

به نام خدا

گزارش تمرین سری اول یادگیری عمیق

طراحی شبکه عصبی MLP

استاد درس:

دکتر حامد ملک

دانشجو:

طراوت پارت

بهار 1401

## مقدمه:

در این پروژه یک شبکه عصبی چند لایه برای انجام تسک طبقه‌بندی چند کلاسه، با استفاده از فریمورک پایتورچ پیاده‌سازی میشود. در ابتدا بعد از بررسی مجموعه داده، پیش‌پردازش‌های لازم روی آن انجام میشود. سپس یک مدل ابتدایی طراحی میشود و آن مدل را آموزش میدهیم تا از کارکرد درست برنامه نوشته شده مطمئن بشویم. در ادامه با بررسی حالت‌های مختلف وزن‌دهی اولیه، تاثیر تابع‌های فعالیت مختلف و تاثیر تعداد لایه‌ها و نورون‌های مختلف بهترین شبکه عصبی را پیدا میکنیم. همچنین با استفاده از dropout و data augmentation سعی میکنیم جلوی overfit شدن شبکه را بگیریم.

## 1. آماده‌سازی دیتاست:

- تابع `__getitem__(self, index)`:

در این تابع تنها یک داده و لیبل متناظر آن را از دیتاست میخوانیم، پیش‌پردازش‌های مورد نیاز را روی آنها انجام میدهیم و داده‌ی پیش‌پردازش شده و لیبل متناظر را ریترن میکنیم.

- تابع `__len__(self)`:

تعداد داده‌های موجود در دیتاست را ریترن میکند.

- تابع `get_transformation()`:

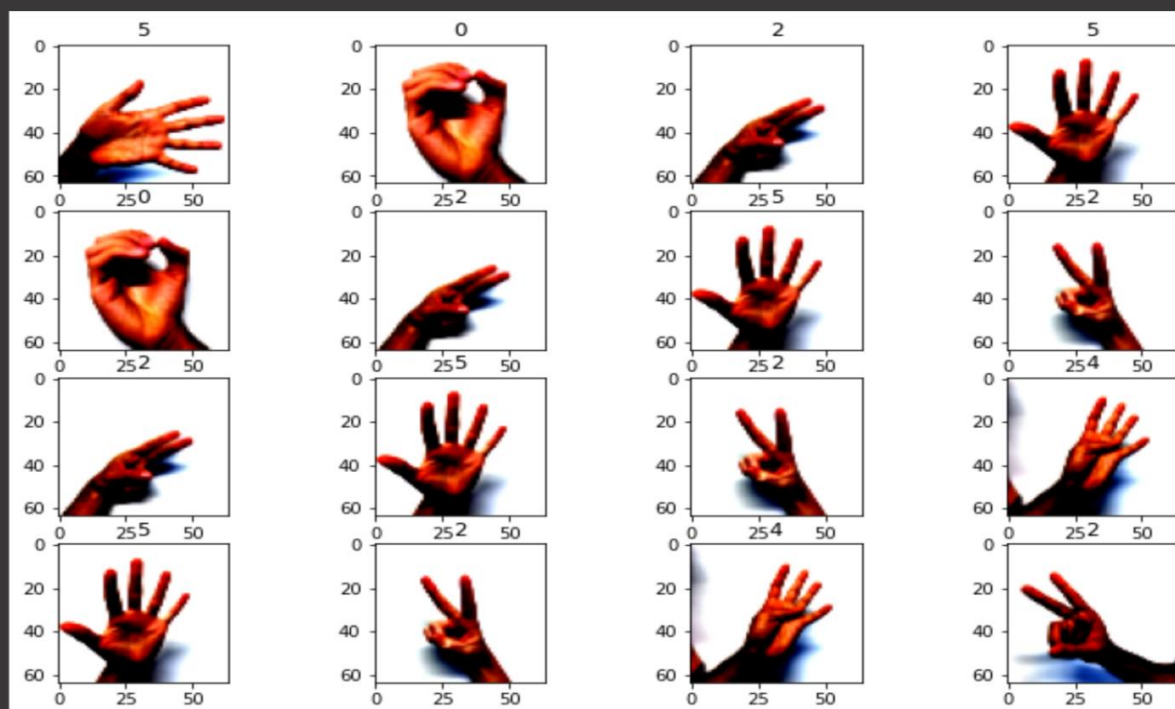
تغییراتی که میخواهیم روی مجموعه داده اعمال کنیم را پیاده‌سازی میکنیم. این تغییرات شامل تبدیل کردن داده به تنسور، نرمالایز کردن داده و `flatten` کردن داده میشود. همچنین در صورتی که نیاز به `data augmentation` داشته باشیم، تغییرات دیگری از جمله `randomRotation` و `Randomhorizontalflip` روی آن اعمال میکنیم.

- تابع `get_one_hot(label, num_classes)`:

این تابع لیبل عددی را ورودی میگیرد و با استفاده از کتابخانه‌های آماده‌ی پایتورچ، `onehot encoding` روی لیبل اعمال میکند.

- تابع `visualize_samples()`:

این تابع تعدادی از داده‌های پردازش شده به همراه لیبل آنها را نمایش میدهد.



شکل 1- خروجی تابع visualize\_samples

## 2. شبکه عصبی پیاده‌سازی شده ابتدایی:

در فایل model.py یک شبکه عصبی ابتدایی پیاده‌سازی شده است. در ادامه با آزمون و خطا پارامترهای این شبکه را تغییر می‌دهیم و بهینه‌ترین شبکه را پیدا می‌کنیم.

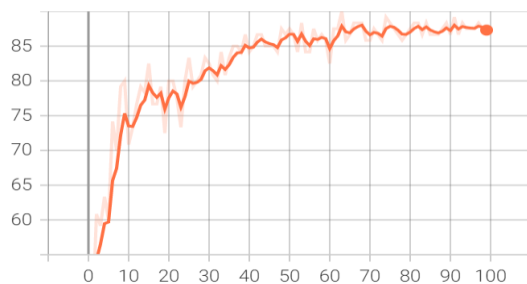
در این قسمت شبکه‌ای با سه لایه Linear که هر کدام به ترتیب شامل 25، 12 و 6 نورون هستند طراحی شده است. تعداد نورون‌ها در لایه آخر برابر تعداد کلاس‌های موجود در دیتاست است. در ادامه‌ی هر کدام از لایه‌های Liner میانی یک تابع فعالیت ReLU قرار گرفته است. نکته قابل توجه این است که در ادامه‌ی لایه خروجی هیچ لایه فعالیت قرار نگرفته است چون خود تابع خطای crossEntropy که از آن استفاده خواهیم کرد، تابع softmax را روی خروجی شبکه اعمال خواهد کرد.

## 3. آموزش شبکه عصبی ابتدایی:

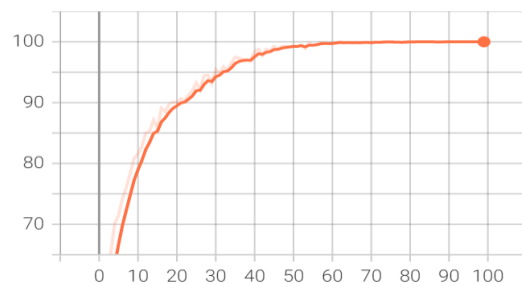
در این قسمت با کامل کردن فایل train.py، شبکه طراحی شده ابتدایی را آموزش می‌دهیم. برای آموزش از تابع خطای crossEntropy و بهینه‌سازی SGD استفاده شده است. تعداد اپاک برابر 150، نرخ آموزش برابر 0.001، اندازه بچ برابر 8 در نظر گرفته شده است.

