1. What is the fundamental difference between shallow and deep learning?

Shallow learning	Deep learning
Shallow learning is a machine learning technique with fewer stages for processing the data.	Deep learning uses neural networks where there are many hidden layers , input and output layers.
This involves less complex data for training.	The data involved in deep learning are complex.
The model is less efficient to learn patterns from complex data.	The model is more efficient in learning complex patterns from data.
Example: linear regression, logistic regression etc.,	Example: Convolutional Neural Network(CNN), Recurrent Neural Network(RNN) etc.,

2. Can you explain the concept of backpropagation and its significance in training neural networks?

Backpropagation is the process of optimization in neural networks where the accuracy for the model is increased by adjusting weight and bias. This makes the deep learning model more efficient than traditional machine learning techniques. This makes them learn from the complex data effectively.

This involves optimizer for adjusting the weight and bias according to the loss. The various optimizers are gradient descent, stochastic gradient descent, mini batch gradient descent, RMSprop, adam optimizer etc,.

3. What is the vanishing gradient problem, and how does it affect training in deep neural networks?

The vanishing gradient problem occurs when the gradient becomes extremely small as they back propagate through the neural network. When the value becomes extremely small for the weight update, there will be no change and the model cannot learn from the data. This leads to a decrease in the performance of the model.

4. Describe the purpose and function of activation functions in neural networks.

Activation function introduces non-linearity to the model. This helps in learning complex patterns from the large data. The main purpose of the activation function is to make the model learn from complex data, as the data may not always have the linear relationship between the input and the output. Some of the common activation function are Step function, Sigmoid, tanh, ReLU, leaky ReLU, ELU etc.,

5. What are some common activation functions used in deep learning, and when would you choose one over another?

Step function:

Formula:

$y = 0$ if $x<0$, 1 if $x\geq 0$

Range: (0, 1)

This is a very basic activation function. This is used in binary classification models where the output is 0 or 1. This is not widely used now-a-days.

Sigmoid:

Formula:

$y = 1/(1+e^{-x})$

Range: (0, 1)

This forms a s-shaped curve. This is computationally expensive because of the exponent. This is also used for binary classification in the output layer.

Tanh (Hyperbolic tangent function):

Formula:

$y = (e^{x}-e^{-x})/(e^{x}+e^{-x})$

Range: (-1, 1)

This also forms a s-shaped curve and is symmetric around the origin. This is also computationally expensive because of the exponent. This is often used in Convolutional Neural network(CNN) and Recurrent neural network(RNN).

ReLU(Rectified Linear Unit):

Formula:

$y = \max(0, x)$

Range: (0, +inf)

This is generally the first choice of any deep learning model because of its simplicity.

Leaky ReLU:

Formula:

$y = x$ if $x>0$, ax if $x\leq 0$

Range: (-inf, +inf)

This is used over ReLU to deal with dying neuron problems.

ELU(Exponential Linear Unit):

$y = x$ if $x>0$, $a(e^{x}-1)$ if $x\leq 0$

Range: (-a, +inf)

6. Explain the concept of overfitting in deep learning models and methods to prevent it.

The overfitting happens when the model performs well on training data but fails to perform when unseen data is given. This occurs in deep learning when the model captures noise(outliers) and learns from them. The various techniques to overcome this are regularization , dropout etc.,

7. What is dropout regularization, and how does it work to prevent overfitting?

This is an effective technique to overcome overfitting. This involves randomly deactivating the fraction of neurons in the layer and updates the weight through backpropagation during training. During testing all neurons are used.

8. What is the role of convolutional layers in convolutional neural networks (CNNs), and how do they differ from fully connected layers?

Convolutional layers are the primary layer in CNN. This involves learning patterns from the visual data(images). Fully connected layers are applied after convolution and pooling layers. The perform operation of an artificial neural network where every neuron in each layer is connected to every other neuron in other layers.

9. What is the purpose of pooling layers in CNNs, and how do they help in feature extraction?

Pooling layer reduces the spatial dimension while retaining the important information of the visual data(images). This helps in feature extraction from the images. The various pooling techniques are max pooling and average pooling.

10. Describe the architecture of a recurrent neural network (RNN) and its applications in sequential data analysis.

RNN consists of input layer where the sequential data is fed, hidden layer where it processes the current input and also stores memory of past input and output layer which produces output based on current input and hidden state.

The applications of RNN are related to sequential data like text, speech recognition and time series data(Weather forecasting).

11. Explain YoLo Algorithm in depth along with it's real life applications

Yolo(You Only Look Once) is an object detection algorithm that processes within a single pass and creates a bounding box and probability for the object present within the image.

The various application of YOLO are vehicle detection during traffic, traffic sign detection etc.,