

## شبکه عصبی تمرین دوم

سحر محمدی ۴۰۲۱۳۹۰۱۰۹

هدف این تمرین حل یک مسئله طبقه بندی برای دو کلاس پیرهن و شلوار از داده‌های Fashion MNIST به کمک شبکه MLP است. به همراه بررسی برخی موارد از جمله توابع فعال سازی مختلف، روش‌های وزن دهی اولیه و... .

### شرح پیاده سازی

طبق موارد خواسته شده، پیاده سازی انجام شد.

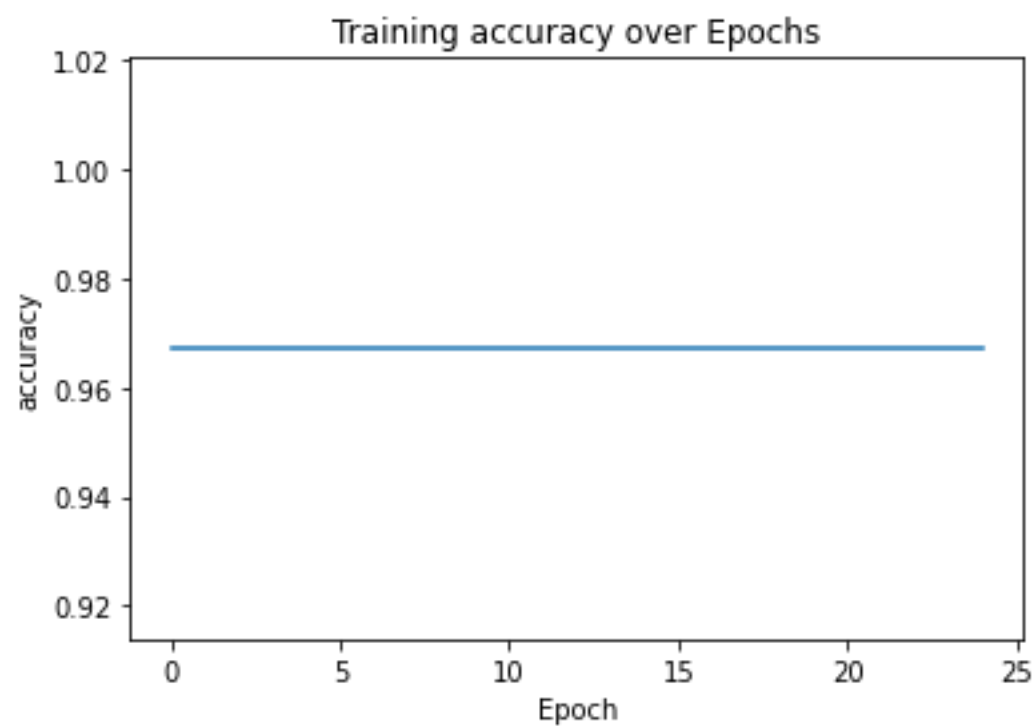