## Accuracy

Accuracy/test  
tag: Accuracy/test

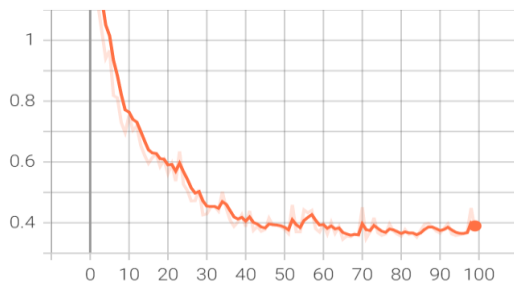


Accuracy/train  
tag: Accuracy/train

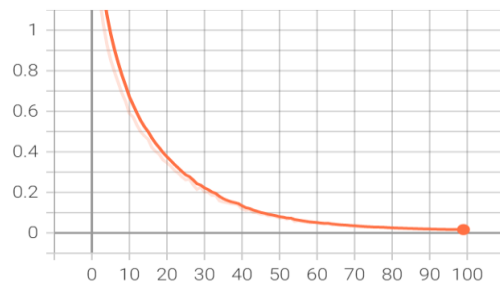


## Loss

Loss/test  
tag: Loss/test



Loss/train  
tag: Loss/train



شکل 2- نمودار های دقت و خطای آموزش شبکه عصبی ابتدایی

```
Epoch [1/100]: Train Loss: 1.6834, Train Accuracy: 31.1111, Test Loss: 1.4195, Test Accuracy:50.0000
Epoch [11/100]: Train Loss: 0.5897, Train Accuracy: 81.5741, Test Loss: 0.7529, Test Accuracy:70.8333
Epoch [21/100]: Train Loss: 0.3462, Train Accuracy: 90.0926, Test Loss: 0.5619, Test Accuracy:80.0000
Epoch [31/100]: Train Loss: 0.2023, Train Accuracy: 95.4630, Test Loss: 0.4291, Test Accuracy:82.5000
Epoch [41/100]: Train Loss: 0.1177, Train Accuracy: 98.4259, Test Loss: 0.3876, Test Accuracy:84.1667
Epoch [51/100]: Train Loss: 0.0728, Train Accuracy: 99.3519, Test Loss: 0.3780, Test Accuracy:87.5000
Epoch [61/100]: Train Loss: 0.0491, Train Accuracy: 99.7222, Test Loss: 0.3957, Test Accuracy:82.5000
Epoch [71/100]: Train Loss: 0.0342, Train Accuracy: 99.8148, Test Loss: 0.4508, Test Accuracy:85.8333
Epoch [81/100]: Train Loss: 0.0242, Train Accuracy: 100.0000, Test Loss: 0.3556, Test Accuracy:87.5000
Epoch [91/100]: Train Loss: 0.0186, Train Accuracy: 100.0000, Test Loss: 0.3656, Test Accuracy:86.6667
```

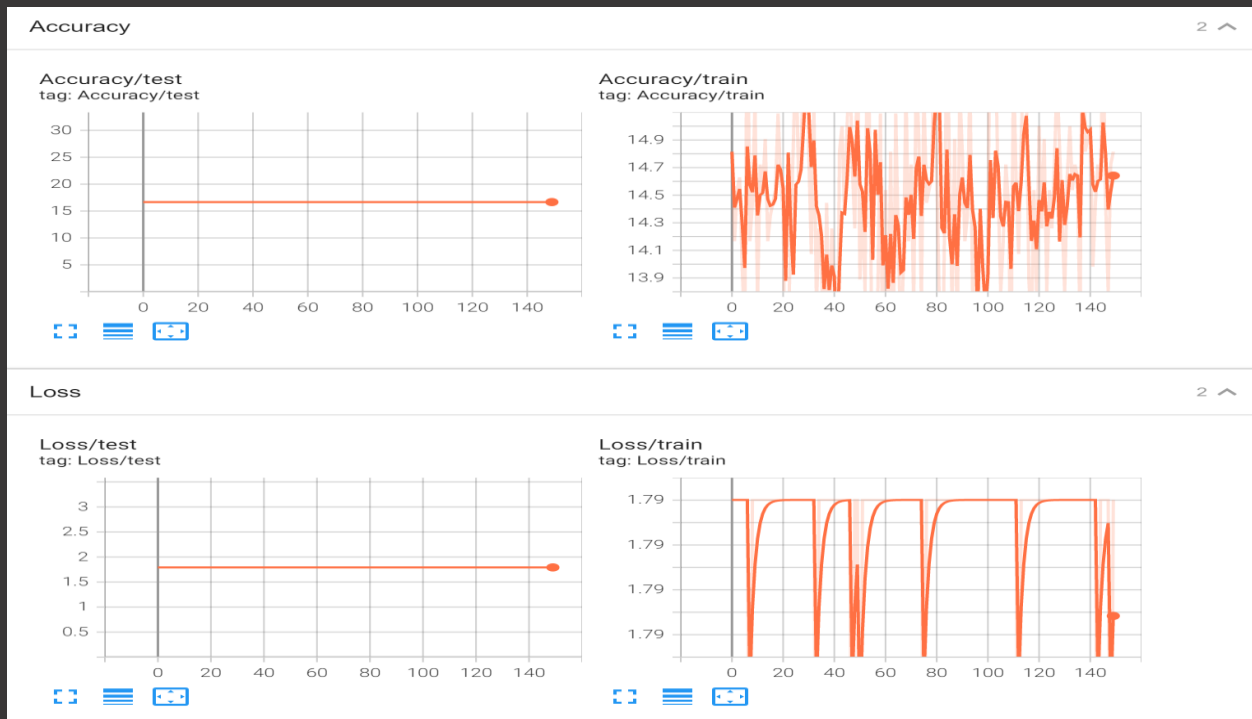
شکل 3- نتایج آموزش شبکه عصبی ابتدایی

## 4.1. مقدار دهی اولیه وزن‌ها:

مقداردهی اولیه صحیح وزن‌ها باعث می‌شود تا مدل سریعتر همگرا شود و مشکلاتی ازجمله exploding gradient یا vanishing gradient یا symmetry updating برای آن رخ ندهد. در ادامه حالت‌های مختلف مقداردهی وزن‌ها را بررسی میکنیم.

### مقداردهی اولیه با zero-constant:

در این حالت تمام پارامترهای مدل با صفر مقدار دهی می‌شود. در صورت استفاده از تابع فعالیت ReLU، خروجی تمام نورون‌ها صفر می‌شود و مدل مقدار صفر را پیشبینی میکند. (البته تابع softmax باعث می‌شود که مقدار خروجی تمام نورون‌ها برابر 1/6 بشود) و در حالتی که تمام نورون‌های عدد صفر را پیشبینی میکنند مطابق با فرمول‌های آپدیت وزن در backpropagation همواره مقدار گرادیان صفر خواهد شد و وزن‌ها آپدیت نمیشوند. در صورت استفاده از تابع فعالیت سیگموئید، تمام نورون‌های میانی عدد  $\frac{1}{2}$  را پیشبینی خواهند کرد. در این حالت که تمام نورون‌ها یک عدد برابر را پیشبینی میکنند، مقدار تمام گرادیان‌ها نیز با هم برابر خواهد شد. در نتیجه مطابق با فرمول آپدیت وزن‌ها، تمام وزن‌ها به یک مقدار آپدیت شوند. (symmetry problem) در نتیجه در صورت استفاده از مقدار دهی zero-constant، مدل آموزش نمیبیند.



شکل 4- نمودار دقت و خطای آموزش شبکه عصبی در حالت استفاده از مقداردهی اولیه zero-constant

```

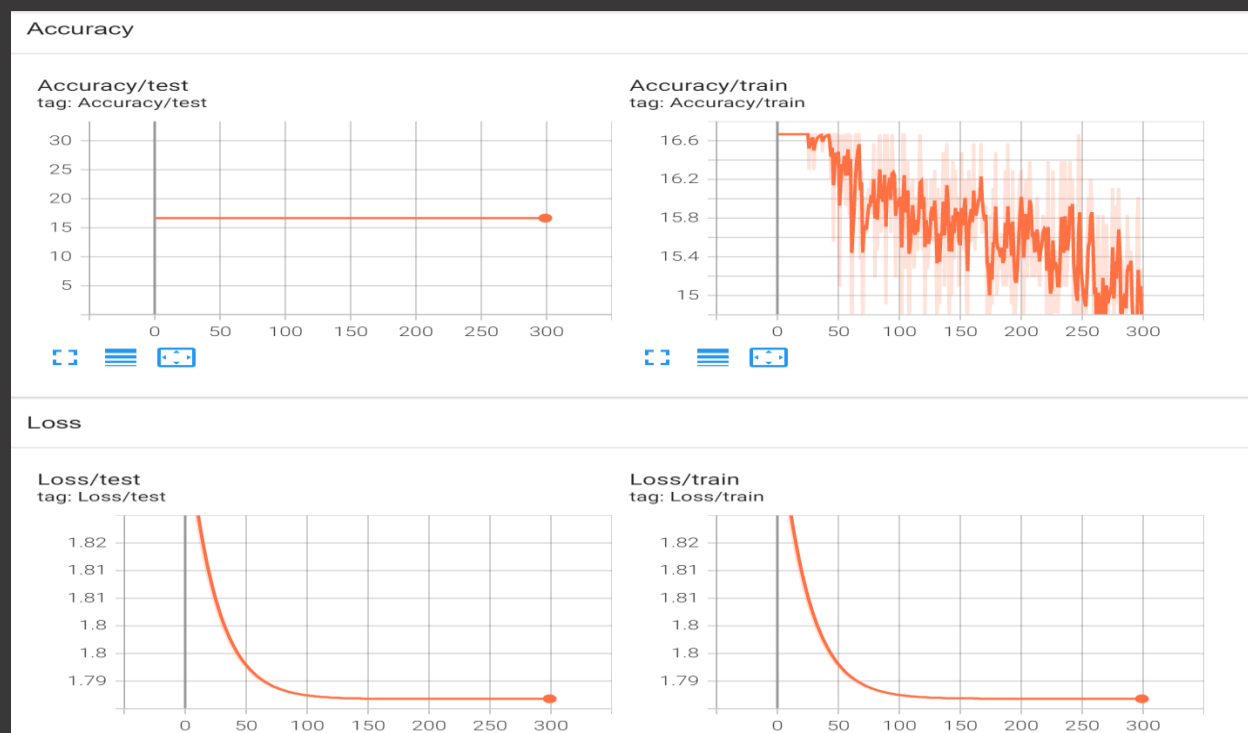
Epoch [1/150]: Train Loss: 1.7920, Train Accuracy: 14.8148, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [11/150]: Train Loss: 1.7920, Train Accuracy: 13.7037, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [21/150]: Train Loss: 1.7920, Train Accuracy: 14.3519, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [31/150]: Train Loss: 1.7920, Train Accuracy: 14.8148, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [41/150]: Train Loss: 1.7920, Train Accuracy: 13.7963, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [51/150]: Train Loss: 1.7919, Train Accuracy: 13.8889, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [61/150]: Train Loss: 1.7920, Train Accuracy: 14.5370, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [71/150]: Train Loss: 1.7920, Train Accuracy: 14.7222, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [81/150]: Train Loss: 1.7920, Train Accuracy: 15.5556, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [91/150]: Train Loss: 1.7920, Train Accuracy: 14.8148, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [101/150]: Train Loss: 1.7920, Train Accuracy: 14.2593, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [111/150]: Train Loss: 1.7920, Train Accuracy: 15.4630, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [121/150]: Train Loss: 1.7920, Train Accuracy: 15.0000, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [131/150]: Train Loss: 1.7920, Train Accuracy: 13.7963, Test Loss: 1.7918, Test Accuracy:16.6667
Epoch [141/150]: Train Loss: 1.7920, Train Accuracy: 15.0000, Test Loss: 1.7918, Test Accuracy:16.6667

