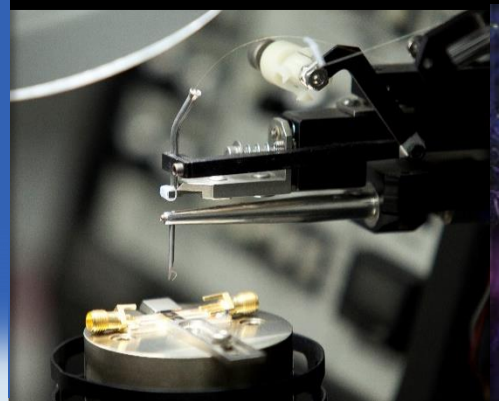
Centro Brasileiro de Pesquisas Físicas

Redes Neurais profundas e aplicações Deep Learning

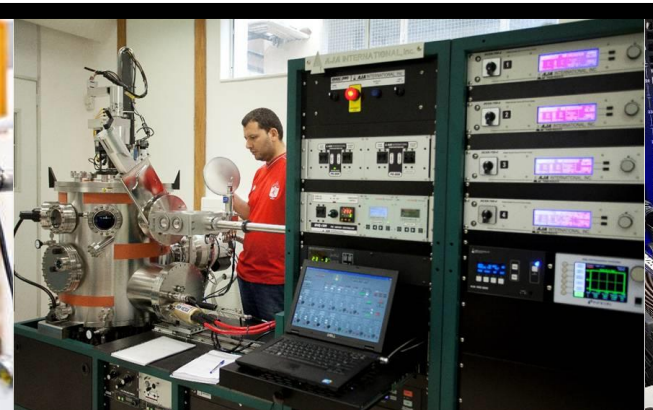
Clécio Roque De Bom – debom@cbpf.br

CBPF

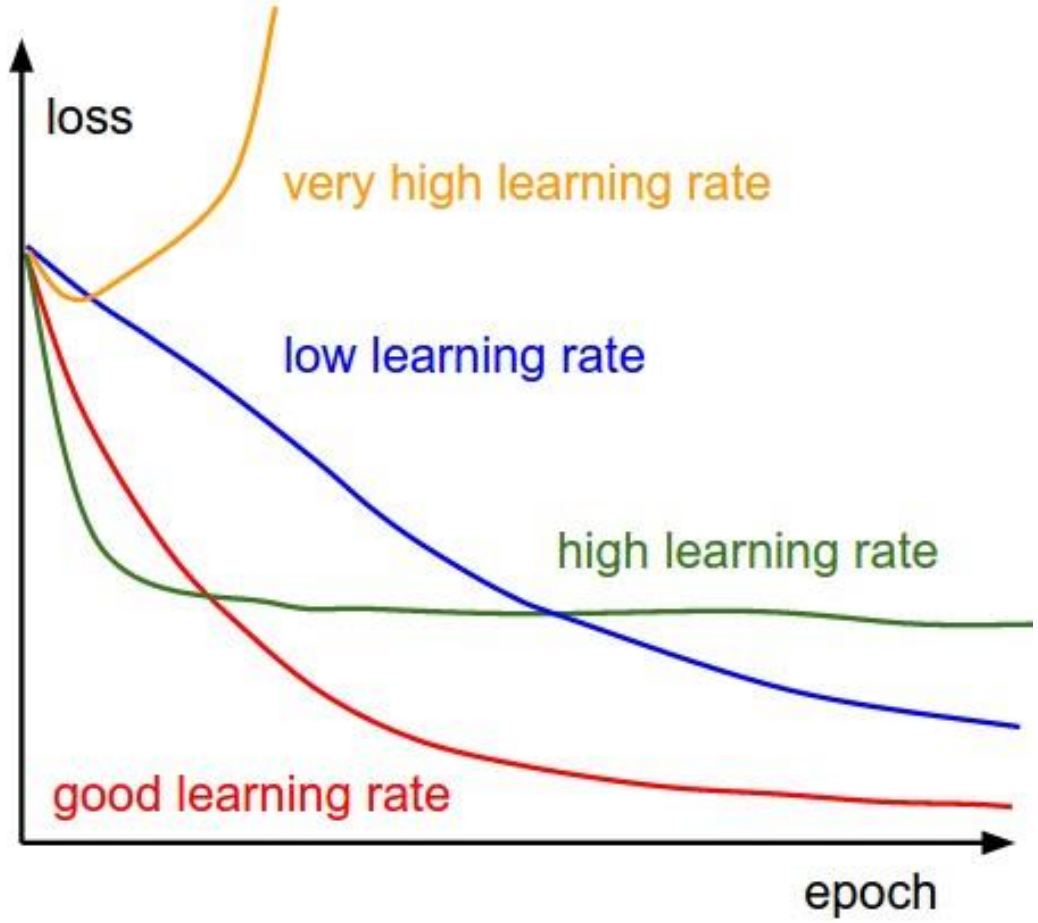
clearnightsrthebest.com





