

باسمه تعالی



دانشگاه صنعتی شریف

دانشکده مهندسی برق

درس یادگیری ماشین

گزارش تمرین شماره پنج

علی محرابیان 96102331

استاد: دکتر صالح کلیبر

زمستان 1398



Constant، مقادیری هستند که ثابت می باشند. variable، مقادیری هستند که مقدار آن ها با توجه به آپدیت شدن گراف تغییر می کند. placeholder، مقادیری هستند که در آینده تعریف می شوند و نوع و اندازه آن ها را در ابتدا تعیین می کنیم.

دو تانسور خواسته شده را ساخته و درهم ضرب می کنیم. گراف آنها به صورت زیر است.

Main GraphAuxiliary Nodes



خروجی ضرب به صورت زیر است.

```
[[ 2 21 12]
 [ 6 36 54]
 [ 6  0  8]]
```



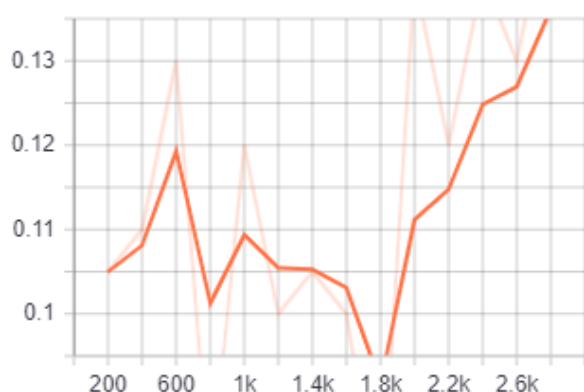
دیتاست mnist را دانلود می کنیم.

در شبکه عصبی یک لایه، شبکه را با داده های train آموزش می دهیم.

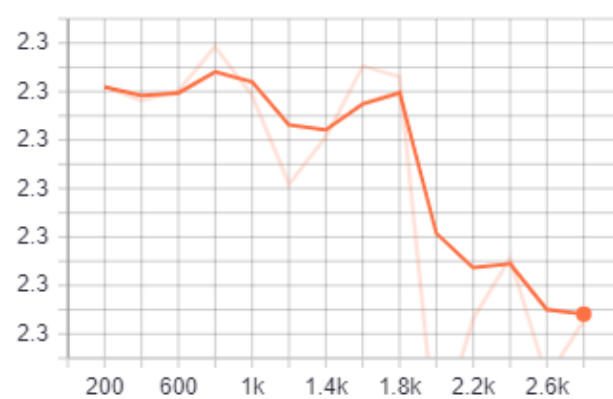
برای انحراف معیار 0:

Accuracy on test set: 0.1135
Loss on test set: 2.3010364

Accuracy



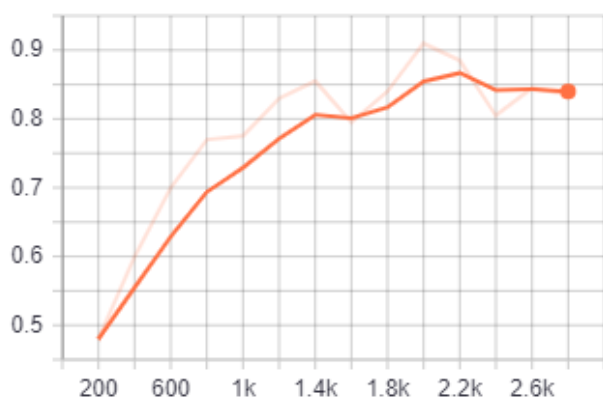
Loss



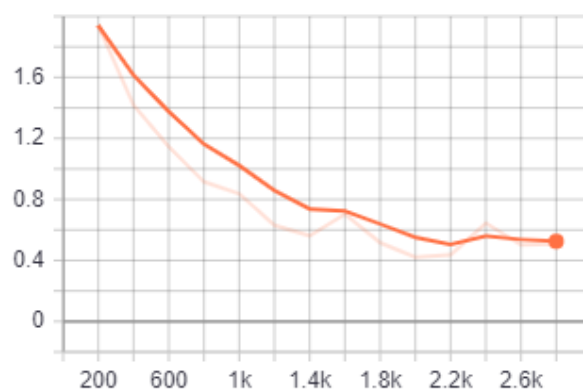
Accuracy on test set: 0.8878
Loss on test set: 0.39976

برای انحراف معیار 0.1:

Accuracy



Loss



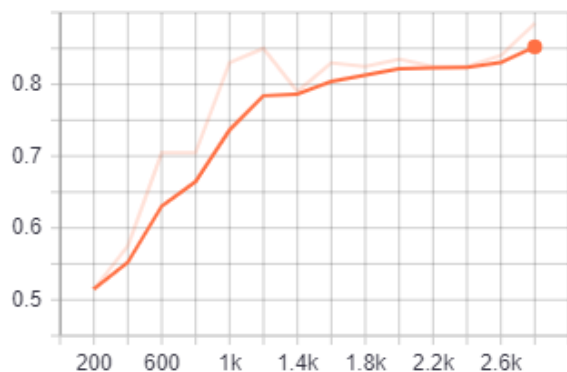


برای انحراف معیار 0.3:

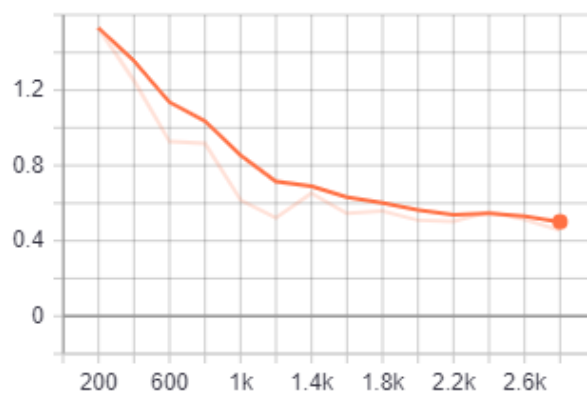
Accuracy on test set: 0.8536

Loss on test set: 0.48152286

Accuracy



Loss

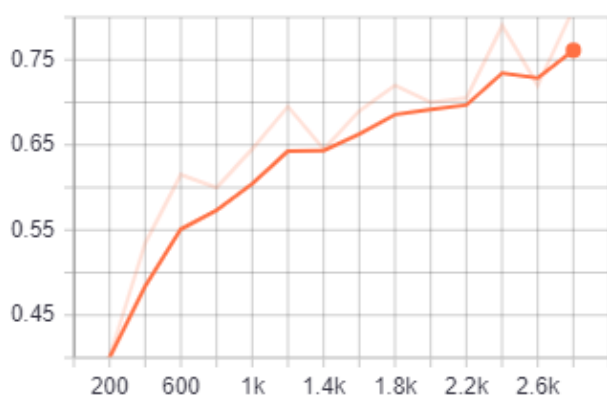


برای انحراف معیار 0.8:

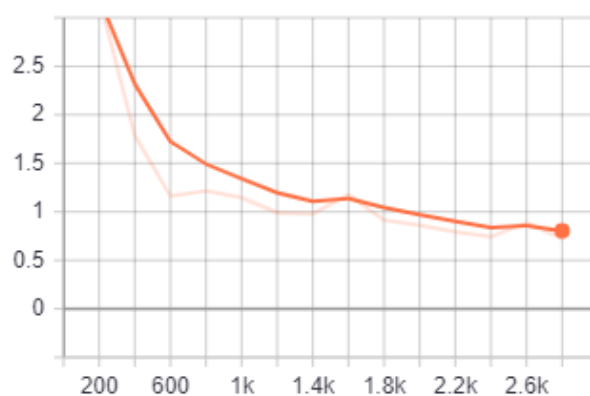
Accuracy on test set: 0.7676

Loss on test set: 0.72626936

Accuracy



Loss



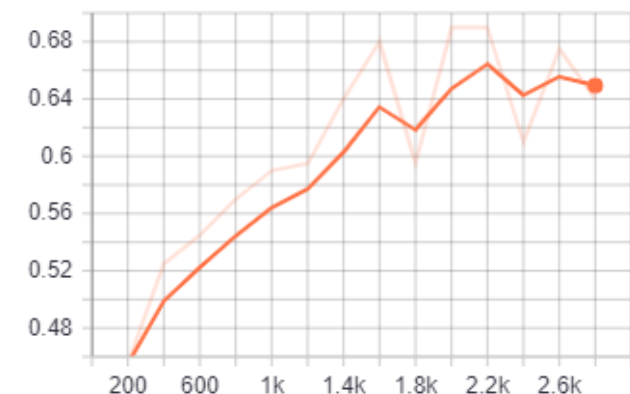


برای انحراف معیار 1.5:

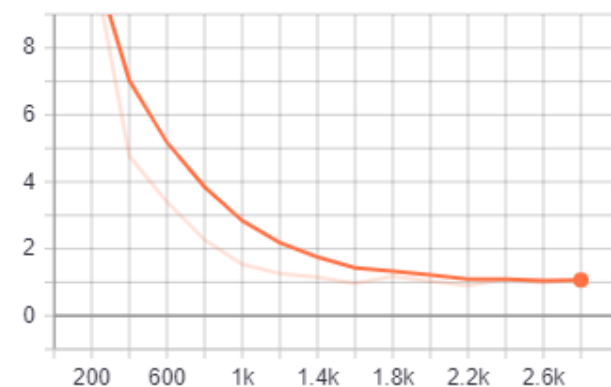
Accuracy on test set: 0.6871

Loss on test set: 0.9348227

Accuracy



Loss

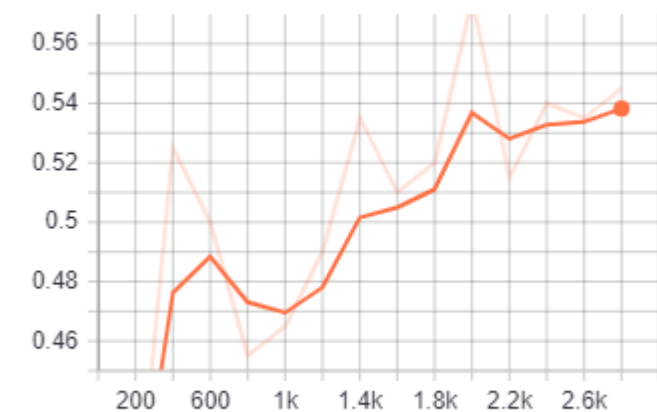


برای انحراف معیار 3:

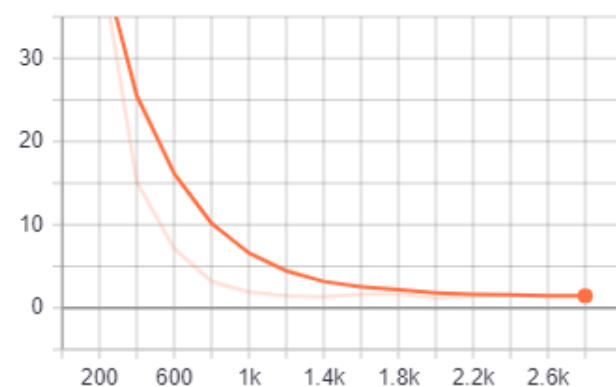
Accuracy on test set: 0.5717

Loss on test set: 1.3035204

Accuracy



Loss



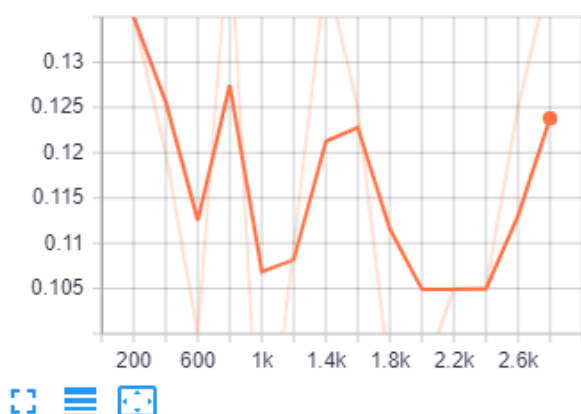


مشاهده می شود که هرچه انحراف معیار بیشتر می شود، دقت بر روی داده های تست کمتر می شود و خطای آن ها بیشتر می شود. هم چنین این که خطای اولیه هنگام آموزش افزایش می یابد.

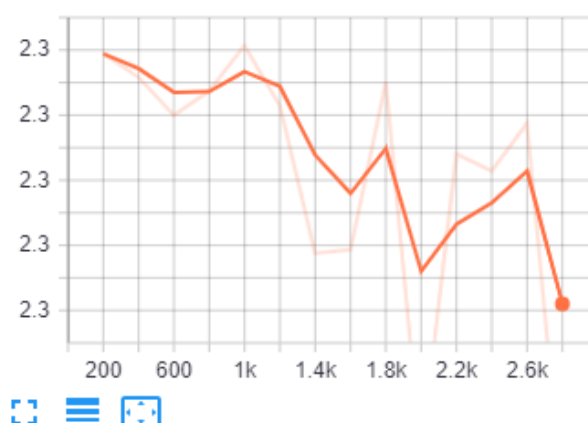
در شبکه عصبی با یک لایه مخفی، وزن ها را با انحراف معیارهای مختلف مقداردهی می کنیم. تمامی موارد زیر پس از آموزش بر روی داده های validation هستند.
برای انحراف معیار 0:

Accuracy on test set: 0.1135
Loss on test set: 2.301048

Accuracy



Loss



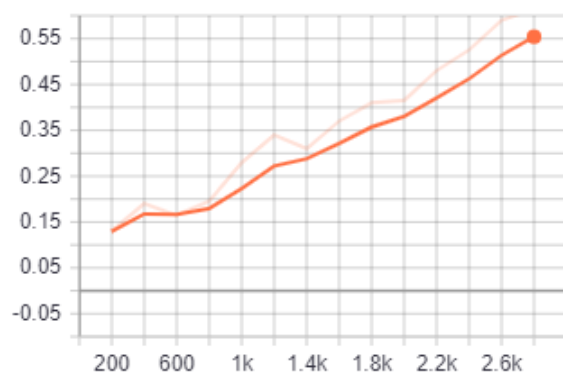


برای انحراف معیار 0.1:

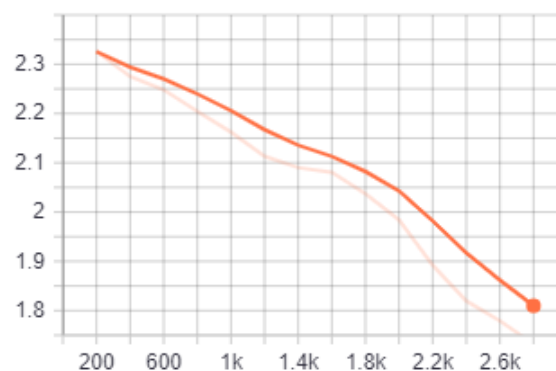
Accuracy on test set: 0.583

Loss on test set: 1.7010049

Accuracy



Loss

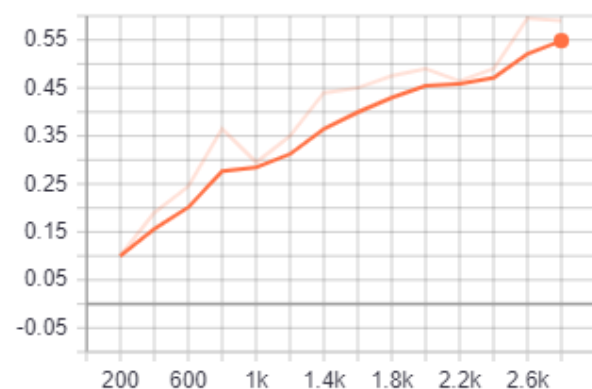


برای انحراف معیار 0.3:

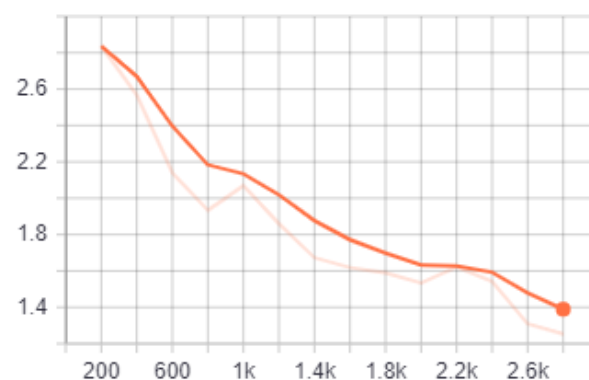
Accuracy on test set: 0.5472

Loss on test set: 1.3821694

Accuracy



Loss

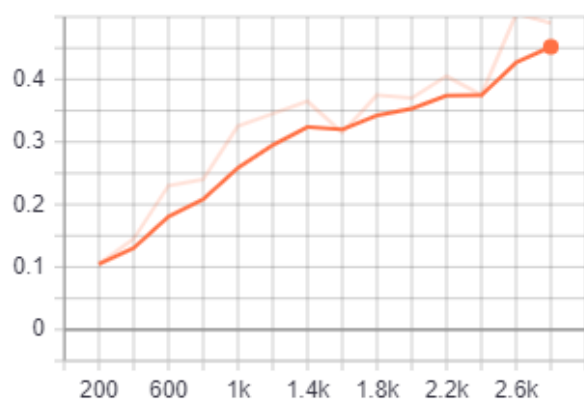




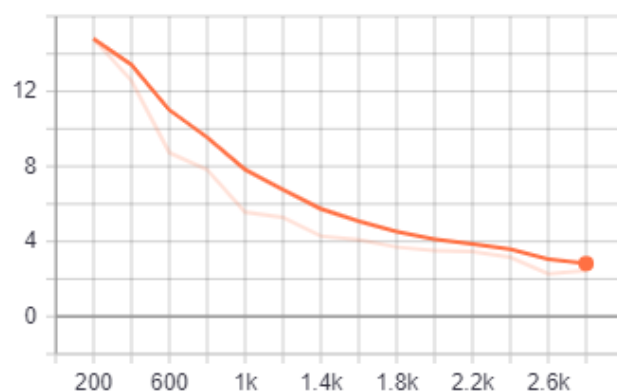
برای انحراف معیار 0.8:

Accuracy on test set: 0.485
Loss on test set: 2.3347065

Accuracy



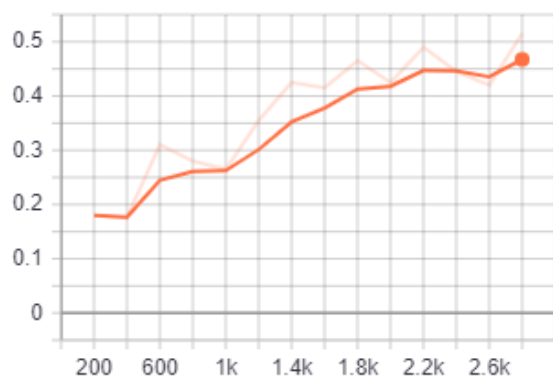
Loss



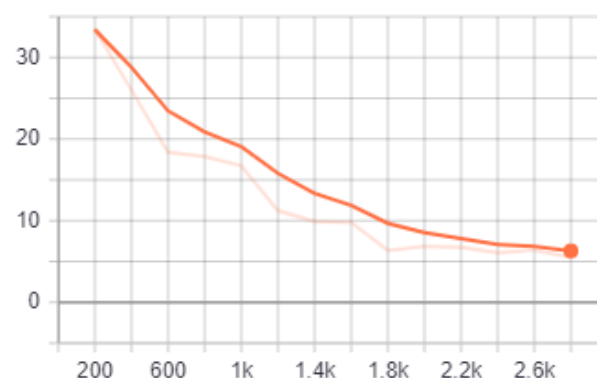
برای انحراف معیار 1.5:

Accuracy on test set: 0.504
Loss on test set: 5.0438156

Accuracy



Loss

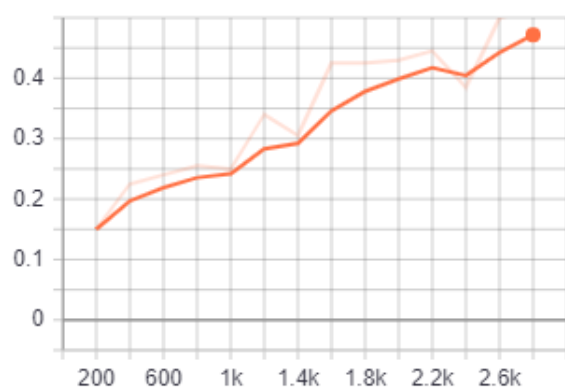




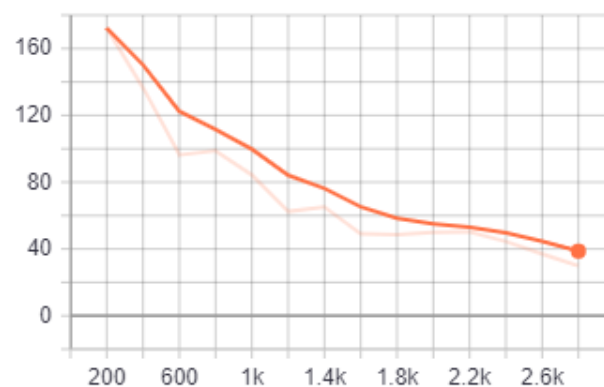
انحراف معیار 3:

Accuracy on test set: 0.4946
Loss on test set: 29.71853

Accuracy



Loss



همان طور که مشاهده می شود، با افزایش انحراف معیار، خطای اولیه برای داده های آموزش بسیار زیاد بوده و دقت نهایی نیز کمتر است.

