**VISHNU INSTITUTE OF TECHNOLOGY (A)**

**DEPARTMENT OF AI & DS, AI & ML**

**DEEP LEARNING**

**III B.Tech II SEM MID-I Objective Bits**

**Unit-1**

1. Why is deep learning gaining popularity in recent years? [ ]

A) It requires less computational power than traditional machine learning.

B) It can automatically extract hierarchical features from data.

C) It is limited to small datasets.

D) It relies solely on rule-based systems.

1. Which of the following is a potential application of deep learning? [ ]

A) Solving symbolic logic problems

B) Image and speech recognition

C) Only working with structured data

D) Handling rule-based systems efficiently

1. In supervised learning, what is the role of the model? [ ]

A) To predict labels or outcomes based on input data

B) To generate random data points

C) To classify data into clusters

D) To find patterns in unlabeled data

1. What characterizes unsupervised learning? [ ]

A) The model is provided with labeled training data.

B) The model learns from input data without explicit labels.

C) The model requires both input and output pairs.

D) Unsupervised learning is not applicable in deep learning.

1. Reinforcement learning involves: [ ]

A) Learning from labeled data.

B) Learning from reward-based feedback.

C) Memorizing input-output pairs.

D) Ignoring the environment's response.

1. What is a common challenge in deep learning related to overfitting? [ ]

A) Underfitting the data

B) Memorizing noise in the training data

C) Not utilizing enough layers in the network

D) Relying too much on rule-based systems

1. What is the vanishing gradient problem in deep learning? [ ]

A) Rapid convergence of the optimization algorithm

B) Gradients becoming too large during backpropagation

C) Gradients approaching zero, hindering weight updates

D) Only using shallow networks in the learning process

1. Why is the choice of an appropriate loss function crucial in deep learning? [ ]

A) It has no impact on model performance.

B) It determines the learning rate.

C) It guides the optimization process during training.

D) It influences the number of layers in the network.

1. What is the basic building block of biological neural networks? [ ]

A) Nodes

B) Synapses

C) Weights

D) Activation functions

1. How do artificial neural networks attempt to mimic biological neural networks? [ ]

A) By using labeled data for training

B) By incorporating layers of nodes and connections

C) By ignoring activation functions

D) By relying solely on supervised learning

1. What is a perceptron in the context of neural networks? [ ]

A) A type of activation function

B) The output layer of a neural network

C) The basic building block of a neural network

D) A type of regularization technique

1. What is the main purpose of the perceptron learning algorithm? [ ]

A)To classify input data into two categories

B) To extract hierarchical features from data

C) To minimize the vanishing gradient problem

D) To perform clustering on input data

1. What is the role of an activation function in a neural network? [ ]

A) To determine the learning rate

B) To introduce non-linearity to the model

C) To define the loss function

D) To control the number of layers in the network

1. Which activation function is commonly used in the output layer for binary classification problems? [ ]

A) ReLU

B) Sigmoid

C) Tanh

D) Softmax

1. How does the Rectified Linear Unit (ReLU) activation function differ from the sigmoid function? [ ]

A) ReLU is linear, while sigmoid is non-linear.

B) ReLU is prone to vanishing gradients.

C) ReLU outputs zero for negative inputs.

D) Sigmoid saturates for extreme values.

1. What is the purpose of the hidden layers in an artificial neural network? [ ]

A) To determine the learning rate

B) To represent the input data

C) To extract hierarchical features from data

D) To connect nodes with synapses

1. How do artificial neural networks generalize to unseen data? [ ]

A) By memorizing the entire training dataset

B) By learning specific patterns without abstraction

C) By adapting to noise in the training data

D) By extracting features and relationships from the training data

1. Why is the XOR problem considered a challenging task for a single-layer perceptron? [ ]

A) XOR is a linearly separable problem.

B) Single-layer perceptrons lack the capacity for non-linear decision boundaries.

C) XOR has too few training examples.

D) XOR has a simple solution.

1. How does a multi-layer perceptron (MLP) address the limitations of a single-layer perceptron in solving the XOR problem? [ ]

A) By increasing the number of nodes in the input layer

B) By adding more layers and non-linear activation functions

C) By removing the activation functions

D) By using a larger learning rate

1. In a multi-layer perceptron, what is the purpose of the backpropagation algorithm?[ ]

A) To initialize model weights

B) To compute gradients and update weights during training

C) To randomly adjust parameters

D) To memorize input-output pairs

1. The fundamental unit of network is [ ]

a) Brain b) Nucleus c) Neuron d) axon

1. A 3-input neuron is trained to output a zero when the input is 110 and a one when the input is 111. After generalization, the output will be zero when and only when the input is? [ ]
2. 000 or 110 or 011 or 101 b) 010 or 100 or 110 or 101
3. 000 or 010 or 110 or 100 d) 100 or 111 or 101 or 001
4. Which neural network has only one hidden layer between the input and output? [ ]
5. Shallow neural network b) Deep neural network
6. Feed-forward neural networks d) Recurrent neural networks
7. Assume a simple MLP model with 3 neurons and inputs= 1,2,3. The weights to the input neurons are 4,5 and 6 respectively. Assume the activation function is a linear constant value of 3. What will be the output? [ ]
8. 32 b) 64 c) 96 d) 128
9. TensorFlow is developed by \_\_\_\_\_\_\_\_\_\_\_\_\_ [ ]
   1. Microsoft Team b. Google Brain team c. IBM Team d.None of the above
10. In a classification problem, which of the following activation function is most widely used in the output layer of neural networks? [ ]