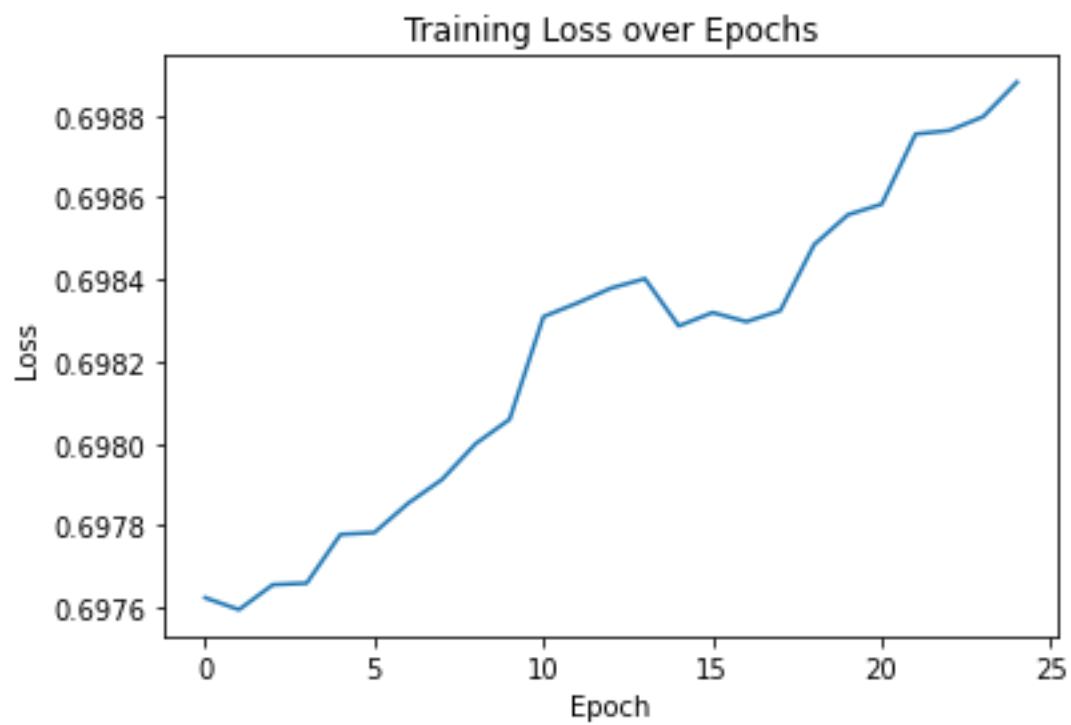
برای هر یک از توابع فعالیت از جمله همانی، سیگموئید، تانژانت هایپربولیک، رلو و لیکی رلو یک تابع تعریف شد. سپس برای هر یک از روش‌های وزن دهی اولیه یک تابع تعریف شد. روش‌ها از جمله تمام صفر، رندوم با توزیع نرمال که مقادیر کوچکی باشند، رندوم با توزیع نرمال که مقادیر بزرگی باشند، روش xavier و He.