```

شکل 5- نتایج آموزش شبکه عصبی در حالت استفاده از مقداردهی اولیه *zero-constant*

## مقداردهی اولیه با *uniform*:

در این حالت تمام وزن‌ها به طور رندوم با تابع توزیع یونیفرم مقداردهی میشوند. تمام اعدادی که در بازه‌ی مشخص شده تابع *uniform* قرار دارد، شانس برابر انتخاب شدن دارند. قانون کلی مقداردهی اولیه وزن‌ها این هست که وزن‌ها را در بازه  $[-y, +y]$  مقداردهی کنیم، بدون اینکه مقدار  $y$  خیلی بزرگ یا خیلی کوچک باشد. معمولاً مقدار  $y$  برابر  $1/\sqrt{\text{input neurons}}$  قرار داده میشود. در این پروژه وزن‌ها در بازه  $[-1, 0]$  مقداردهی شدند که با توجه به تعداد نورون‌های ورودی، پارامترها مقدار کوچکی گرفته‌اند. در نتیجه مشکل *vanishing gradient* به وجود آمده است و مقدار گرادیان‌ها بسیار کم شده است و مدل با سرعت خیلی کم همگرا میشود.



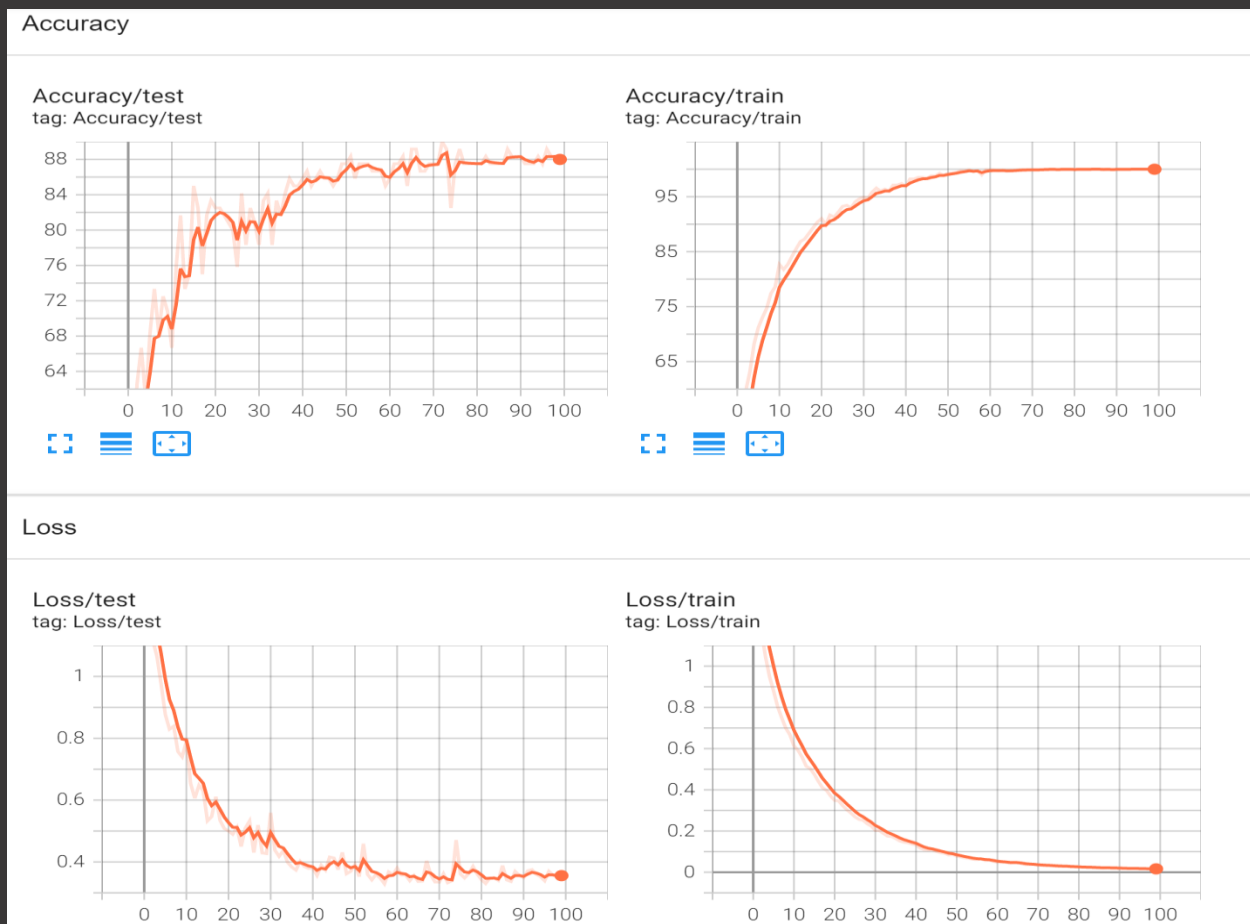
شکل 6- نمودار دقت و خطای آموزش شبکه عصبی در حالت استفاده از مقداردهی اولیه *uniform*



## 4.2. استفاده از تابع‌های فعالیت متفاوت:

در این قسمت عملکرد شبکه عصبی در صورت عدم استفاده از تابع فعالیت، استفاده از تابع ReLU و استفاده از تابع فعالیت سیگموید بررسی شده است. با توجه به نتایج بدست آمده در حالت‌های عدم استفاده از تابع فعالیت و استفاده از تابع ReLU به دقت خوب می‌رسیم. اما در صورت استفاده از تابع فعالیت سیگموید، مدل آموزش نمی‌بیند.