a.Sigmoid function  b. Hyperbolic function

c.Rectifier function d.All the above

1. Suppose the number of nodes in the input layer is 5 and the hidden layer is 10. The maximum number of connections from the input layer to the hidden layer would be- [ ]
   1. More than 50 b. Less than 50 c.50 d.None
2. Weighted sums are passed via activation functions and \_\_\_\_\_\_\_\_\_ is produced? [ ]
   1. Input b.Output c.Network d.Connection
3. In a neural network, which of the following causes the loss not to decrease faster? [ ]
   1. Stuck at a local minima b. High regularization parameter

c. Slow learning rate d. All of the above

1. What distinguishes deep learning from traditional machine learning approaches? [ ]

A) Deep learning uses fewer layers in its models.

B) Deep learning relies on explicit rule-based systems.

C) Deep learning automatically learns hierarchical features from data.

D) Deep learning is only suitable for linearly separable problems.

**Unit-2**

1. What is the primary purpose of a loss function in a deep neural network? [ ]

A) To initialize model weights

B) To measure the accuracy of the model

C) To regularize the model

D) To compute gradients during backpropagation

1. In a classification task, which loss function is commonly used when dealing with binary outcomes?

[ ]

A) Mean Squared Error (MSE)

B) Cross-Entropy Loss

C) Huber Loss

D) L1 Regularization Loss

1. How does the choice of a loss function differ between regression and classification problems?

[ ]

A) It doesn't matter; any loss function can be used for both.

B) Regression uses cross-entropy, while classification uses mean squared error.

C) They typically have different loss functions tailored to their respective objectives.

D) The choice depends on the type of optimizer used.

1. What is the main idea behind the gradient descent optimization algorithm? [ ]

A) Minimizing the loss by adjusting model parameters in the opposite direction of the gradient

B) Maximizing the loss to explore the entire solution space

C) Randomly adjusting parameters to find the global minimum

D) Ignoring gradients and updating parameters randomly

1. How does the learning rate impact the convergence of the gradient descent algorithm? [ ]

A) Higher learning rates always lead to faster convergence.

B) Lower learning rates may slow down convergence but offer more stability.

C) Learning rate has no impact on convergence.

D) Learning rate only affects the accuracy, not the convergence speed.

1. In stochastic gradient descent (SGD), how is the model updated after processing each individual data point? [ ]

A) The weights are updated once after processing all data points.

B) The weights are updated after processing each batch of data points.

C) The weights are updated after processing each individual data point.

D) The weights remain unchanged throughout training.

1. What is the primary advantage of RMSprop over traditional gradient descent? [ ]

A) It adapts the learning rate individually for each parameter.

B) It always uses a fixed learning rate.

C) It updates all model parameters simultaneously.

D) It doesn’t utilize gradients during optimization.

1. How does RMSprop handle the issue of varying scales of gradients for different parameters? [ ]

A) By applying a moving average to the gradients

B) By normalizing gradients

C) By using a fixed learning rate for all parameters

D) By ignoring gradients with small values

1. In the context of neural networks, what is backpropagation? [ ]

A) A method to initialize neural network weights

B) A technique for forward-pass computation

C) The process of computing gradients and updating weights during training

D) A regularization technique for preventing overfitting

1. Why is the chain rule of calculus crucial in the backpropagation algorithm? [ ]

A) It helps in optimizing the loss function directly.

B) It enables the computation of gradients with respect to each layer's input.

C) It is used to calculate the initial weights of the network.

D) It is only relevant for the output layer.

1. What are some challenges associated with training deep neural networks compared to shallow networks? [ ]

A) Shallow networks have more parameters to optimize.

B) Deep networks suffer from the vanishing gradient problem.

C) Shallow networks are more prone to overfitting.

D) Deep networks always converge faster.

1. How does the vanishing gradient problem impact the training of deep models? [ ]

A) It leads to faster convergence.

B) It causes the model to get stuck during training.

C) It has no effect on training.

D) It only affects the output layer.

1. What is the primary goal of regularization in deep learning? [ ]

A) To increase model complexity

B) To reduce model accuracy

C) To prevent overfitting

D) To speed up training

1. How does L2 regularization differ from L1 regularization? [ ]

A) L2 regularization adds the absolute values of the weights to the loss function.

B) L2 regularization encourages sparsity in the weights.

C) L2 regularization penalizes large weights by adding their squared values to the loss.

D) L2 regularization is only applicable to output layers.

1. How does early stopping prevent overfitting in deep neural networks? [ ]

A) By stopping training as soon as the model achieves maximum accuracy

B) By stopping training when the loss on the validation set starts to increase

C) By continuously increasing the learning rate during training

D) By ignoring the validation set during training

1. What are potential drawbacks or limitations of early stopping? [ ]

A) It may lead to overfitting.