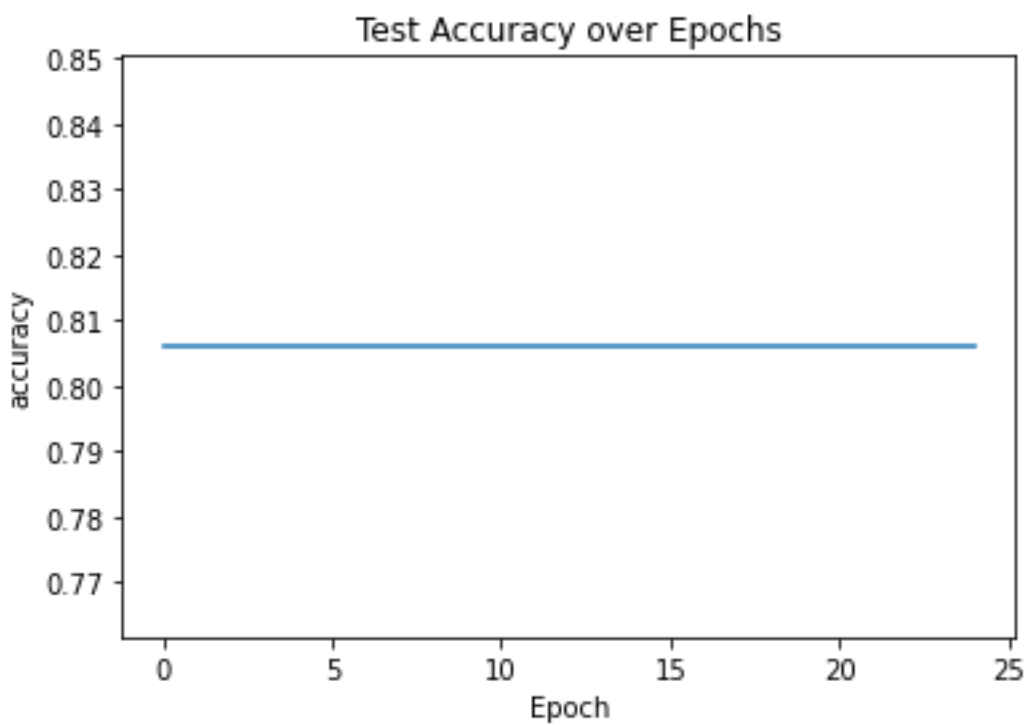
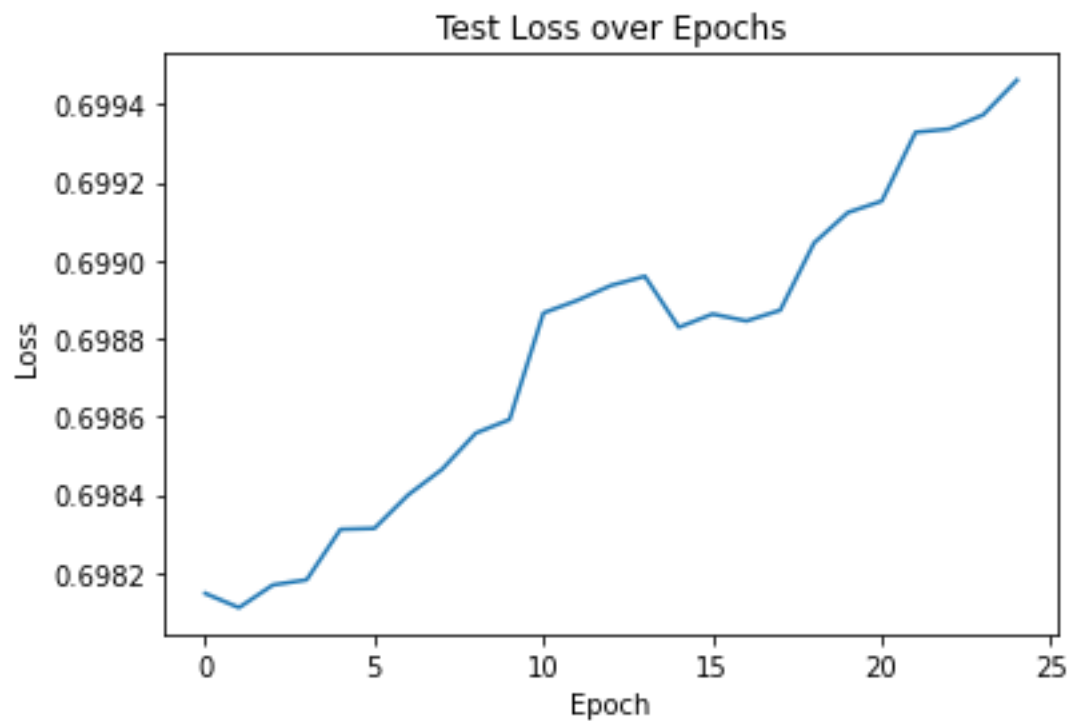
### آزمایشات

