start and end at the same neuron.The diagonal of the weight matrix W may be different from 0. Below diagram shows the direct recurrent network and Hindon diagram,    **(ii)** **Indirect recurrence neural network:**   * If connections are allowed towards the input layer, called indirect recurrences. * A neuron j can use indirect forwards connections to influence itself. * On a feedforward network, now with additional connections between neurons and their preceding layer being allowed. * Therefore, below the diagonal of W is different from 0.     **(iii) Lateral recurrence neural network:**   * Connections between neurons within one layer are called lateral recurrences. * Each neuron often inhibits the other neurons of the layer and strengthens itself. * As a result, only the strongest neuron becomes active (winner-takes-all scheme). * A laterally recurrent network permits connections within one layer.     **3.** **Completely linked neural network:**   * Every neuron is always allowed to be connected to every other neuron, as a result every neuron can become an input neuron. The matrix W may be unequal to 0 everywhere, except along its diagonal. * Permit connections between all neurons, except for direct network with laterally recurrent neurons.     . | **2**  **2**  **2**  **6** |
| **17b** | **Explain various function aspects of artificial neuron model with respect to bias, weighted inputs and activation functions.**  A bias neuron is a neuron whose offset value is always 1. The threshold value is an activation function parameter of a neuron that indicates the activity of a neuron. The bias does not depend on any input value.A bias neuron is used to represent neuron biases as connection weights, which enables any weight training algorithm to train the biases at the same time. Threshold values are implemented as connection weights and can directly be trained together with the connection weights, which considerably facilitates the learning process.  In absence of bias, model will train over point passing through origin only. In accordance with real-world scenario, the model should be more flexible.    **Weighted Input:**  The propagation function converts vector inputs to scalar network inputs.  For a neuron j, The propagation function receives the outputs    From other neurons i1, i2, . . . , in (which are connected to j). Transforms them. The connecting weights wi,j into the network input netj. Processed by the activation function. The network input is the result of the propagation function.  **The activation function**  When comparing with a neuron-based model that is in our brains, the activation function is “what is to be fired to the next neuron”. An activation function is a function that is added into an ANN to help the ”network learn complex patterns in the data”. Activation function decides, whether a neuron should be activated or not by calculating weighted sum and further adding bias with it. Determines the activation of a neuron dependent on network input and threshold value. The activation aj of a neuron j depends on the previous activation state of the neuron and the external input.  The activation function transforms the network input netj, as well as the previous activation state aj(t - 1) into a new activation state aj(t), with the threshold value Θ.    The most important feature in an activation function is its ability to add non-linearity into a neural network. A neural network without an activation function is essentially just a linear regression model with limited abilities.  **Linear Activation Function: Non-Linear Function** | **6**  **2**  **4** |
| **18a** | **Distinguish between Supervised Learning and Unsupervised Learning in ANN.**  **Supervised learning:**   * Supervised learning is a process of providing labelled input data as well as correct output data to the machine learning model. * The aim of a supervised learning algorithm is to find a mapping function to map the input variable(x) with the output variable(y). * In supervised learning, the training set consists of input patterns as well as their correct results in the form of the precise activation of all output neurons. * Thus, for each training set that is fed into the network the output, for instance, can directly be compared with the correct solution and the network weights can be changed according to their difference. * The training data provided to the machines work as the supervisor that teaches the machines to predict the output correctly. * It applies the same concept as a student learns in the supervision of the teacher.   **Types of supervised machine learning algorithms:**  Supervised learning can be further divided into two types of problems**:**  **(i) Classification:**  Classification algorithms are used when the output variable is categorical, which means there are two classes such as Yes-No, True-false, etc.  Below are some popular Classification algorithms which come under supervised learning:  K-nearest neighbor (KNN),Decision Trees,Naive Bayes,Support vector Machines (SVM)  **(ii) Regression**:  Regression algorithms are used if there is a relationship between the input variable and the output variable. Below are some popular Regression algorithms which come under supervised learning:  Linear Regression, Logistic Regression, Ridge Regression ,Lasso Regression, Polynomial Regression.  **Unsupervised learning:**   * Unsupervised learning is a machine learning technique in which models are not supervised using training dataset. * Instead, models itself find the hidden patterns and insights from the given data. * Unsupervised learning works on unlabelled and uncategorized data which make unsupervised learning more important. * The training set only consists of input patterns, the network tries by itself to detect similarities and to generate pattern classes. * The goal of unsupervised learning is to find the underlying structure of dataset, group that data according to similarities, and represent that dataset in a compressed format. * In real-world, we do not always have input data with the corresponding output so to solve such cases, we need unsupervised learning. * It can be compared to learning which takes place in the human brain while learning new things.   **Types of Unsupervised machine learning algorithm:**  The unsupervised learning algorithm can be further categorized into two types of problems:  **Clustering:**  Clustering is a method of grouping the objects into clusters such that objects with most similarities remains into a group and has less or no similarities with the objects of another group.  Cluster analysis finds the commonalities between the data objects and categorizes them as per the presence and absence of those commonalities.  Below are some popular Clustering algorithms which come under unsupervised learning:  Centroid-based Clustering, Density-based Clustering, Distribution-based Clustering, Hierarchical Clustering.  **Association:**  An association rule is an unsupervised learning method which is used for finding the relationships between variables in the large database.  It determines the set of items that occurs together in the dataset.Such as people who buy X item (suppose a bread) are also tend to purchase Y (Butter/Jam) item.  Below is the popular Association algorithm which come under unsupervised learning:  Apriori algorithm,  **Applications of unsupervised learning algorithms:**  Fraud detection, Malware detection, Identification of human errors during data entry, Conducting accurate basket analysis, etc. | **4**  **2**  **4**  **2** |
| **18b** | **Demonstrate in detail about Single Layer Perceptron (SLP) learning algorithm for linearly separable classification.**  A single layer perceptron (SLP) is a perceptron having only one layer of variable weights and one layer of output neurons Connections with trainable weights go from input layer to an output neuron Ω, which returns the information whether the pattern entered at the input neurons was recognized or not. Certainly, the existence of several output neurons Ω 1, Ω2. . . Ω n does not considerably change the concept of the perceptron. A perceptron with several output neurons can also be regarded as several different perceptron with the same input.    **Perceptron learning algorithm:**   * The original perceptron learning algorithm with binary neuron activation function is described in algorithm. * It has been proven that the algorithm converges in finite time, so in finite time the perceptron can learn anything. * Suppose that we have a single layer perceptron with randomly set weights which we want to teach a function by means of training samples. * The set of these training samples is called P. * It contains, as already defined, the pairs (p, t) of the training samples p and the associated teaching input t.     x is the input vector,y is the output vector of a neural network.  Output neurons are referred to as    i is the input value of a neuron,o is the output value of a neuron.  The error vector Ep represents the difference (t−y) under a certain training sample p, O be the set of output neurons, I be the set of input neurons.  Our learning target will be certainly be, that for all training samples the output y of the network is approximately the desired output t, | **4**  **2**  **4**  **2** |

**HOD/CSE**