B) It can slow down the training process.

C) It might stop training too early, missing the optimal model.

D) It is only effective for small datasets.

1. In the context of deep learning, what is data augmentation? [ ]

A) Reducing the size of the training dataset

B) Modifying the input data to create new training samples

C) Increasing the dimensionality of the input features

D) Removing outliers from the dataset

1. How does data augmentation contribute to regularization in deep learning? [ ]

A) By introducing noise into the input data

B) By increasing the size of the model parameters

C) By creating diverse training samples from existing data

D) By decreasing the learning rate during training

1. What is the purpose of dropout in a neural network? [ ]

A) To increase the size of the model

B) To add noise to the input features

C) To randomly deactivate neurons during training

D) To decrease the learning rate

1. How does dropout prevent overfitting in deep neural networks? [ ]

A) By increasing the number of training epochs

B) By reducing the number of neurons in each layer

C) By preventing reliance on specific neurons for the entire task

D) By ignoring the validation set during training

1. Which of the following steps can be taken to prevent overfitting in a neural network? [ ]

A) Dropout of neurons

B) Loss functions

C) Gradient descent

D) None

1. What is the purpose of a loss function? [ ]
2. Calculate the error value of the forward network
3. Optimize the error values according to the error rate
4. Both A and B
5. None
6. Suppose you have a dataset from where you have to predict three classes. Then which of the following configuration you should use in the output layer? [ ]
7. Activation function = softmax, loss function = cross entropy
8. Activation function = sigmoid, loss function = cross entropy
9. Activation function = softmax, loss function = mean squared error
10. Activation function = sigmoid, loss function = mean squared error
11. What does a gradient descent algorithm do? [ ]
12. Tries to find the parameters of a model that minimizes the cost function
13. Adjusts the weights at the input layers
14. Both
15. None
16. In a neural network, which of the following causes the loss not to decrease faster? [ ]
17. Stuck at a local minima
18. High regularization parameter
19. Slow learning rate
20. All of the above
21. Which of the following is true about dropout? [ ]
22. Applied in the hidden layer nodes
23. Applied in the output layer nodes
24. Both A and B
25. None

**Unit-3**

1. What is the primary purpose of convolution operation in a Convolutional Neural Network ? [ ]

A) To flatten the input data

B) To reduce the dimensionality of the data

C) To extract features from the input data

D) To apply activation functions to the input

1. In a CNN, what is the role of a convolutional layer? [ ]

A) Aggregating features

B) Reducing dimensionality

C) Extracting spatial hierarchies of features

D) Connecting all nodes in the network

1. What is the purpose of pooling in a CNN? [ ]

A) Increasing the spatial dimensions of the input

B) Reducing the number of parameters

C) Adding noise to the feature maps

D) Applying non-linear transformations to the data

1. Where fully connected layers are typically placed in CNN architecture? [ ]

A) At the beginning

B) In the middle

C) At the end

D) After pooling layers

1. Why is padding used in convolutional layers? [ ]

A) To increase computational efficiency

B) To reduce the receptive field

C) To preserve spatial dimensions

D) To remove redundant features

1. What is the purpose of transposed convolutional layers in a CNN? [ ]

A) Downsampling the feature maps

B) Increasing the receptive field

C) Upsampling the spatial resolution

D) Applying global pooling

1. Why are activation functions like ReLU commonly used in convolutional layers? [ ]

A) To introduce non-linearity

B) To reduce overfitting

C) To increase interpretability

D) To speed up training convergence

1. What is the purpose of a 1x1 convolutional layer in a CNN? [ ]

A) Feature reduction

B) Spatial hierarchy extraction

C) Dimensionality expansion

D) Global pooling

1. How can CNNs handle input images of different sizes without altering the network architecture?

[ ]

A) Using fully connected layers

B) Applying global average pooling

C) Increasing the filter size

D) Removing pooling layers

1. How does the stride parameter in the convolution operation affect the output size? [ ]

A) It increases the output size

B) It decreases the output size

C) It has no effect on the output size

D) It depends on the filter size

1. What is the main difference between max pooling and average pooling? [ ]

A) Max pooling adds noise to the data

B) Average pooling reduces dimensionality

C) Max pooling selects the maximum value, while average pooling computes the average

D) Average pooling always results in larger values than max pooling

1. How does the receptive field change as you go deeper into the layers of a CNN? [ ]

A) It remains constant

B) It increases

C) It decreases

D) It depends on the activation function

1. What is the impact of using 'valid' padding in convolution? [ ]

A) It increases the spatial dimensions

B) It reduces the receptive field

C) It removes unnecessary features

D) It leads to smaller output sizes

1. How do multiple convolutional layers contribute to hierarchical feature learning in CNNs? [ ]

A) They focus on global features

B) They learn fine-grained features first

C) They progressively learn more abstract features

D) They have no impact on feature hierarchy

1. What is the primary effect of pooling on the spatial dimensions of the input volume? [ ]

A) It increases spatial resolution

B) It reduces spatial resolution

C) It has no effect on spatial dimensions

D) It depends on the type of pooling used