۱. بررسی تاثیر تابع فعالیت

- در آزمایش اول از یک شبکه با ۳ لایه مخفی که در همه لایه‌ها تابع فعالیت همانی بود استفاده شد.
- سپس یک شبکه ۴ لایه با توابع لیکی رلو، همانی، همانی و سیگموئید.

```
train loss and accuracy
0.6988813340630882 0.9671666666666666
test loss and accuracy
0.6994614016259664 0.806
```

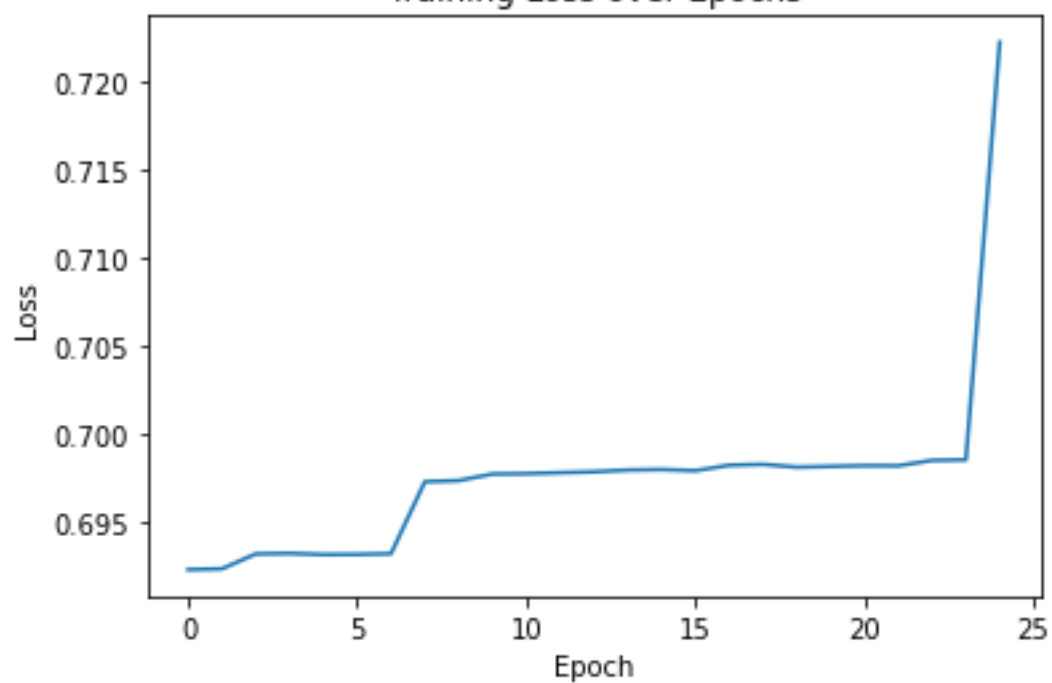




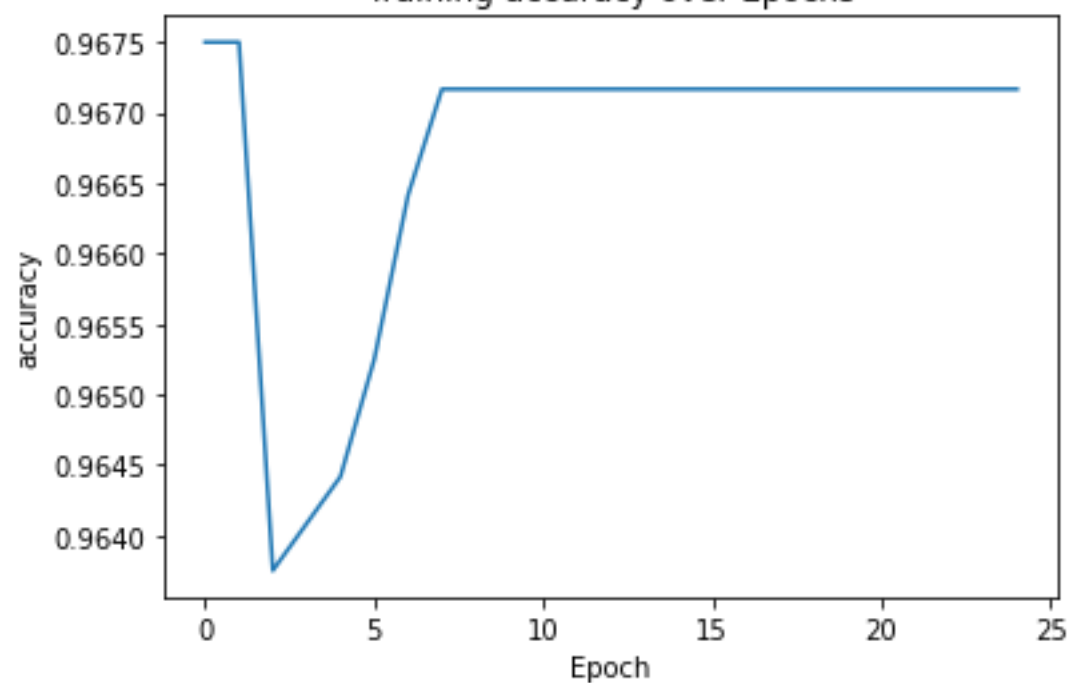
• ۴ لایه با رلو، رلو، همانی و سیگموید.

