Cross Validated

How is it possible that validation loss is increasing while validation accuracy is increasing as well

Asked 2 years, 8 months ago Active 18 days ago Viewed 38k times

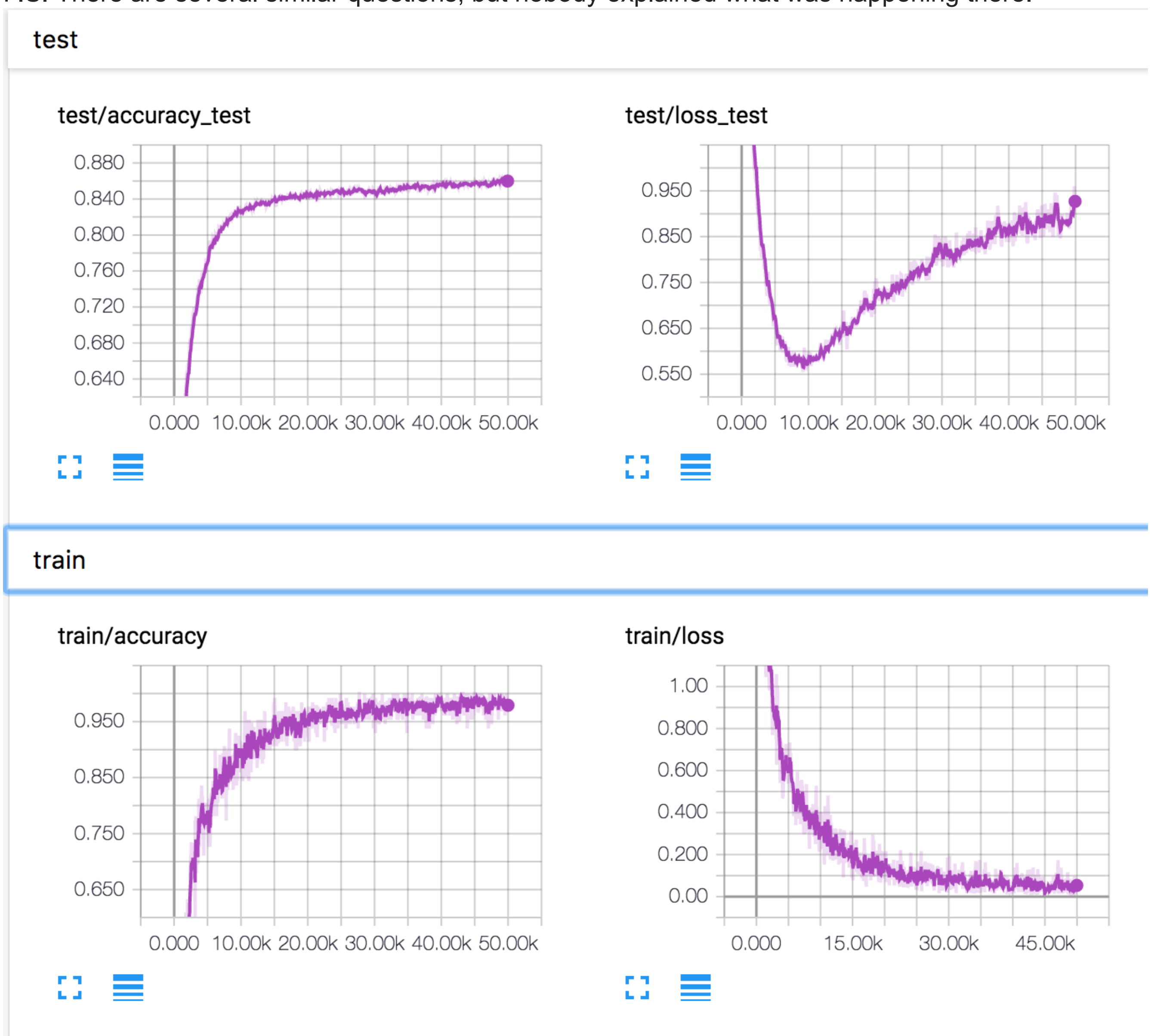
35

23

I am training a simple neural network on the CIFAR10 dataset. After some time, validation loss started to increase, whereas validation accuracy is also increasing. The test loss and test accuracy continue to improve.

How is this possible? It seems that if validation loss increase, accuracy should decrease.