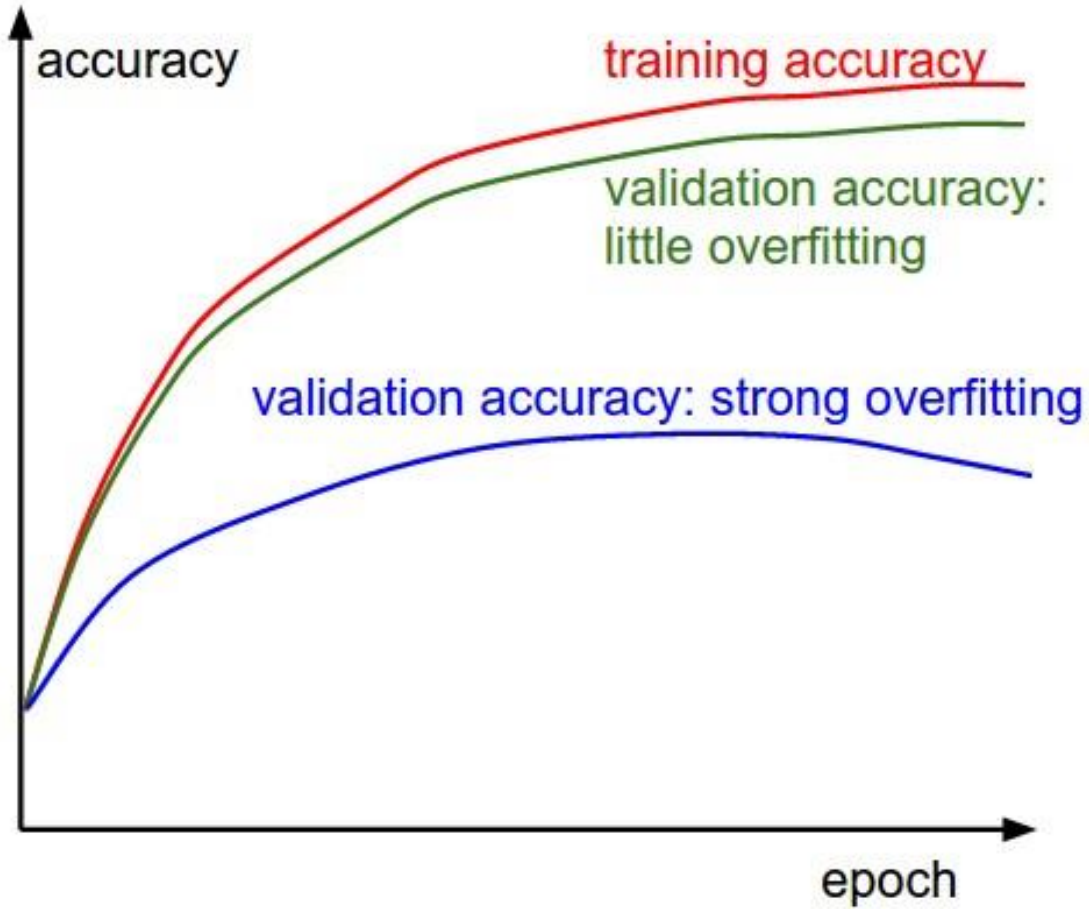



Loss Intuition



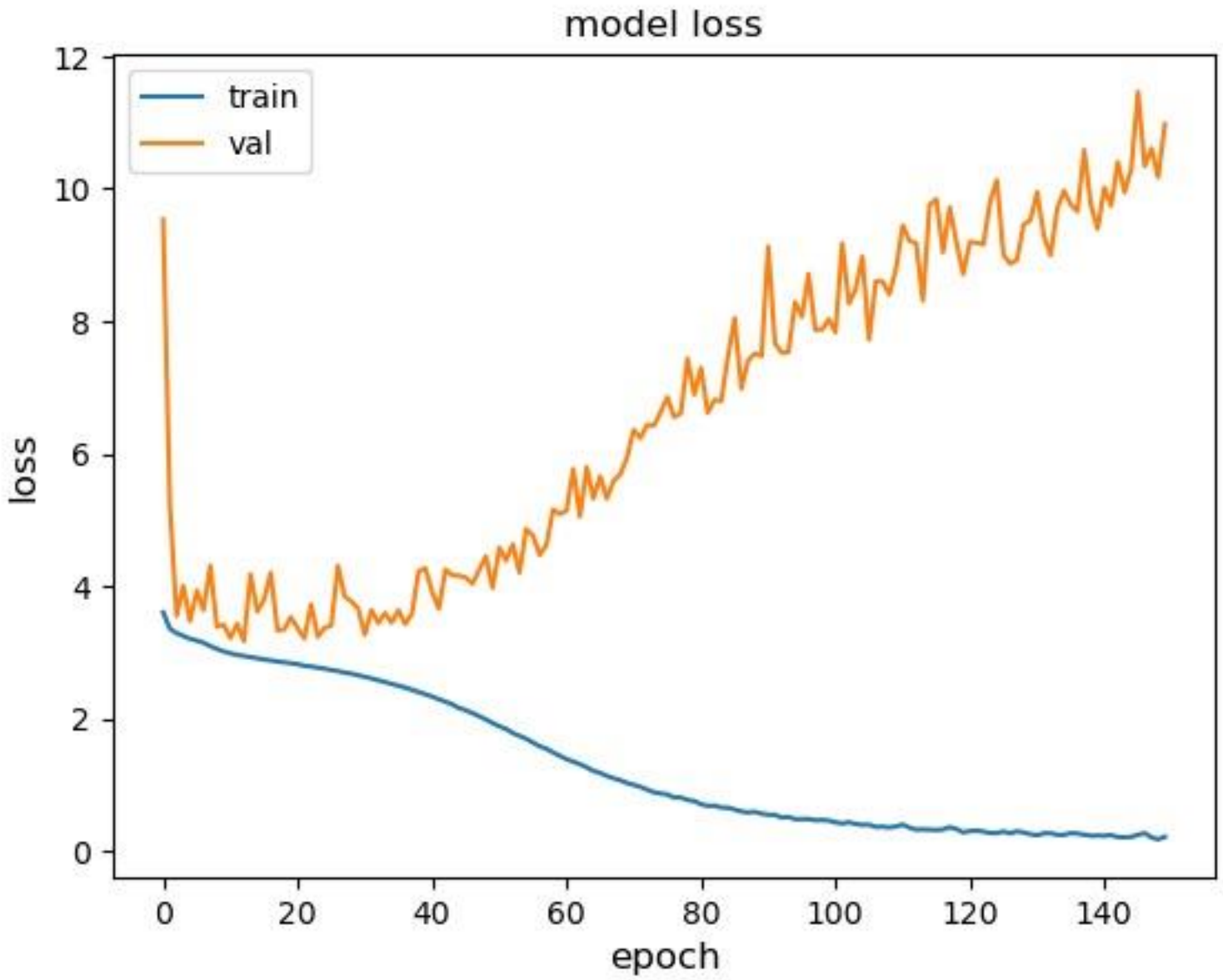
Source: http://cs231n.github.io/neural-networks-3/#baby