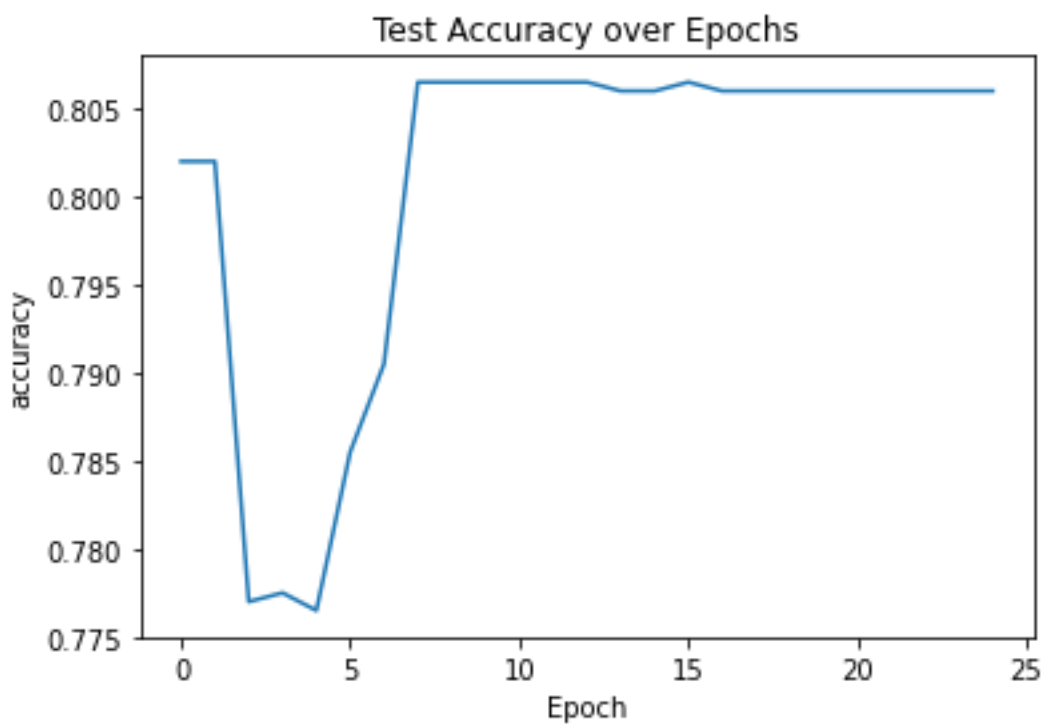
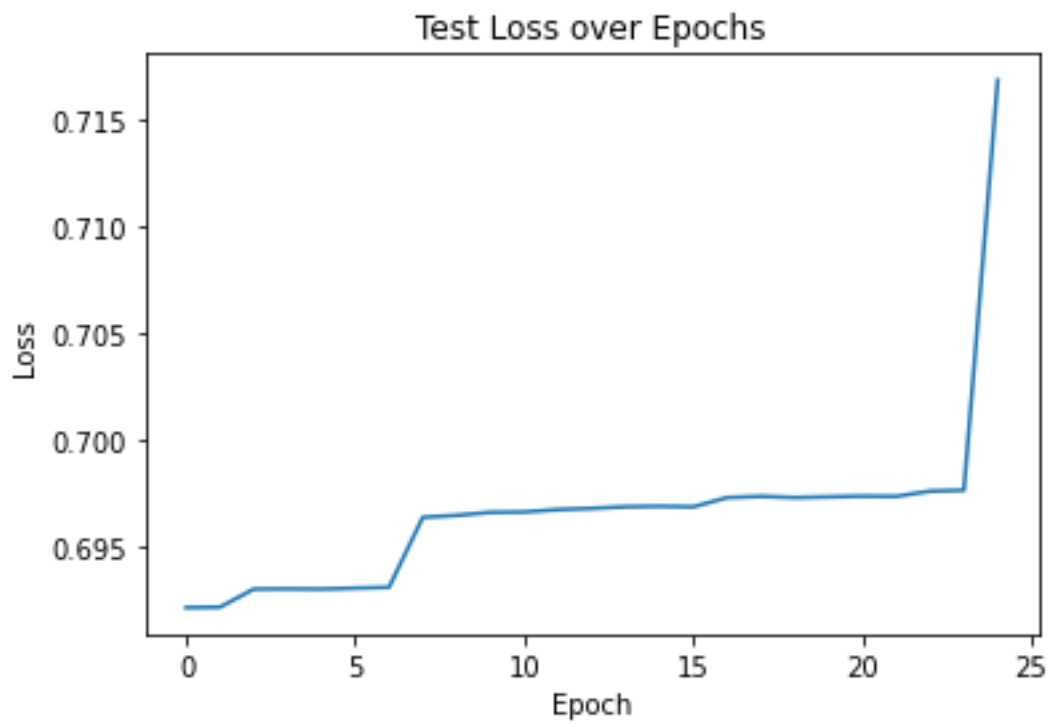
```
train loss and accuracy
0.7222022557370538 0.9671666666666666
test loss and accuracy
0.7168333453816205 0.806
```