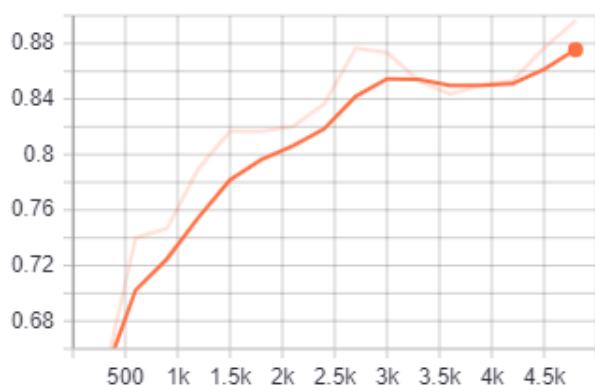


برای جلوگیری از مفصل شدن گزارش، برای شبکه عصبی دو لایه، داده های train فقط برای سه انحراف معیار و validation برای یک انحراف معیار رسم می کنیم. موارد زیر برای داده های train هستند.

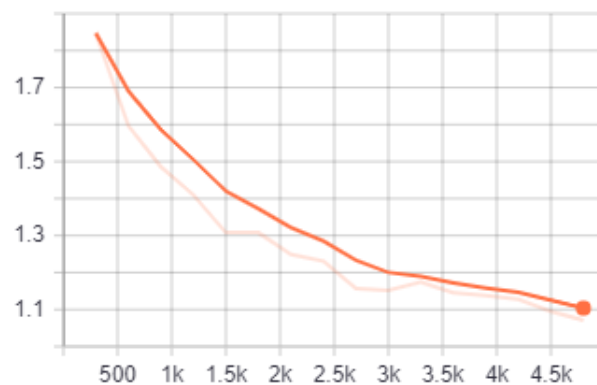
برای انحراف معیار 0.1:

Accuracy on test set: 0.8868
Loss on test set: 1.0689563

Accuracy



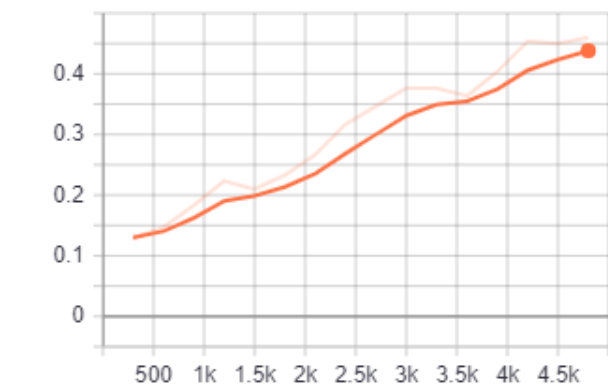
Loss



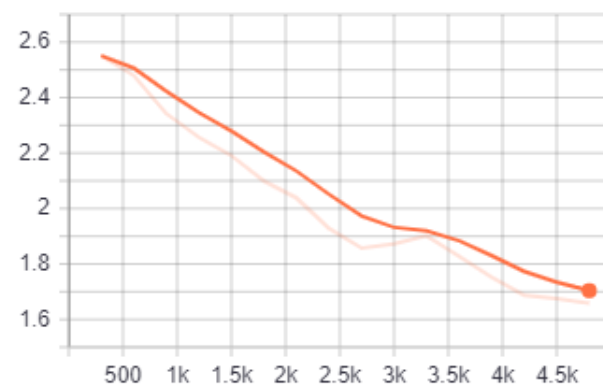
برای انحراف معیار 0.8:

Accuracy on test set: 0.497
Loss on test set: 1.5961471

Accuracy



Loss

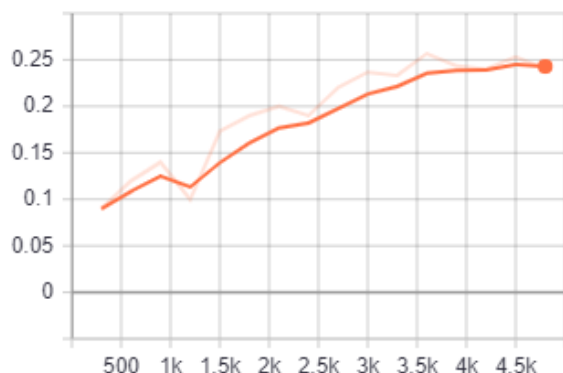




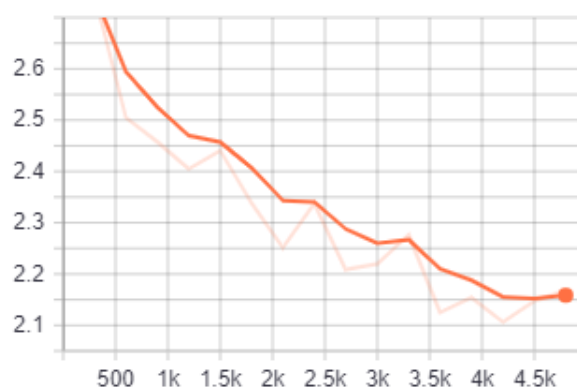
Accuracy on test set: 0.2673
Loss on test set: 2.0969214

