**DEEPLEARNING**(Worksheet-3)

1. (B)
2. (C)
3. (A)
4. (D)
5. (C)
6. (B)
7. (B)
8. (A)
9. (A)
10. (B)
11. By calculating weighted sum with adding bias, Activation function decides that whether neuron should be activated or not. It makes the backpropagation possible because of the gradients are supplied with the errors to updates weights and biases. It does the non-linear transformation to the input making it capable to learn and perform more complex tasks. So without activation function this can’t happen and it remains a simple linear regression model only.
12. Forward propagation does the calculation and storage of intermediate variables and outputs in the format of input layer to output layer. When we implemented the algorithm, we focused on the calculations involved in forward propagation through the model.

Backward propagation does the calculation of gradient of neural network parameters. It traverses the network in reverse order that is from output layer to input layer.

1. **Stochastics gradient descent**: it does the updates after every input samples. This produces noisy or fluctuating loss outputs. This method is slow as compared to mini batch gradient descent. It is an iterative method for optimizing a differentiable objective function

**Batch gradient descent**: it is much slow on very large training data because it involves calculations over the full training set at each step.so it becomes very computationally expensive to do batch gradient descent. This is good for convex or relatively smooth error manifolds.

**Minibatch gradient descent**: it takes a batch of input samples and updates the gradient after that batch is processed. Because of this, it leads to much faster training.

1. By using minibatch gradient descent we can process a large number of input samples in a very short time period.

Mini batch allows for a high-quality gradient and this will be very useful for using high learning rates.

Sometimes it converges faster.

1. Transfer learning is a research problem in machine learning which focuses on storing knowledge gained while solving one problem and applying it to a different but related problem.