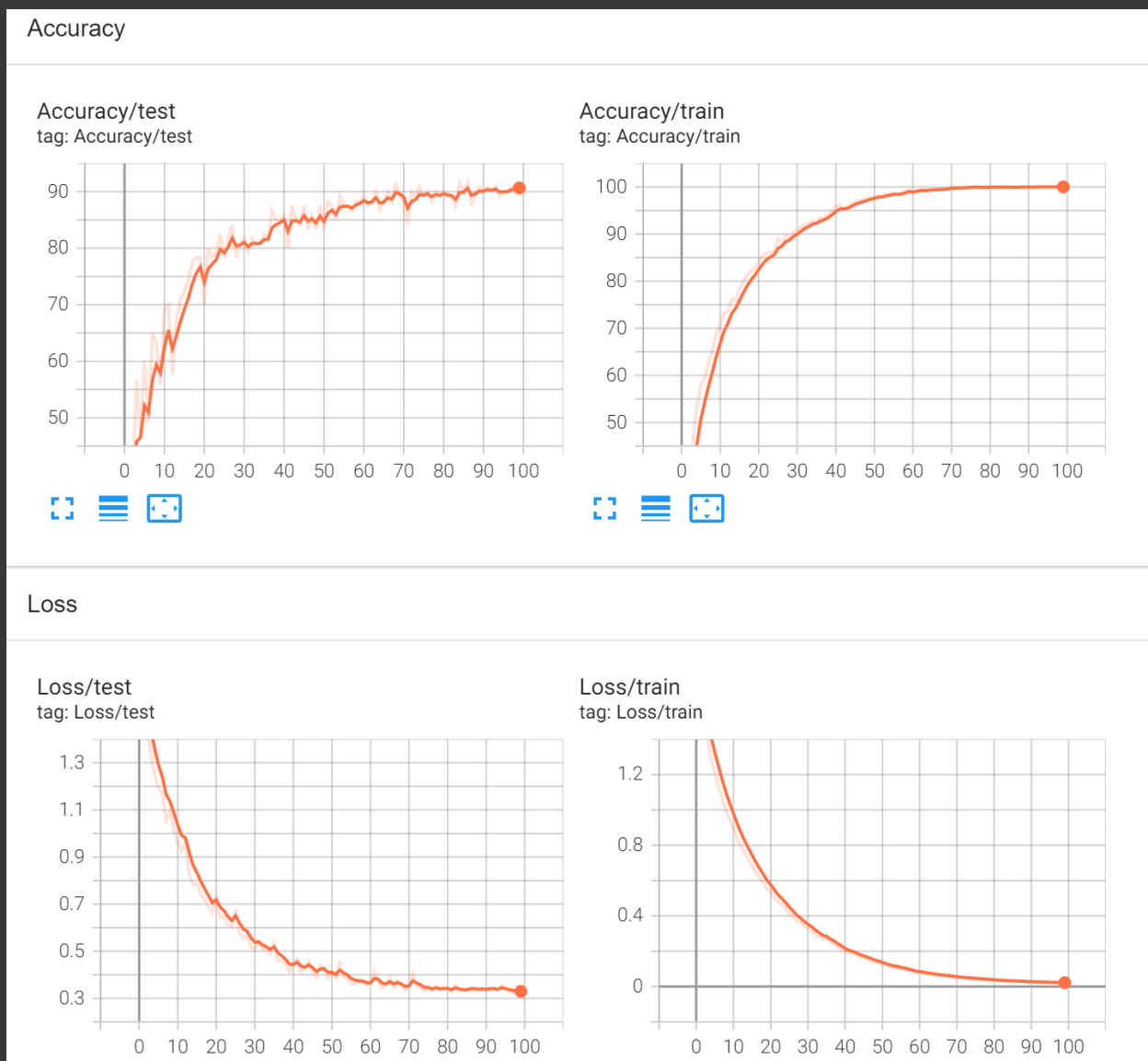
به نظر می‌رسد که تابع فعالیت سیگموید مشکل *vanishing gradient* را ایجاد می‌کند. تابع سیگموید خروجی نوره‌ها را در بازه  $[0,1]$  قرار می‌دهد. به عبارتی تغییرات زیاد در ورودی تابع سیگموید، تغییرات کم در خروجی خواهد داشت. در نتیجه مشتق تابع سیگموید عدد کوچکی است. گرادیان کم باعث می‌شود که وزن‌ها به مقدار کمی آپدیت شوند و آموزش خیلی کند اتفاق بیفتد و مدل همگرا نشود.



شکل 7- نمودار دقت و خطای آموزش شبکه عصبی در حالت عدم استفاده از تابع فعالیت

Epoch [1/100]: Train Loss: 1.6734, Train Accuracy: 33.0556, Test Loss: 1.4291, Test Accuracy:44.1667  
 Epoch [11/100]: Train Loss: 0.6135, Train Accuracy: 82.6852, Test Loss: 0.7902, Test Accuracy:66.6667  
 Epoch [21/100]: Train Loss: 0.3487, Train Accuracy: 91.0185, Test Loss: 0.5006, Test Accuracy:82.5000  
 Epoch [31/100]: Train Loss: 0.2029, Train Accuracy: 94.9074, Test Loss: 0.5589, Test Accuracy:78.3333  
 Epoch [41/100]: Train Loss: 0.1326, Train Accuracy: 96.9444, Test Loss: 0.3796, Test Accuracy:85.8333  
 Epoch [51/100]: Train Loss: 0.0773, Train Accuracy: 99.2593, Test Loss: 0.3919, Test Accuracy:87.5000  
 Epoch [61/100]: Train Loss: 0.0483, Train Accuracy: 99.9074, Test Loss: 0.3804, Test Accuracy:85.8333  
 Epoch [71/100]: Train Loss: 0.0337, Train Accuracy: 99.9074, Test Loss: 0.3347, Test Accuracy:87.5000  
 Epoch [81/100]: Train Loss: 0.0250, Train Accuracy: 99.9074, Test Loss: 0.3426, Test Accuracy:87.5000  
 Epoch [91/100]: Train Loss: 0.0192, Train Accuracy: 100.0000, Test Loss: 0.3490, Test Accuracy:88.3333