P.S. There are several similar questions, but nobody explained what was happening there.



neural-networks deep-learning conv-neural-network overfitting

edited Jan 3 '19 at 1:08
pglpm
517 4 13

asked May 28 '17 at 14:13
Konstantin Solomatov
483 1 5 8

2 You can check some hints to understand in my answer here: stats.stackexchange.com/questions/258166/… – ahstat May 28 '17 at 14:20

1 @ahstat I understand how it's technically possible, but I don't understand how it happens here. – Konstantin Solomatov May 28 '17 at 15:16

The 'illustration 2' is what I and you experienced, which is a kind of overfitting. For my particular problem, it was alleviated after shuffling the set. – ahstat May 28 '17 at 15:28

@ahstat There're a lot of ways to fight overfitting. For example, I might use dropout. What I am interesting the most, what's the explanation for this. I.e. why is it increasing so gradually and only up. – Konstantin Solomatov May 28 '17 at 17:14

4 Answers

13

Building on Ankur's answer and the comment underneath it, I think the following scenario is possible, while I have no proof of it. Two phenomenons might be happening at the same time :

1. Some examples with borderline predictions get predicted better and so their output class changes (eg a cat image predicted at 0.4 to be a cat and 0.6 to be a horse becomes predicted 0.4 to be a horse and 0.6 to be a cat). Thanks to this, accuracy increases while loss decreases.
2. Some examples with very bad predictions keep getting worse (eg a cat image predicted at 0.8 to be a horse becomes predicted at 0.9 to be a horse) AND/OR (more probable, in particular for multi-class ?) some examples with very good predictions get a little worse (eg a cat image predicted at 0.9 to be a cat becomes predicted at 0.8 to be a cat). With this phenomenon, loss increases while accuracy stays the same.

So if phenomenon 2 kicks in at some point, on lots of examples (eg for a specific class which is not well understood for some reason) and/or with a loss increase stonger than the loss decrease you gain from 1., then you might find yourself in your scenario.
Once again, maybe this is not what's happening, but I think that being able to come up with such scenarios must remind us of the sometimes slippery relationship between (cross-entropy)loss and accuracy.

edited Jul 8 '19 at 13:27

answered Apr 17 '18 at 13:45
Soltius
338 2 9

19

Accuracy of a set is evaluated by just cross-checking the highest softmax output and the correct labeled class.It is not depended on how **high** is the softmax output. To make it clearer, here are some numbers.

Suppose there are 3 classes- dog, cat and horse. For our case, the correct class is **horse** . Now, the output of the softmax is [0.9, 0.1]. For this loss ~0.37. The classifier will predict that it is a horse. Take another case where softmax output is [0.6, 0.4]. Loss ~0.6. The classifier will still predict that it is a horse. But surely, the loss has increased. So, it is all about the output distribution.

answered Dec 30 '17 at 16:05

ANKUR SATYA
191 1 5

4 Observation: in your example, the accuracy doesnt change. It's still 100%. Do you have an example where loss decreases, and accuracy decreases too? – Hugh Perkins Dec 30 '17 at 17:04

2

From Ankur's answer, it seems to me that:

> **Accuracy** measures the percentage *correctness* of the prediction i.e. $\frac{correct-classes}{total-classes}$

while

> **Loss** actually tracks the *inverse-confidence* (for want of a better word) of the prediction. A high *Loss* score indicates that, even when the model is making good predictions, it is $less$ sure of the predictions it is making...and vice-versa.

So...

> **High Validation Accuracy** + **High Loss Score** vs **High Training Accuracy** + **Low Loss Score** suggest that the model may be over-fitting on the training data.

answered Feb 23 '19 at 13:19

Ignatius Ezeani
21 2

2

Many answers focus on the mathematical calculation explaining how is this possible. But they don't explain why it becomes so. And they cannot suggest how to digger further to be more clear.

I have 3 hypothesis. And suggest some experiments to verify them. Hopefully it can help explain this problem.

1. Label is noisy. Compare the false predictions when val_loss is minimum and val_acc is maximum. Check whether these sample are correctly labelled.
2. [Less likely] The model doesn't have enough aspect of information to be certain. Experiment with more and larger hidden layers.
3. [A very wild guess] This is a case where the model is less certain about certain things as being trained longer. Such situation happens to human as well. When someone started to learn a technique, he is told exactly what is good or bad, what is certain things for (high certainty). When he goes through more cases and examples, he realizes sometimes certain border can be

> blur (less certain, higher loss), even though he can make better decisions (more accuracy). And he may eventually gets more certain when he becomes a master after going through a huge list of samples and lots of trial and errors (more training data). So in this case, I suggest experiment with adding more noise to the training data (not label) may be helpful.

Don't argue about this by just saying if you disagree with these hypothesis. It will be more meaningful to discuss with experiments to verify them, no matter the results prove them right, or prove them wrong.

edited Jan 17 at 7:44

answered Sep 3 '19 at 8:06

Diansheng
119 4

1 have this same issue as OP, and we are experiencing scenario 1. – teng Jan 16 at 18:03