Loss Intuition



Source: http://cs231n.github.io/neural-networks-3/#baby

Loss Intuition



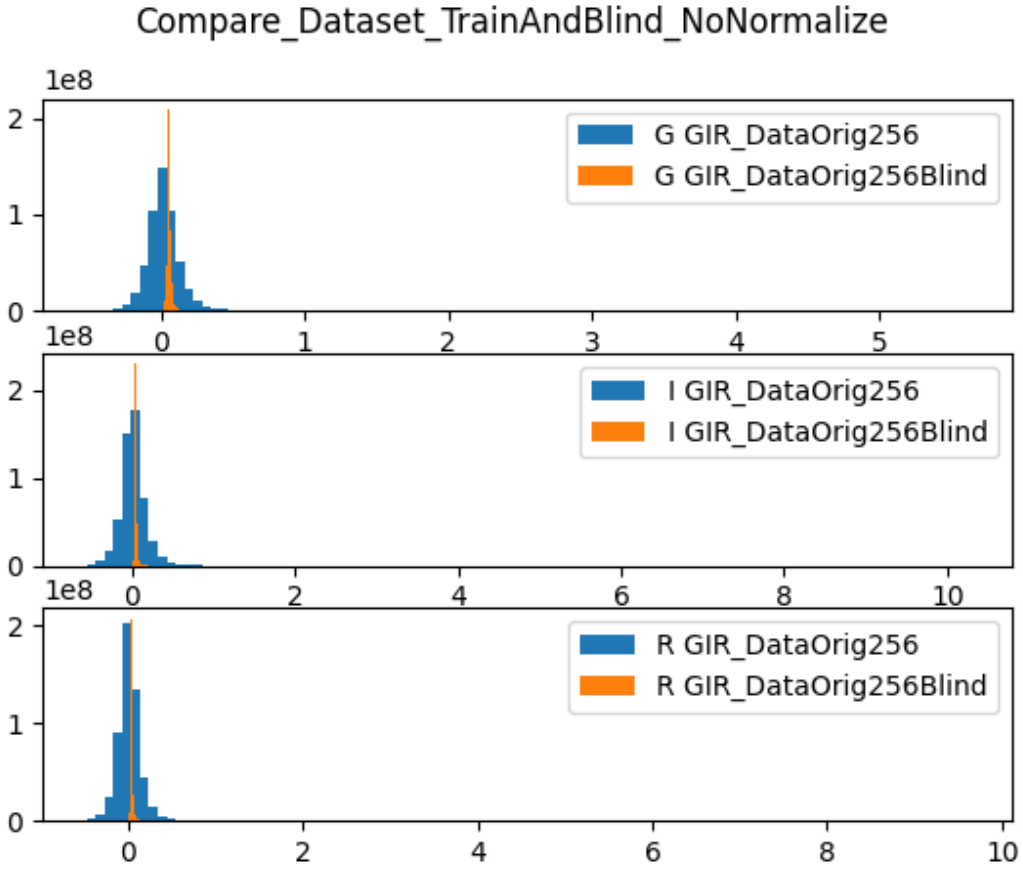
It overfits! What else can I do? The Loss is high/negative? What did I Missed?

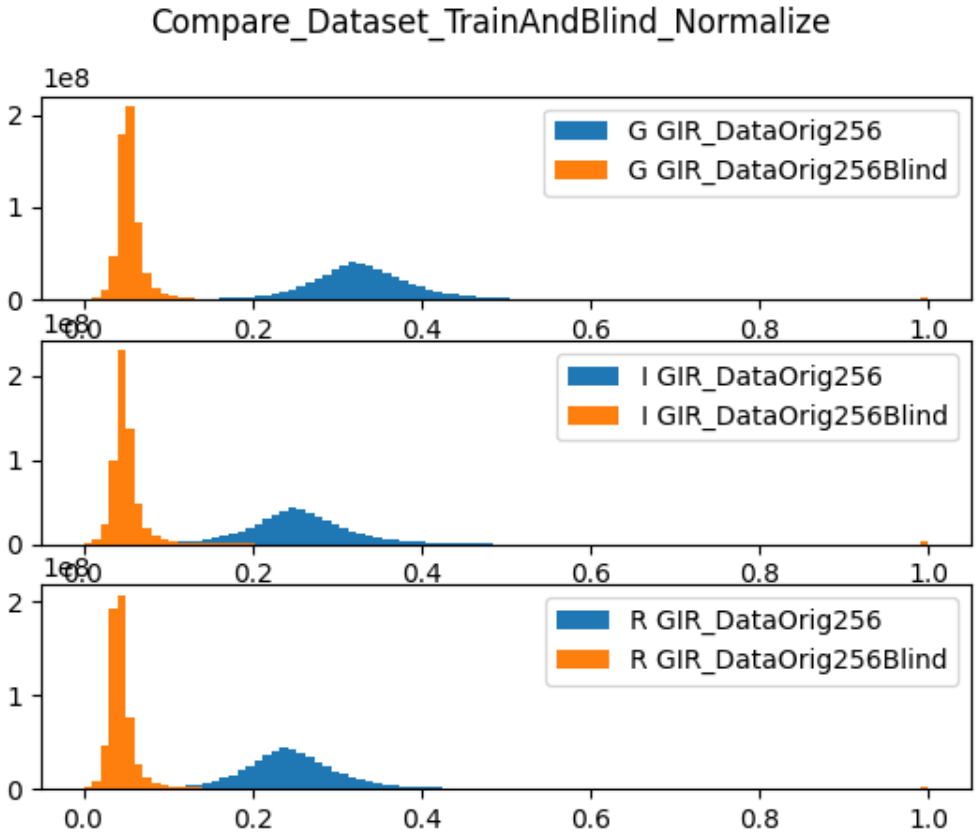
Be aware on what you want the DNN to outputs. E.g. If you want the AE to outputs a number between 0 and 255 you can not let the last layer be sigmoid/softmax since it outputs a number in the interval [0,1[.