شکل 8- نتایج آموزش شبکه عصبی در حالت عدم استفاده از تابع فعالیت

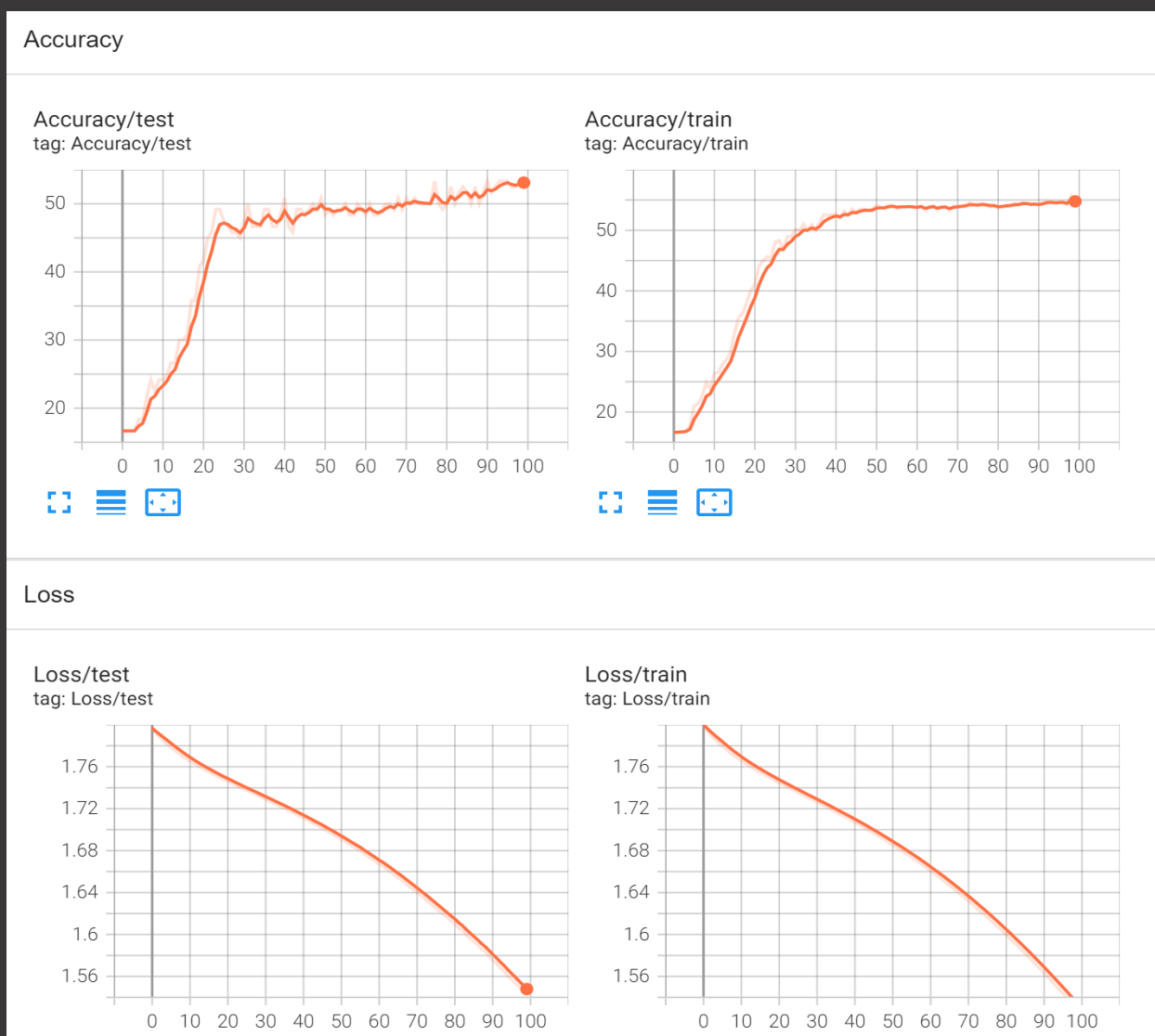


شکل 9- نمودار دقت و خطای آموزش شبکه عصبی در حالت استفاده از تابع فعالیت ReLU



Epoch [1/100]: Train Loss: 1.7723, Train Accuracy: 23.5185, Test Loss: 1.7316, Test Accuracy:21.6667  
 Epoch [11/100]: Train Loss: 0.8937, Train Accuracy: 70.8333, Test Loss: 0.9574, Test Accuracy:69.1667  
 Epoch [21/100]: Train Loss: 0.5371, Train Accuracy: 84.1667, Test Loss: 0.7363, Test Accuracy:70.0000  
 Epoch [31/100]: Train Loss: 0.3282, Train Accuracy: 90.8333, Test Loss: 0.5104, Test Accuracy:81.6667  
 Epoch [41/100]: Train Loss: 0.1945, Train Accuracy: 95.8333, Test Loss: 0.4425, Test Accuracy:85.8333  
 Epoch [51/100]: Train Loss: 0.1260, Train Accuracy: 97.8704, Test Loss: 0.4118, Test Accuracy:83.3333  
 Epoch [61/100]: Train Loss: 0.0825, Train Accuracy: 98.7037, Test Loss: 0.3640, Test Accuracy:89.1667  
 Epoch [71/100]: Train Loss: 0.0523, Train Accuracy: 99.9074, Test Loss: 0.3573, Test Accuracy:88.3333  
 Epoch [81/100]: Train Loss: 0.0367, Train Accuracy: 100.0000, Test Loss: 0.3433, Test Accuracy:90.0000  
 Epoch [91/100]: Train Loss: 0.0266, Train Accuracy: 99.9074, Test Loss: 0.3349, Test Accuracy:90.0000

شکل 10- نتایج آموزش شبکه عصبی در حالت استفاده از تابع فعالیت ReLU



شکل 11- نمودار دقت و خطای آموزش شبکه عصبی در حالت استفاده از تابع فعالیت sigmoid

Epoch [1/100]:	Train Loss: 1.8003, Train Accuracy: 16.6667, Test Loss: 1.7968, Test Accuracy:16.6667
Epoch [11/100]:	Train Loss: 1.7658, Train Accuracy: 26.3889, Test Loss: 1.7654, Test Accuracy:24.1667
Epoch [21/100]:	Train Loss: 1.7446, Train Accuracy: 41.0185, Test Loss: 1.7462, Test Accuracy:41.6667
Epoch [31/100]:	Train Loss: 1.7264, Train Accuracy: 50.0926, Test Loss: 1.7290, Test Accuracy:47.5000
Epoch [41/100]:	Train Loss: 1.7071, Train Accuracy: 52.6852, Test Loss: 1.7111, Test Accuracy:50.8333
Epoch [51/100]:	Train Loss: 1.6855, Train Accuracy: 54.1667, Test Loss: 1.6908, Test Accuracy:48.3333
Epoch [61/100]:	Train Loss: 1.6607, Train Accuracy: 53.7037, Test Loss: 1.6673, Test Accuracy:48.3333
Epoch [71/100]:	Train Loss: 1.6320, Train Accuracy: 53.9815, Test Loss: 1.6403, Test Accuracy:50.8333
Epoch [81/100]:	Train Loss: 1.5996, Train Accuracy: 53.6111, Test Loss: 1.6105, Test Accuracy:50.0000
Epoch [91/100]:	Train Loss: 1.5634, Train Accuracy: 54.1667, Test Loss: 1.5754, Test Accuracy:53.3333

شکل 12- نتایج آموزش شبکه عصبی در حالت استفاده از تابع فعالیت *sigmoid*

### 4.3. استفاده از معماری‌های متفاوت:

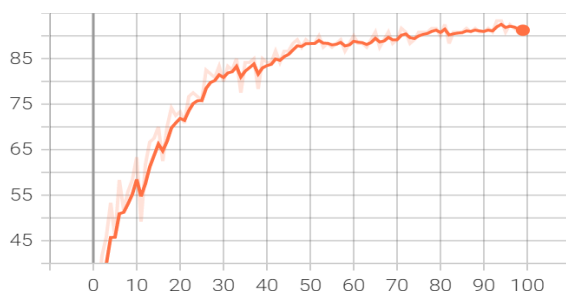
تعداد لایه‌ها و پارامترهای خیلی کم در شبکه عصبی باعث میشود که مدل حتی داده‌های آموزشی را به خوبی یاد نگیرد و به دقت بدی روی داده‌های تست برسد. از طرف دیگر معمولا با افزایش بیش از اندازه تعداد لایه‌ها و به طور کل پارامترهای شبکه، مدل روی داده‌های آموزشی *overfit* میکند و دقت آن روی داده‌های تست کاهش میابد. همچنین مدل با تعداد پارامترهای بیشتر نیاز به محاسبات بیشتری دارد. در نتیجه فرآیند آموزش مدل های عمیق تر زمانبر تر است.

در ادامه دو شبکه مختلف بر روی دیتاست آموزش میبندد. شبکه اول شامل سه لایه مخفی با 500، 100 و 50 نورون است. شبکه دوم دارای دو لایه مخفی با 100 و 50 نورون است. با بررسی نتایج مشاهده میشود که بعد از 100 اپیاک هر دو شبکه به دقت‌های خوبی میرسند و شبکه با تعداد لایه‌های بیشتر منجر به *overfit* نشده است. تحلیل این اتفاق میتواند به *data distribution* در داده‌های آموزشی و تست مرتبط باشد. به نظر میرسد که در دیتاست این پروژه داده‌های تست بسیار مشابه داده‌ها آموزش هستند چون بعد از اینکه مدل کامل روی داده‌های آموزشی فیت میشود (دقت 100%) دقت روی داده‌های تست کاهش نمیابد.

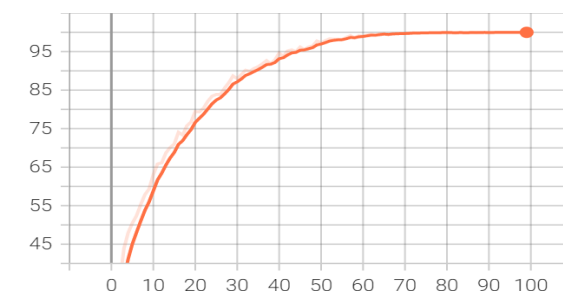
با توجه به مشاهدات انجام شده بهتر است با مدلی که دارای تعداد پارامتر های کمتر است ادامه بدهیم. چون با توجه به دیتاست موجود تعداد پارامتر های کم برای رسیدن به دقت مناسب کفایت میکند و پارامتر های زیاد صرفا هزینه زمانی و محاسباتی خواهد داشت.

