What are linear models in deep learning?

The term linear model implies that **the model is specified as a linear combination of features**. Based on training data, the learning process computes one weight for each feature to form a model that can predict or estimate the target value.

Normalization is **a data pre-processing tool used to bring the numerical data to a common scale without distorting its shape**. Generally, when we input the data to a machine or deep learning algorithm we tend to change the values to a balanced scale.

Normalization is an approach which is applied during the preparation of data in order to change the values of numeric columns in a dataset to use a common scale when the features in the data have different ranges.

Batch Normalization

The advantages of batch normalization are mentioned below:

* Batch normalization reduces the internal covariate shift (ICS) and accelerates the training of a deep neural network
* This approach reduces the dependence of gradients on the scale of the parameters or of their initial values which result in higher learning rates without the risk of divergence
* Batch Normalisation makes it possible to use saturating nonlinearities by preventing the network from getting stuck in the saturated modes

Weight Normalization

* Weight normalization improves the conditioning of the optimisation problem as well as speed up the convergence of stochastic gradient descent.
* It can be applied successfully to recurrent models such as LSTMs as well as in deep reinforcement learning or generative models

Layer Normalization

The advantages of layer normalization are mentioned below:

* Layer normalization can be easily applied to recurrent neural networks by computing the normalization statistics separately at each time step
* This approach is effective at stabilising the hidden state dynamics in recurrent networks

Group Normalization

The advantages of group normalization are mentioned below:

* It has the ability to replace batch normalization in a number of deep learning tasks
* It can be easily implemented in modern libraries with just a few lines of codes

Instance Normalization

* This normalization simplifies the learning process of a model.
* The instance normalization can be applied at test time.