برای انحراف معیار 1.5:

Accuracy



Loss

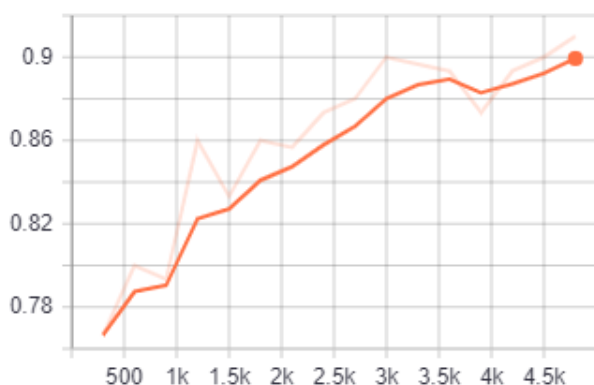


به طور کلی موارد حالت قبل برقرار هستند. به طور میانگین، از مقایسه دو شبکه می توان نتیجه گرفت که شبکه اول خروجی دقیق تری به ما می دهد.

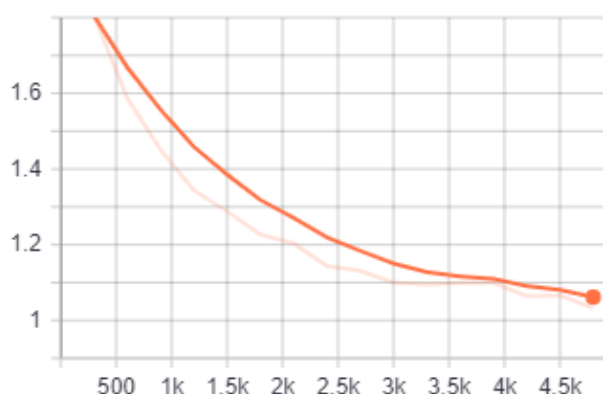
Accuracy on test set: 0.8925
Loss on test set: 1.0619289

برای داده های validation و انحراف معیار 0.1:

Accuracy

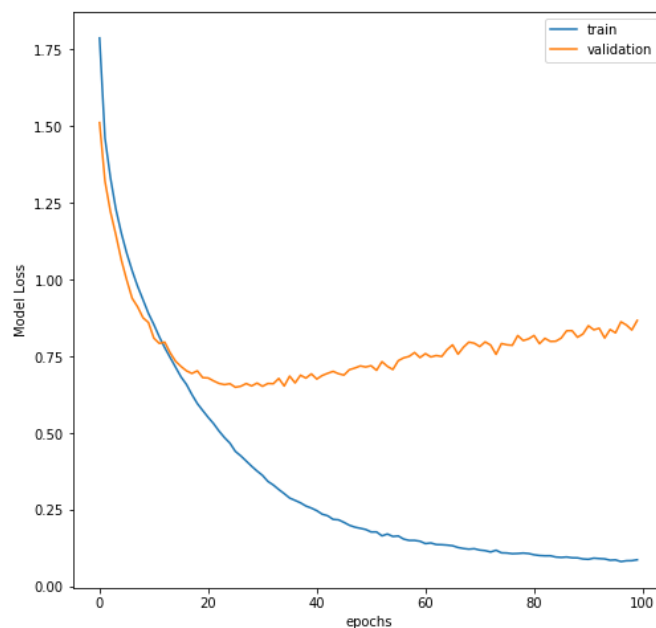
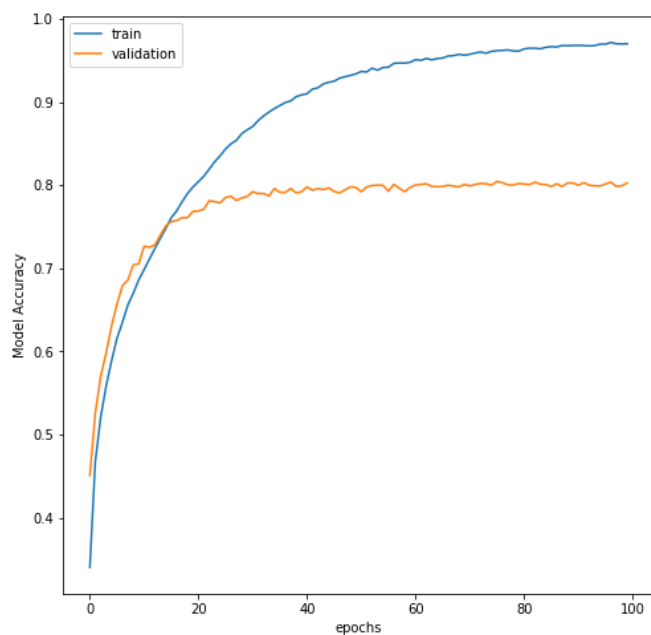


Loss





دیتاست CIFAR داندلود کرده و شبکه خواسته شده را آموزش می دهیم. accuracy و loss به صورت زیر هستند.