## Accuracy

Accuracy/test  
tag: Accuracy/test

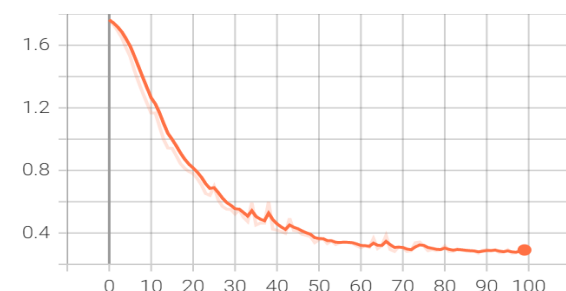


Accuracy/train  
tag: Accuracy/train

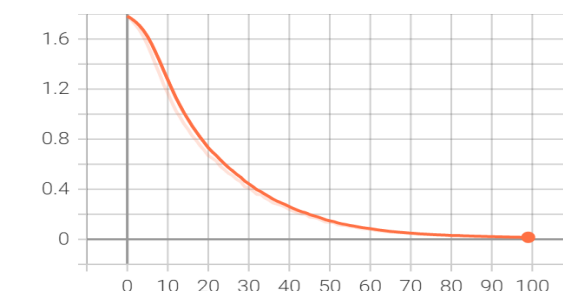


## Loss

Loss/test  
tag: Loss/test



Loss/train  
tag: Loss/train



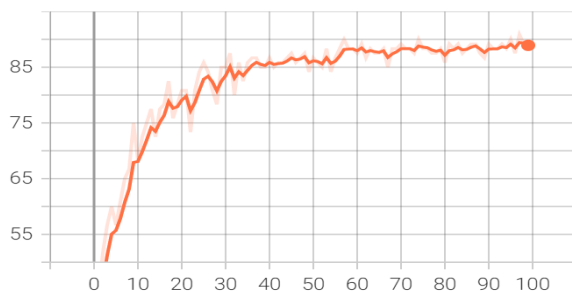
شکل 13- نمودار دقت و خطای آموزش شبکه عصبی با معماری سه لایه مخفی

```
Epoch [1/100]: Train Loss: 1.7835, Train Accuracy: 18.7037, Test Loss: 1.7629, Test Accuracy:20.8333
Epoch [11/100]: Train Loss: 1.1667, Train Accuracy: 63.0556, Test Loss: 1.1673, Test Accuracy:63.3333
Epoch [21/100]: Train Loss: 0.6681, Train Accuracy: 79.4444, Test Loss: 0.7799, Test Accuracy:73.3333
Epoch [31/100]: Train Loss: 0.4100, Train Accuracy: 87.9630, Test Loss: 0.5219, Test Accuracy:80.0000
Epoch [41/100]: Train Loss: 0.2376, Train Accuracy: 94.6296, Test Loss: 0.4183, Test Accuracy:84.1667
Epoch [51/100]: Train Loss: 0.1356, Train Accuracy: 97.3148, Test Loss: 0.3564, Test Accuracy:88.3333
Epoch [61/100]: Train Loss: 0.0796, Train Accuracy: 99.0741, Test Loss: 0.3064, Test Accuracy:90.0000
Epoch [71/100]: Train Loss: 0.0456, Train Accuracy: 99.7222, Test Loss: 0.3041, Test Accuracy:89.1667
Epoch [81/100]: Train Loss: 0.0278, Train Accuracy: 100.0000, Test Loss: 0.3177, Test Accuracy:90.0000
Epoch [91/100]: Train Loss: 0.0208, Train Accuracy: 99.9074, Test Loss: 0.2942, Test Accuracy:90.8333
```

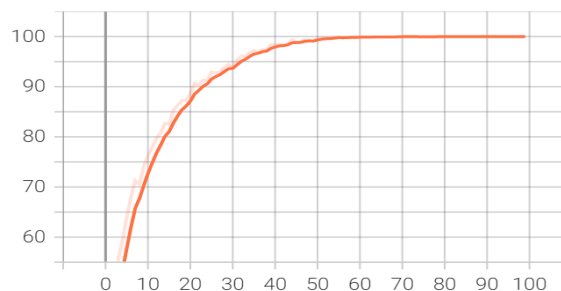
شکل 14- نتایج آموزش شبکه عصبی با معماری سه لایه مخفی

## Accuracy

Accuracy/test  
tag: Accuracy/test

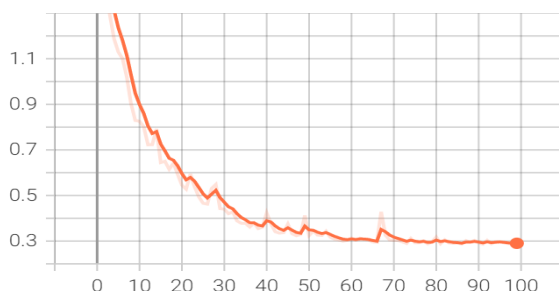


Accuracy/train  
tag: Accuracy/train

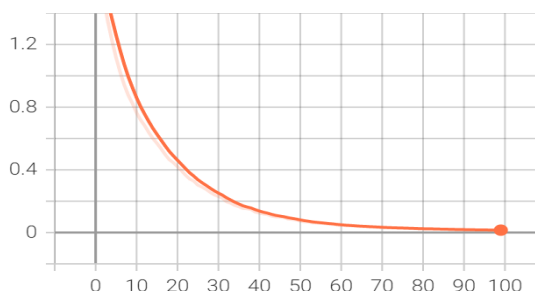


## Loss

Loss/test  
tag: Loss/test



Loss/train  
tag: Loss/train



شکل 15- نمودار دقت و خطای آموزش شبکه عصبی با معماری دو لایه مخفی

```
Epoch [1/100]: Train Loss: 1.7455, Train Accuracy: 27.5926, Test Loss: 1.6762, Test Accuracy:36.6667
Epoch [11/100]: Train Loss: 0.7646, Train Accuracy: 76.2963, Test Loss: 0.8257, Test Accuracy:68.3333
Epoch [21/100]: Train Loss: 0.4200, Train Accuracy: 88.6111, Test Loss: 0.5452, Test Accuracy:80.8333
Epoch [31/100]: Train Loss: 0.2302, Train Accuracy: 93.7963, Test Loss: 0.4400, Test Accuracy:85.0000
Epoch [41/100]: Train Loss: 0.1243, Train Accuracy: 98.3333, Test Loss: 0.4235, Test Accuracy:86.6667
Epoch [51/100]: Train Loss: 0.0726, Train Accuracy: 99.6296, Test Loss: 0.3243, Test Accuracy:86.6667
Epoch [61/100]: Train Loss: 0.0462, Train Accuracy: 99.9074, Test Loss: 0.3134, Test Accuracy:87.5000
Epoch [71/100]: Train Loss: 0.0315, Train Accuracy: 100.0000, Test Loss: 0.3034, Test Accuracy:89.1667
Epoch [81/100]: Train Loss: 0.0232, Train Accuracy: 100.0000, Test Loss: 0.3179, Test Accuracy:85.8333
Epoch [91/100]: Train Loss: 0.0181, Train Accuracy: 100.0000, Test Loss: 0.2920, Test Accuracy:89.1667
```

شکل 16- نتایج آموزش شبکه عصبی با معماری دو لایه مخفی

## 4.4.1. استفاده از dropout:

یکی از راهکارهای حل مشکل overfitting استفاده از لایه‌های dropout است. با عمیق تر کردن مدل و افزایش تعدد پارامترها، مدل داده‌های آموزشی را بهتر یاد می‌گیرد اما باید حواسمون به پدیده overfitting باشد. لایه‌های dropout نوروں‌های شبکه عصبی را با احتمالی از پیش تعیین شده دور میریزد. اینگونه مدل ساده‌تر میشود و میتوانیم شبکه‌های عمیق بدون مشکل overfit داشته باشیم. اما در مسئله‌ی ما در هیچ حالتی overfit رخ نداده است. پس از اعمال dropout دقت مدل خیلی کاهش یافت چون مدل ساده‌تر شده است. به عبارتی مدل روی دیتا underfit کرده است. برای اینکه مدل به دقت قبلی برسد نیاز داریم که مدل را با اپیاک بیشتری آموزش بدهیم. در نتیجه استفاده از لایه‌های dropout لزوما باعث بهتر شده دقت مدل نمیشود و باید به مسئله توجه کرد.