Possible strategies:

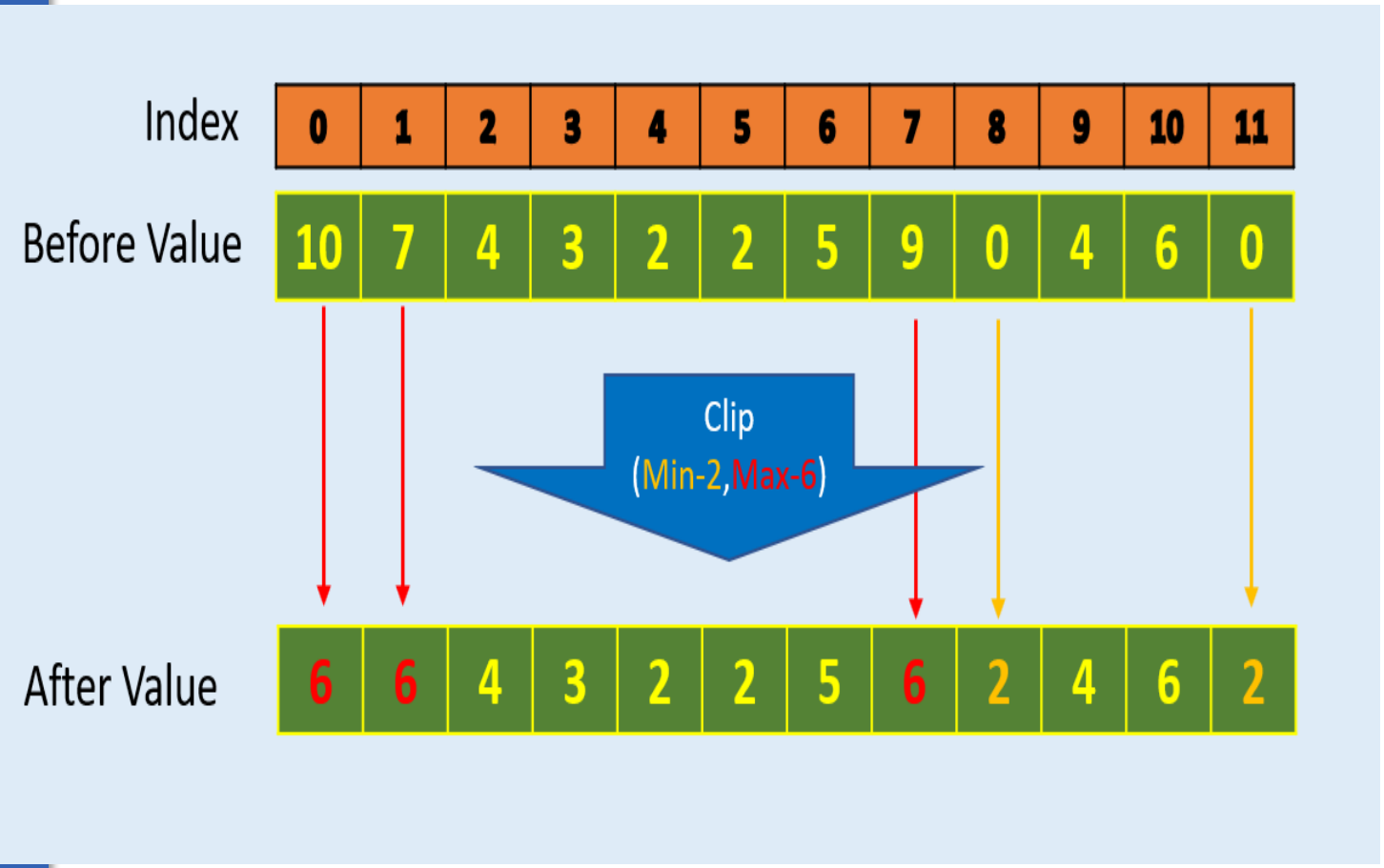
- Normalize your data
- Change your activation functions

