THEORETICAL ASSIGNMENT
NAME : MEETRAJSINH CHUDASAMA
MODULE 8 :NEURAL NETWORKS & DEEP LEARNING

Q-1. Explain the Architecture of Neural Networks.

The Artificial Neural Network (ANN) architecture refers to the structured arrangement of nodes and layers that define how an artificial neural network processes and learns from data. The design of ANN influences its ability to learn complex patterns and perform tasks efficiently. Just like the human brain processes information through interconnected neurons, ANN use layers of artificial neurons to learn patterns and make predictions. The architecture explains how data flows through the network, how neurons are connected, and how the network learns and makes predictions.
Here are the key components of a neural network architecture.

1. Layers :

Neural networks consist of layers of neurons, which include the

- **Input Layers :** This is where the network receives its input data. Each input neuron in the layer corresponds to a feature in the input data.

- **Hidden Layers :** These layers perform most of the computational heavy lifting. A neural network can have one or multiple hidden layers. Each layer consists of neurons that transform the inputs into something that the output layer can use.

- **Output Layers :** The final layer produces the output of the model. The format of these outputs varies depending on the specific task (e.g., classification, regression).

2. Neurons (Nodes) :

Neurons are the basic computational units that perform a weighted sum of their inputs, apply a bias, and pass the result through an activation function.
Weights and biases: Weights represent the strength of the connections between neurons, and biases allow neurons to make predictions even when all inputs are zero.

3. Activation function :

Non-linear functions (like ReLU and Sigmoid) are used to introduce non-linearity into the network, enabling it to model complex relationships.

Q-2. Differences Between CNNs, RNNs, and Transformers.

RNNs are a type of neural network architecture that has shown great success in sequence modeling tasks. They are designed to process sequences of inputs, such as sentences or time-series data. RNNs have a recurrent structure that allows them to capture the context and dependencies between different inputs in a sequence. However, RNNs suffer from the problem of vanishing gradients, which makes it difficult for them to capture long-term dependencies.

CNNs, on the other hand, are a type of neural network architecture that is widely used for image recognition tasks. They are designed to extract features from the input data by applying convolutional filters. CNNs have been adapted for NLP tasks by treating words as images and using 1D convolutional filters to extract features from sentences. However, CNNs have limited ability to capture the relationships between different words in a sentence.

Transformers are a type of neural network architecture that aredesigned to overcome the limitations of RNNs and CNNs in NLP tasks. Transformers are based on the idea of self-attention, which allows them to capture the relationships between different words in a sentence without relying on a recurrent structure.The transformer architecture consists of an encoder and a decoder. The encoder takes a sequence of input tokens and generates a sequence of hidden representations. The decoder takes the hidden representations generated by the encoder and generates a sequence of output tokens.

CNN	RNN	Transformer
Like looking at parts of a photo to understand the whole image.	Like reading a sentence word by word, remembering what came before.	Like reading the entire sentence at once and figuring out relationships between all the words instantly.