شکل 17- نمودار دقت و خطای آموزش شبکه عصبی پس از اضافه کردن لایه‌های dropout

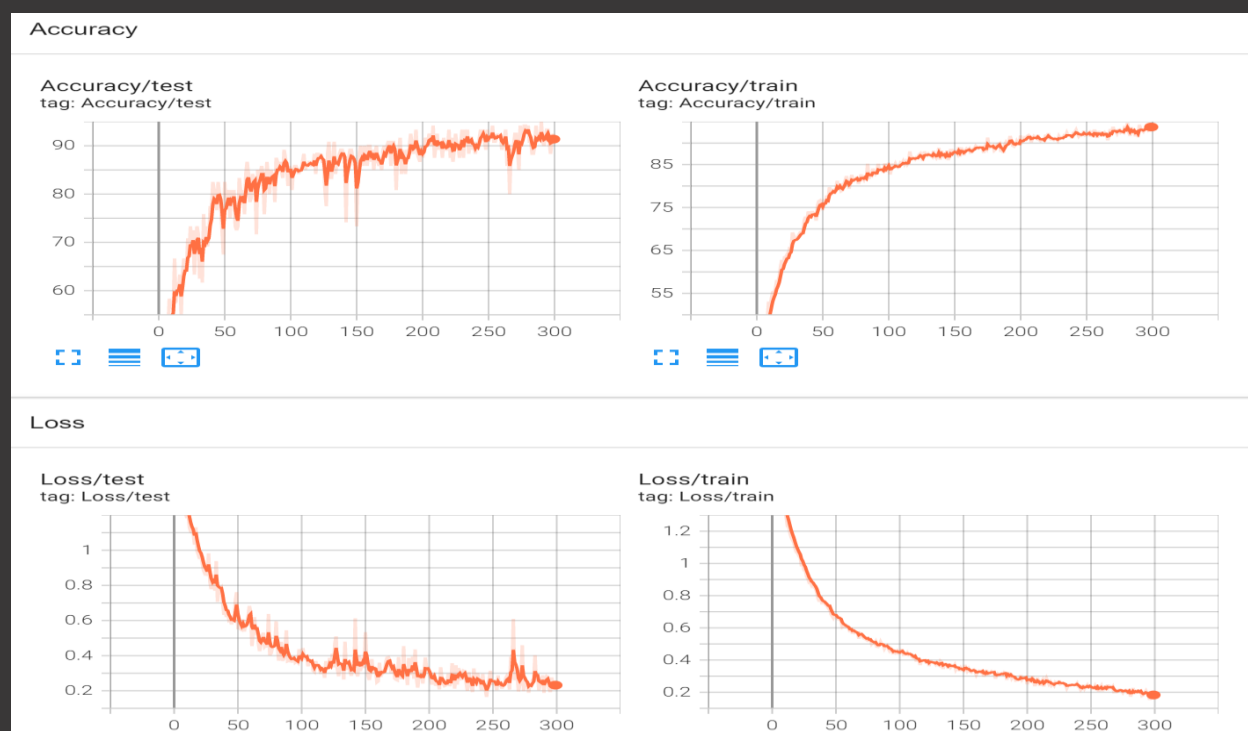
Epoch [1/100]: Train Loss: 1.9049, Train Accuracy: 20.2778, Test Loss: 1.8687, Test Accuracy:18.3333  
 Epoch [11/100]: Train Loss: 1.5929, Train Accuracy: 32.8704, Test Loss: 1.5943, Test Accuracy:30.0000  
 Epoch [21/100]: Train Loss: 1.5207, Train Accuracy: 38.0556, Test Loss: 1.4814, Test Accuracy:36.6667  
 Epoch [31/100]: Train Loss: 1.4101, Train Accuracy: 41.4815, Test Loss: 1.4596, Test Accuracy:35.8333  
 Epoch [41/100]: Train Loss: 1.3888, Train Accuracy: 43.5185, Test Loss: 1.4640, Test Accuracy:40.0000  
 Epoch [51/100]: Train Loss: 1.3108, Train Accuracy: 46.0185, Test Loss: 1.3294, Test Accuracy:42.5000  
 Epoch [61/100]: Train Loss: 1.2962, Train Accuracy: 46.4815, Test Loss: 1.3510, Test Accuracy:41.6667  
 Epoch [71/100]: Train Loss: 1.2772, Train Accuracy: 47.5000, Test Loss: 1.3747, Test Accuracy:39.1667  
 Epoch [81/100]: Train Loss: 1.2494, Train Accuracy: 46.5741, Test Loss: 1.2988, Test Accuracy:45.0000  
 Epoch [91/100]: Train Loss: 1.2323, Train Accuracy: 48.7963, Test Loss: 1.3081, Test Accuracy:43.3333

شکل 18- نتایج آموزش شبکه عصبی پس از اضافه کردن لایه‌های dropout

## 4.4.2. استفاده از data augmentation:

دیتاست‌های بزرگ با تنوع زیاد سبب افزایش دقت مدل بر روی داده‌های تست می‌شود. اما گاهی تهیه دیتاست‌های بزرگ هزینه‌بر و زمان‌بر است. تکنیک‌های data augmentation باعث به ماکم می‌کنند که بدون صرف هزینه حجم داده‌های خود را افزایش بدهیم. اینکار با انجام تبدیلاتی از جمله rotation، flipping، translation و .. انجام می‌شود. همچنین افزایش حجم دیتا می‌تواند مشکلات بیش برآزش را نیز حل کند.

در این پروژه بعد از اعمال data augmentation موفق شدیم که دقت تست را تا 95% افزایش بدهیم.



شکل 18- نمودار دقت و خطای آموزش شبکه عصبی پس از انجام data augmentation



```

Epoch [1/300]: Train Loss: 1.7853, Train Accuracy: 18.9815, Test Loss: 1.7962, Test Accuracy:16.6667
Epoch [11/300]: Train Loss: 1.3232, Train Accuracy: 51.2963, Test Loss: 1.1831, Test Accuracy:56.6667
Epoch [21/300]: Train Loss: 1.0627, Train Accuracy: 60.9259, Test Loss: 0.9510, Test Accuracy:65.8333
Epoch [31/300]: Train Loss: 0.8844, Train Accuracy: 67.5926, Test Loss: 0.7955, Test Accuracy:71.6667
Epoch [41/300]: Train Loss: 0.7531, Train Accuracy: 73.6111, Test Loss: 0.6541, Test Accuracy:76.6667
Epoch [51/300]: Train Loss: 0.6790, Train Accuracy: 74.3519, Test Loss: 0.5569, Test Accuracy:82.5000
Epoch [61/300]: Train Loss: 0.5965, Train Accuracy: 79.5370, Test Loss: 0.6495, Test Accuracy:73.3333
Epoch [71/300]: Train Loss: 0.5642, Train Accuracy: 79.3519, Test Loss: 0.5303, Test Accuracy:79.1667
Epoch [81/300]: Train Loss: 0.5031, Train Accuracy: 81.9444, Test Loss: 0.5958, Test Accuracy:75.8333
Epoch [91/300]: Train Loss: 0.4913, Train Accuracy: 82.6852, Test Loss: 0.3711, Test Accuracy:86.6667
Epoch [101/300]: Train Loss: 0.4436, Train Accuracy: 85.3704, Test Loss: 0.3687, Test Accuracy:85.0000
Epoch [111/300]: Train Loss: 0.4170, Train Accuracy: 85.2778, Test Loss: 0.3068, Test Accuracy:88.3333
Epoch [121/300]: Train Loss: 0.3996, Train Accuracy: 86.2037, Test Loss: 0.3275, Test Accuracy:88.3333
Epoch [131/300]: Train Loss: 0.3674, Train Accuracy: 87.5000, Test Loss: 0.3647, Test Accuracy:83.3333
Epoch [141/300]: Train Loss: 0.3554, Train Accuracy: 88.2407, Test Loss: 0.3288, Test Accuracy:88.3333
Epoch [151/300]: Train Loss: 0.3322, Train Accuracy: 88.2407, Test Loss: 0.5331, Test Accuracy:73.3333
Epoch [161/300]: Train Loss: 0.3084, Train Accuracy: 88.4259, Test Loss: 0.2534, Test Accuracy:90.0000
Epoch [171/300]: Train Loss: 0.3237, Train Accuracy: 88.9815, Test Loss: 0.3049, Test Accuracy:89.1667
Epoch [181/300]: Train Loss: 0.3184, Train Accuracy: 88.2407, Test Loss: 0.4541, Test Accuracy:80.8333
Epoch [191/300]: Train Loss: 0.2920, Train Accuracy: 89.9074, Test Loss: 0.2727, Test Accuracy:90.0000
Epoch [201/300]: Train Loss: 0.2488, Train Accuracy: 91.7593, Test Loss: 0.2947, Test Accuracy:87.5000
Epoch [211/300]: Train Loss: 0.2397, Train Accuracy: 91.5741, Test Loss: 0.2711, Test Accuracy:90.0000
Epoch [221/300]: Train Loss: 0.2502, Train Accuracy: 91.4815, Test Loss: 0.2122, Test Accuracy:93.3333
Epoch [231/300]: Train Loss: 0.2459, Train Accuracy: 91.0185, Test Loss: 0.1995, Test Accuracy:93.3333
Epoch [241/300]: Train Loss: 0.2342, Train Accuracy: 92.4074, Test Loss: 0.2507, Test Accuracy:88.3333
Epoch [251/300]: Train Loss: 0.2278, Train Accuracy: 91.8519, Test Loss: 0.2548, Test Accuracy:90.8333
Epoch [261/300]: Train Loss: 0.2262, Train Accuracy: 92.1296, Test Loss: 0.2710, Test Accuracy:89.1667
Epoch [271/300]: Train Loss: 0.2001, Train Accuracy: 93.7963, Test Loss: 0.1958, Test Accuracy:92.5000
Epoch [281/300]: Train Loss: 0.2201, Train Accuracy: 92.7778, Test Loss: 0.2223, Test Accuracy:94.1667
Epoch [291/300]: Train Loss: 0.1767, Train Accuracy: 94.2593, Test Loss: 0.1881, Test Accuracy:95.0000

