https://www.analyticsvidhya.com/blog/2020/04/comprehensive-popular-deep-learning-interview-questions-answers/

Supervised Learning and Deep learning:-

* Image classification
* Text classification
* Sequence tagging

Unsupervised deep learning techniques:-

* Word embeddings (like Skip-gram and Continuous Bag of Words): Understanding Word Embeddings: From Word2Vec to Count Vectors
* Autoencoders: Learn How to Enhance a Blurred Image using an Autoencoder!

**Weights and Bias with Cricket eg :-**

**Factors** like the weather or temperature might have a higher weight, and other factors like equipment would have a lower weight.

**Bias** lets you assign some threshold which helps you activate a decision-point (or a neuron) only when that threshold is crossed.

* need 1 ball and 6 wickets

**Forward propagation:**  calculate the output of the activation at each node .

**Backpropagation :** minimize the cost function by its understanding of how it changes with changing the weights and biases in a neural network.

Change obtained by gradient descent.

**Batch Normalization** is one of the techniques used for reducing the training time of our deep learning algorithm. Just like normalizing our input helps improve our logistic regression model, we can normalize the activations of the hidden layers in our deep learning model as well

### CNN :- it just identifies curves and edges. Thus, instead of looking at the entire image, it helps to just read the image in parts.

### Element wise operations:-

### CNN_Deep_learning_questions

### Artificial Neural Network (ANN) :- each input is given the same weight and fed to the network at the same time

### I saw the movie and hated it

### Difficult to identify how it is attached to movie to RNN is comes into picture.

### RNN :- The addition of a loop is to denote preserving the previous node’s information for the next node, better for sequential data.

### Valid Padding: When we do not use any padding. The resultant matrix after convolution will have dimensions (n – f + 1) X (n – f + 1)

### Same padding: Adding padded elements all around the edges such that the output matrix will have the same dimensions as that of the input matrix.

### Gradient Descent algorithm tries to minimize the error by taking small steps towards the minimum value. These steps are used to update the weights and biases in a neural network.

### exploding gradient :- the steps become too large and this results in larger updates to weights and bias terms

### vanishing gradient : steps are too small and this leads to minimal changes in the weights and bias terms – even negligible changes at times.

### LSTM – Input gate, Output gate, forget gate

### GRU :- Update Gate, Ouput gate

### GRU less complex and faster than LSTM

### Transformers use what is called the attention mechanism. This basically means mapping dependencies between all the parts of a sentence

https://www.edureka.co/blog/interview-questions/deep-learning-interview-questions/