BATCH NORMALIZATION

Agenda

- What is Batch Normalization?
- Principle of BN
- How does Batch Normalization work?
- When and How to use BN?
- How to evaluate BN results?
- Regularization in BN and Normalization of evaluation
- Hypothesis in BN
- Why is this method so important?
- What are the side-effects of BN?
- Advantages of Batch Normalization

What is batch normalization ?

What is Batch Normalization ?

- Normalization is a pre-processing tool for the data that without disturbing the shape of the data it bring the numerical data to the common scale.
- Batch normalization is a process which is used to make the neural networks more stable and faster.
- For making the neural networks become more stable and faster, in a deep neural network extra layers are added.

How does batch normalization Work ?

How does Batch Normalization work ?

The points for the working of batch normalization are mentioned below :

a. By subtracting the layer input by the mini batch mean and then dividing it by the mini batch standard deviation, the batch normalization normalises the layers

Mini batch means a subset for the complete training data.

When and How to use BN ?

When and How to use BN

When you want to normalize the output of the previous layers, then the Batch Normalization is used.

Learning becomes more efficient while using Batch Normalization, so to avoid Overfitting it can be used as the regularizer.

For standardizing the input and output, the layer is added to the sequential Model.

How to evaluate BN results ?

How to evaluate BN results ?

A transformation is applied in Batch normalization which maintains either of
The two results :

a. The mean output close to 0

b. The output standard deviation close to 1

And the BN works differently during inference and during training

Regularization and Normalization in BN

Regularization and Normalization in BN

Convergence and Generalization in training neural networks can eb improved using the BN.
BN is analyzed by a basic block which consists of a kernel layer, a nonlinear activation function and a BN layer

Three aspects are used for understanding these phenomenon :

a. BN is viewed as an implicit regularizer, and as an explicit regularization BN can be decomposed into PN and gamma decay.

b. The regularization shows the converged training with effective and large max learning rate

c. By using statistical mechanics, generalization of Bn is explored

Why is this method so important ?

Why is this method so important ?

This method is important for the following reasons :

It helps to solve a problem know as internal covariate shift.

It helps to solve the vanishing gradient problem

For all the layers of neural network it helps to learn at the normalized Rate.

What is the side-effect of BN ?

What is the side-effect of BN ?

Regularization is the side effect of BN

For avoiding the overfitting, you should not rely on BN.

Between the training examples, batch normalization breaks the Independence.

For language modelling and time series tasks, batch normalization requires the special attention.

Advantages of using BN

Advantages of using BN

- By using BN the , it is simplified and easy to create the deeper networks
- By using the BN, the training process speeds up
- By using the BN, its easy to handle the internal covariate shift
- By using the BN, the loss parameter gets smoothen.

Summary

Let's take a quick recap

Thank You