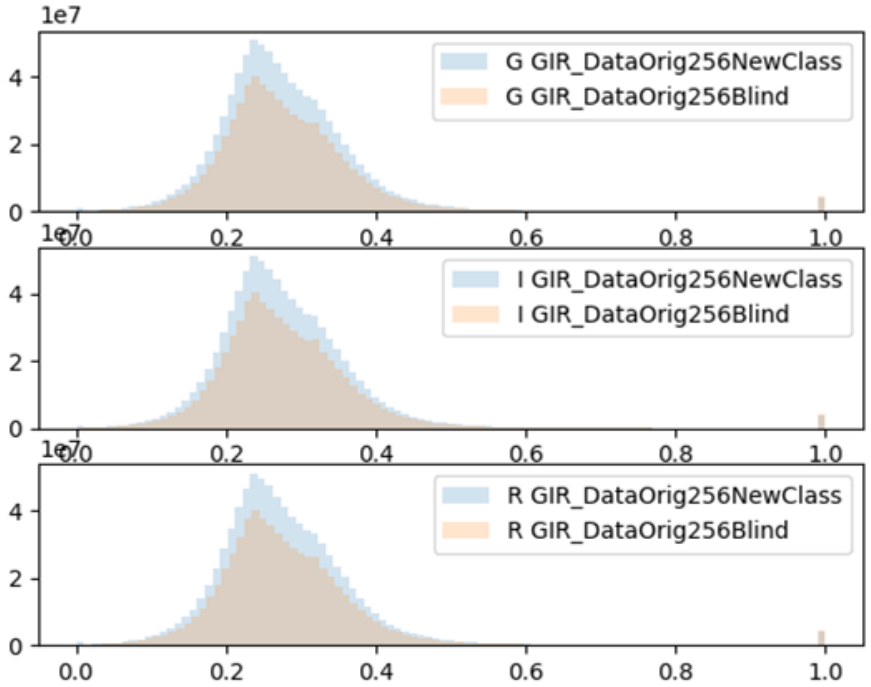
Data Assessment





Data Assessment





It overfits! What else can I do? The Loss is high/negative? What did I Missed?

Be aware on what you want the DNN to outputs. E.g. If you want the AE to outputs a number between 0 and 255 you can not let the last layer be sigmoid/softmax since it outputs a number in the interval [0,1[.

Possible strategies:

- Normalize your data
- Change your activation functions

My loss results does not make sense!

Take a closer look, does your loss fits into your problem? For instance, a loss binary_cross_entropy can not be use in a categorical one. Are your inputs in the categorical format?
Are you trying to make regression using loss function tuned for probability? E.g. Binary_cross_entropy assumes a probability distribution rather than a unbounded variable. Try rms instead

It overfits! What else can I do? The Loss is high/negative? What did I Missed?

Be aware on what you want the DNN to outputs. E.g. If you want the AE to outputs a number between 0 and 255 you can not let the last layer be sigmoid/softmax since it outputs a number in the interval [0,1[.

Possible strategies:

- Normalize your data
- Change your activation functions

My loss results does not make sense!

Take a closer look, does your loss fits into your problem? For instance, a loss binary_cross_entropy can not be use in a categorical one. Are your inputs in the categorical format?
Are you trying to make regression using loss function tuned for probability? E.g. Binary_cross_entropy assumes a probability distribution rather than a unbounded variable. Try rms instead

It overfits! What else can I do? The Loss is high/negative? What did I Missed?

Be aware on what you want the DNN to outputs. E.g. If you want the AE to outputs a number between 0 and 255 you can not let the last layer be sigmoid/softmax since it outputs a number in the interval [0,1[.