برای داده های تست، مقادیر به صورت زیر است.

```
10000/10000 [=====] - 1s 73us/step  
Test loss: 0.8983099212646485  
Test accuracy: 0.7892
```

خوب بودن درصد تست به این معنا است که هنگام training، دچار overfitting نشده ایم.

Callback مجموعه ای از توابع می باشد که در زمان تعیین شده ای هنگام آموزش بر روی داده ها اعمال

می شود. به طور مثال در این قسمت، در epoch های معینی، learning rate را عوض می کنیم تا فرآیند یادگیری بهتر انجام شود.



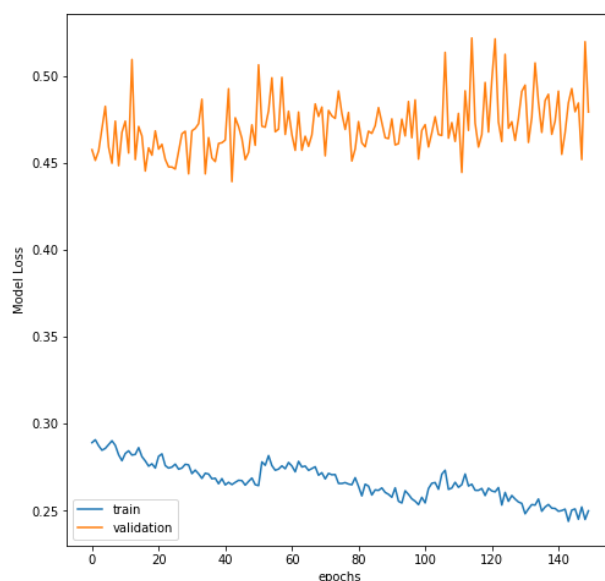
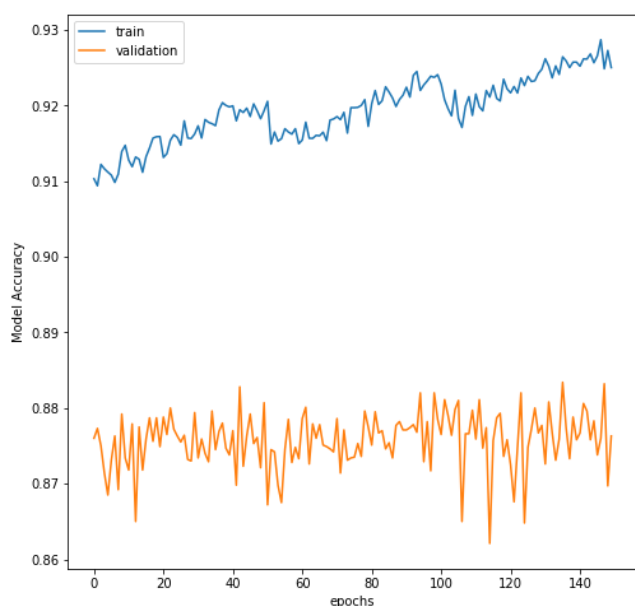
ضریب dropout، به این منظور استفاده می شود که از پیچیدگی شبکه جلوگیری شود. یعنی با احتمال $1-p$ ، بعضی از node ها به طور رندوم در هر مرحله آموزش حذف می شوند.

Normalize کردن دیتا نیز یعنی داده ها منهای میانگین و سپس تقسیم بر واریانس آن ها بکنیم تا از پراکندگی و بزرگ بودن آن ها جلوگیری شود.

Data augmentation، موارد زیادی همچون شیفต์ دادن عکس، دوران عکس به صورت رندوم، scale کردن عکس و مواردی از این قبیل را شامل می شود.

Kernel regularization یعنی این که چه قدر می توانیم خطای قابل صرف نظر داشته باشیم.

دقت و خطا به صورت زیر است.



خروجی بر روی داده های تست به صورت زیر است.

```
10000/10000 [=====] - 1s 121us/step  
Test loss: 0.4792442533493042  
Test accuracy: 0.8763
```