```

شکل 19- نتایج آموزش شبکه عصبی پس از انجام *data augmentation*

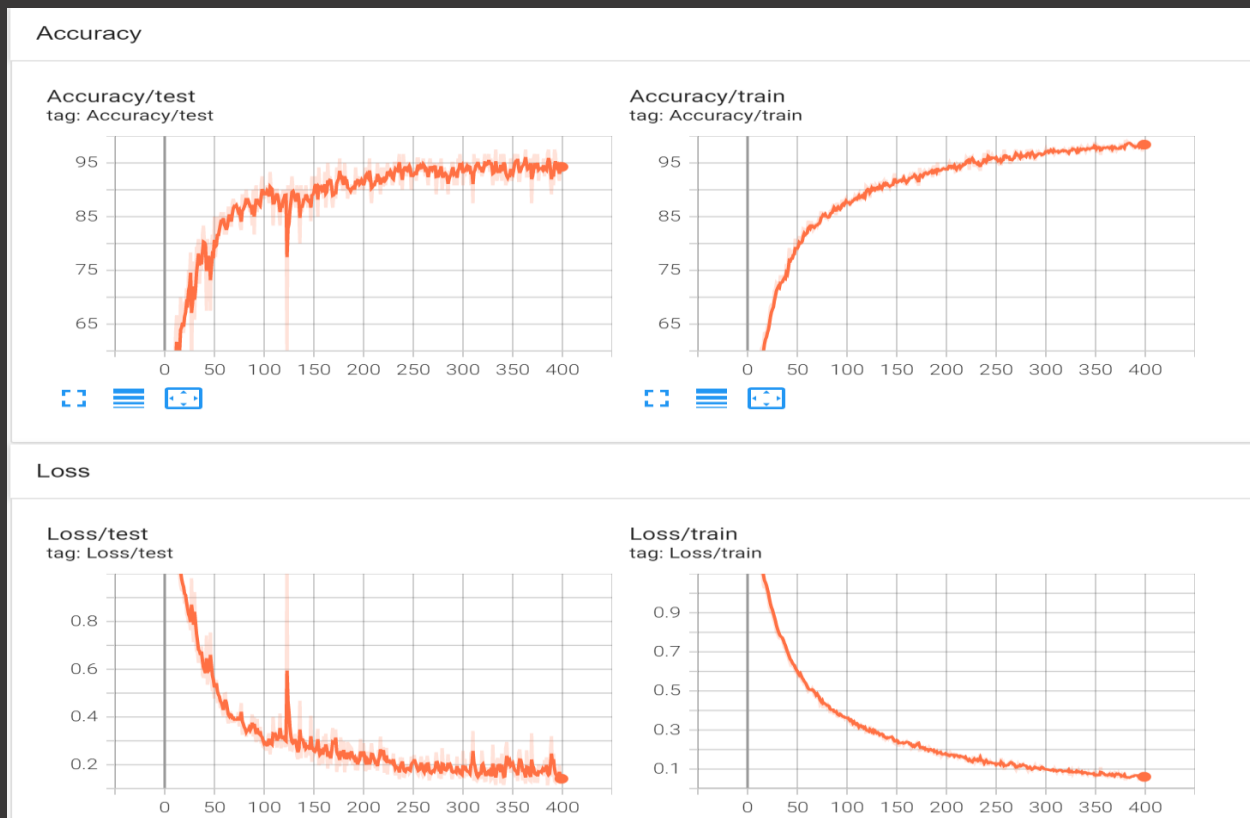
## بهترین شبکه عصبی نهایی: (optional)

پس از انجام آزمایشات متعدد، شبکه‌ای با تنظیمات زیر به عنوان بهترین شبکه انتخاب شد. این شبکه در نهایت به دقت 95% بر روی داده‌های تست میرسد.

```

1  - activation: relu
2      augmentation: true
3      batch_size: 8
4      hidden_layers: [50, 100]
5      img_size: 64
6      initialization: null
7      learning_rate: 0.001
8      n_classes: 6
9      n_epochs: 400
10     print_every: 10

```



شکل 20- نمودار دقت و خطای بهترین شبکه عصبی تنظیم شده

```
Epoch [41/400]: Train Loss: 0.6809, Train Accuracy: 73.8889, Test Loss: 0.5816, Test Accuracy:80.0000
Epoch [51/400]: Train Loss: 0.5923, Train Accuracy: 78.9815, Test Loss: 0.4770, Test Accuracy:85.0000
Epoch [61/400]: Train Loss: 0.5066, Train Accuracy: 82.2222, Test Loss: 0.4964, Test Accuracy:81.6667
Epoch [71/400]: Train Loss: 0.4943, Train Accuracy: 83.3333, Test Loss: 0.3760, Test Accuracy:87.5000
Epoch [81/400]: Train Loss: 0.4156, Train Accuracy: 84.9074, Test Loss: 0.3369, Test Accuracy:87.5000
Epoch [91/400]: Train Loss: 0.3697, Train Accuracy: 87.5926, Test Loss: 0.3984, Test Accuracy:81.6667
Epoch [101/400]: Train Loss: 0.3592, Train Accuracy: 87.1296, Test Loss: 0.2711, Test Accuracy:88.3333
Epoch [111/400]: Train Loss: 0.3399, Train Accuracy: 87.9630, Test Loss: 0.3274, Test Accuracy:87.5000
Epoch [121/400]: Train Loss: 0.2884, Train Accuracy: 90.5556, Test Loss: 0.3460, Test Accuracy:89.1667
Epoch [131/400]: Train Loss: 0.2771, Train Accuracy: 90.5556, Test Loss: 0.2975, Test Accuracy:88.3333
Epoch [141/400]: Train Loss: 0.2537, Train Accuracy: 91.1111, Test Loss: 0.2569, Test Accuracy:90.0000
Epoch [151/400]: Train Loss: 0.2557, Train Accuracy: 91.1111, Test Loss: 0.2424, Test Accuracy:90.0000
Epoch [161/400]: Train Loss: 0.2438, Train Accuracy: 90.7407, Test Loss: 0.3189, Test Accuracy:88.3333
Epoch [171/400]: Train Loss: 0.2351, Train Accuracy: 91.2037, Test Loss: 0.2902, Test Accuracy:89.1667
Epoch [181/400]: Train Loss: 0.2067, Train Accuracy: 92.6852, Test Loss: 0.2990, Test Accuracy:90.0000
Epoch [191/400]: Train Loss: 0.1949, Train Accuracy: 92.7778, Test Loss: 0.2191, Test Accuracy:90.8333
Epoch [201/400]: Train Loss: 0.1647, Train Accuracy: 94.0741, Test Loss: 0.2904, Test Accuracy:90.0000
Epoch [211/400]: Train Loss: 0.1721, Train Accuracy: 93.4259, Test Loss: 0.2115, Test Accuracy:91.6667
Epoch [221/400]: Train Loss: 0.1366, Train Accuracy: 96.0185, Test Loss: 0.2735, Test Accuracy:90.8333
Epoch [231/400]: Train Loss: 0.1426, Train Accuracy: 94.8148, Test Loss: 0.2475, Test Accuracy:91.6667
Epoch [241/400]: Train Loss: 0.1529, Train Accuracy: 94.8148, Test Loss: 0.1510, Test Accuracy:96.6667
Epoch [251/400]: Train Loss: 0.1110, Train Accuracy: 96.2963, Test Loss: 0.2137, Test Accuracy:92.5000
Epoch [261/400]: Train Loss: 0.1144, Train Accuracy: 96.2963, Test Loss: 0.1522, Test Accuracy:95.8333
Epoch [271/400]: Train Loss: 0.0957, Train Accuracy: 97.5000, Test Loss: 0.1813, Test Accuracy:95.0000
Epoch [281/400]: Train Loss: 0.1157, Train Accuracy: 95.9259, Test Loss: 0.2012, Test Accuracy:91.6667
Epoch [291/400]: Train Loss: 0.1039, Train Accuracy: 96.4815, Test Loss: 0.1626, Test Accuracy:95.0000
Epoch [301/400]: Train Loss: 0.0972, Train Accuracy: 97.5926, Test Loss: 0.2580, Test Accuracy:90.8333
Epoch [311/400]: Train Loss: 0.0904, Train Accuracy: 97.0370, Test Loss: 0.3415, Test Accuracy:87.5000
Epoch [321/400]: Train Loss: 0.0883, Train Accuracy: 97.2222, Test Loss: 0.2107, Test Accuracy:93.3333
Epoch [331/400]: Train Loss: 0.0816, Train Accuracy: 97.4074, Test Loss: 0.2843, Test Accuracy:90.8333
Epoch [341/400]: Train Loss: 0.0816, Train Accuracy: 97.1296, Test Loss: 0.1305, Test Accuracy:95.0000
Epoch [351/400]: Train Loss: 0.0658, Train Accuracy: 97.9630, Test Loss: 0.2157, Test Accuracy:92.5000
Epoch [361/400]: Train Loss: 0.0635, Train Accuracy: 97.6852, Test Loss: 0.1349, Test Accuracy:95.0000
Epoch [371/400]: Train Loss: 0.0841, Train Accuracy: 97.6852, Test Loss: 0.1575, Test Accuracy:93.3333
Epoch [381/400]: Train Loss: 0.0429, Train Accuracy: 99.2593, Test Loss: 0.1969, Test Accuracy:93.3333
Epoch [391/400]: Train Loss: 0.0697, Train Accuracy: 97.5926, Test Loss: 0.1488, Test Accuracy:95.8333
write successful
```

شکل 19- نتایج آموزش بهترین شبکه عصبی تنظیم شده