Possible strategies:

- Normalize your data
- Change your activation functions

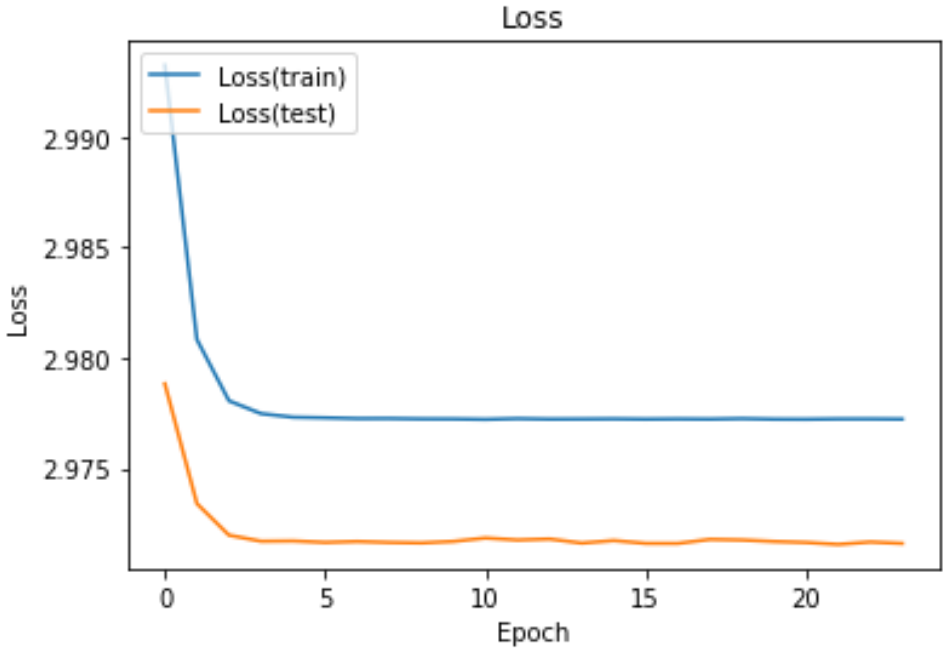
My loss results does not make sense!

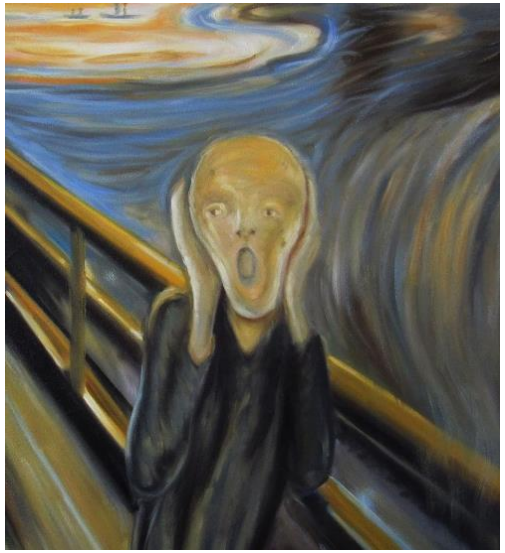
Take a closer look, does your loss fits into your problem? For instance, a loss binary_cross_entropy can not be use in a categorical one. Are your inputs in the categorical format?
Are you trying to make regression using loss function tuned for probability? E.g. Binary_cross_entropy assumes a probability distribution rather than a unbounded variable. Try rms instead

My validation accuracy is higher than the training accuracy

That is very sensible....

To be sure about this statement first one must do a k-fold cross validation





Check in particular Accuracy instead of loss that can give you more intuition.

Smile credit: Pinterest

"The Scream, Munch"

My validation accuracy is higher than the training accuracy

If the problem persists, you exclude the two initial explanations:

- Training set had too many 'hard' cases to learn
- Validation set had mostly 'easy' cases to predict

Possible scenarios on a training:

Underfitting , i.e. Validation and training error high.

Overfitting, i.e. Validation error is high and training error low.

Resonable/fair fit, i.e. Validation error low, slightly higher than the training error.

Unknown fit, i.e. Validation error low, training error high.

“O desespero”, Munch.

My validation accuracy is higher than the training accuracy

Check your dropouts! They might be tricking you.
You can also: evaluate the training set after each epoch with the check pointed saved hd5, on this case it should be a exact model used to validation set

If the problem persists:

Check your batch size, it might be too small...

If nothing works ... Go for other lib.

“A melancolia”, Munch.

Centro Brasileiro de Pesquisas Físicas

Redes Neurais profundas e aplicações Deep Learning

Clécio Roque De Bom – debom@cbpf.br

CBPF

clearnightsrthebest.com