**PROJECT : SESSION 2 REPORT**

what is optimizer?

>>>>>

The goal of machine learning and deep learning is to reduce the difference between the predicted output and the actual output.

This is also called as a Cost function(C) or Loss function. Loss function are convex functions.As our goal is to minimize the Loss function by finding the optimized value for weights.

We also need to ensure that the algorithm generalizes well.

optimizers:-

#SGD

#RMSprop

#Adam

#Adadelta

#Adagrad

#Adamax

#Nadam

#Ftrl

what is activation function and why do we need it?

>>>>>

Definition of activation function:- It’s just a thing function that you use to get the output of node. It is also known as Transfer Function. Activation function decides, whether

a neuron should be activated or not by calculating weighted sum and further adding bias with it.The purpose of the activation function is to introduce non-linearity into the output of

a neuron.

Explanation :-

We know, neural network has neurons that work in correspondence of weight, bias and their respective activation function. In a neural network, we would update the weights and

biases of the neurons on the basis of the error at the output. This process is known as back-propagation. Activation functions make the back-propagation possible since the gradients

are supplied along with the error to update the weights and biases.

Why do we need activation function :-

A neural network without an activation function is essentially just a linear regression model. The activation function does the non-linear transformation to the input making it

capable to learn and perform more complex tasks.It maps the resulting values in between 0 to 1 or -1 to 1 etc.(depending upon the function). Neural networks have to implement complex

mapping functions hence they need activation functions that are non-linear in order to bring in the much needed non-linearity property that enables them to approximate any function.

A neuron without an activation function is equivalent to a neuron with a linear activation function given by.

why use softmax and not sigmoid in cnn?

>>>>>

The soft-max layer outputs a probability distribution, i.e. the values of the output sum to 1. The sigmoid function outputs marginal probabilities and therefore can be used for

multiple-class classification, when the classes are not mutually exclusive. Additionally the soft-max layer is soft version of the max-output layer so it is differentiable and also

resilient to outliers. A problem with sigmoids is that as you reach saturation (values get close to 1 or 0), the gra.dients vanish. This is detrimental to optimization speed and soft-max

doesn't have this problem. Another interpretation is soft-max as a generalization of sigmoid, actually when there are two classes they are the same.

.

|  |  |
| --- | --- |
| Soft-Max   1. Used for multi-classification in logistic regression model 2. the probabilities sum will be 1 . 3. Used in the different layers of neural networks. 4. The high value will have the higher probability than other values . | sigmoid  1. Used for binary classification in logistic regression model.  2. The probabilities sum need not be 1.  3. Used as activation function while building neural networks.  4. The high value will have the high probability but not the higher. |

g