Data Science

Keras difference beetween val_loss and loss during training

Asked 3 years, 1 month ago Active 4 months ago Viewed 49k times

23

8

What is the difference between `val_loss` and `loss` during training in Keras?

E.g.

```
Epoch 1/20
1000/1000 [==============================] - 1s - loss: 0.1760, val_loss: 0.2032
```

On some sites I read that on validation, dropout was not working.

machine-learning deep-learning keras

Share Improve this question Follow

edited Dec 31 '17 at 13:15
Neil Slater
24.5k 3 59 87

asked Nov 30 '17 at 19:33
Vladitapalov
381 1 2 8

What you read about dropout is probably that, when dropout is used (i.e. `dropout` is not `None`), dropout is only applied during training (i.e. no dropout applied during validation). As such, one of the differences between validation loss (`val_loss`) and training loss (`loss`) is that, when using dropout, validation loss can be lower than training loss (usually not expected in cases where dropout is not used). – Psi Aug 27 '19 at 13:01

2 Answers

Active Oldest Votes

19

`val_loss` is the value of cost function for your cross-validation data and loss is the value of cost function for your training data. On validation data, neurons using drop out do not drop random neurons. The reason is that during training we use drop out in order to add some noise for avoiding over-fitting. During calculating cross-validation, we are in the recall phase and not in the training phase. We use all the capabilities of the network.

Thanks to one of our dear friends, I quote and explain the contents from here which are useful.

> `validation_split` : Float between 0 and 1. The fraction of the training data to be used as validation data. The model will set apart this fraction of the training data, will not train on it, and will evaluate the loss and any model metrics on this data at the end of each epoch. The validation data is selected from the last samples in the *x* and *y* data provided, before shuffling.

> `validation_data` : tuple (x_val, y_val) or tuple (x_val, y_val, val_sample_weights) on which to evaluate the loss and any model metrics at the end of each epoch. The model will not be trained on this data. This will override validation_split.

As you can see

```
fit(self, x=None, y=None, batch_size=None, epochs=1, verbose=1, callbacks=None,
validation_split=0.0, validation_data=None, shuffle=True, class_weight=None,
sample_weight=None, initial_epoch=0, steps_per_epoch=None, validation_steps=None)
```

`fit` method used in `Keras` has a parameter named validation_split, which specifies the percentage of data used for evaluating the model which is created after each epoch. After evaluating the model using this amount of data, that will be reported by `val_loss` if you've set verbose to `1`; moreover, as the documentation clearly specifies, you can use either `validation_data` or `validation_split`. Cross-validation data is used to investigate whether your model over-fits the data or does not. This is what we can understand whether our model has generalization capability or not.

Share Improve this answer Follow

edited Jan 21 '20 at 13:17
sophros
199 9

answered Nov 30 '17 at 20:03
Media
12k 9 41 83

This is misleading because it does not have to be about `cross-validation` – HashRocketSyntax Jun 16 '20 at 0:51

Would you elaborate? – Media Jun 16 '20 at 9:43

2

When fitting a model, you have the option to specify a portion of the training dataset that is not trained upon. This is separate from the test dataset.

The "val" in `val_loss` stands for "validation." It's a shame they didn't spell it out.

```
model.fit(validation_split=0.1)
# By default, this float is `0.`
```

Or you can explicitly provide the data to `model.fit(validation_data=<see soruce code below>)` . This would seem to be a good idea if you want a stratified (equally distributed) validation set.

training.py

The use case for this is knowing when your model is appropriately fit to your dataset. Otherwise, you are just looking back and forth at the loss and accuracy of your train and test set wondering if it is balanced.

Share Improve this answer Follow

edited Aug 20 '20 at 20:07
Andrea Blengino
1,005 3 6 19

answered Jun 16 '20 at 1:06
HashRocketSyntax
452 2 12