Training Loss over Epochs



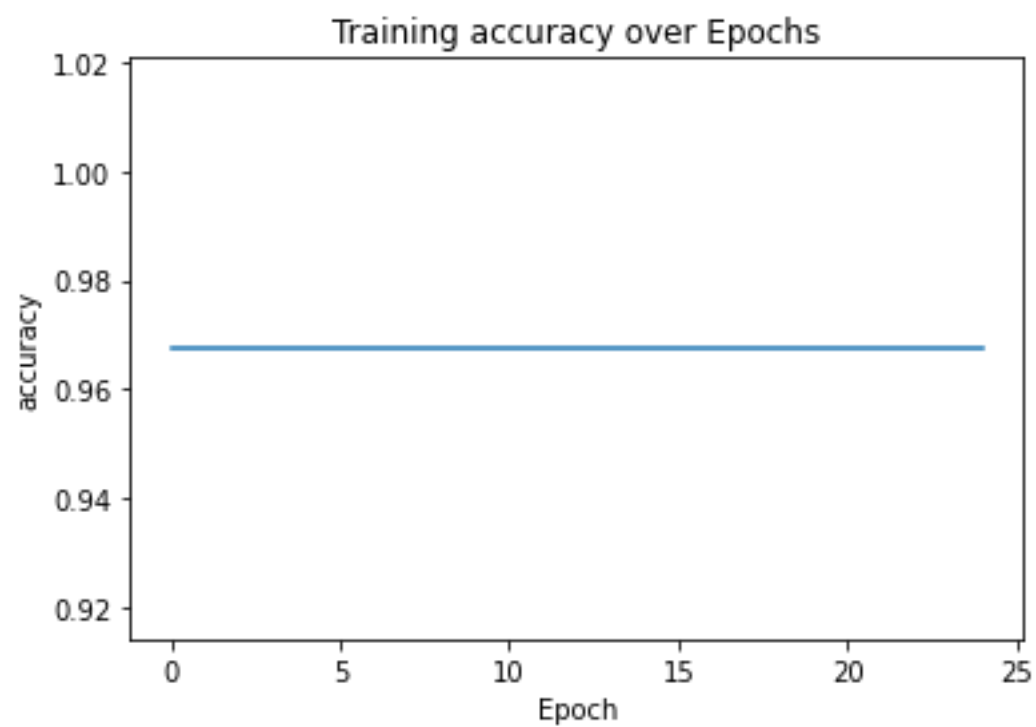
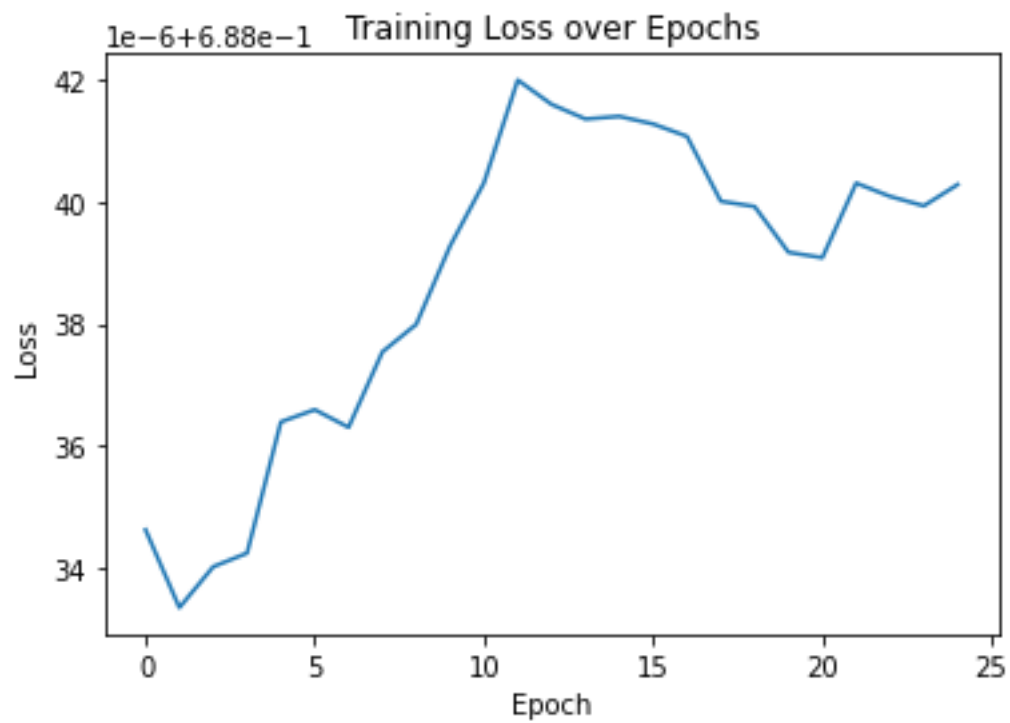
Training accuracy over Epochs

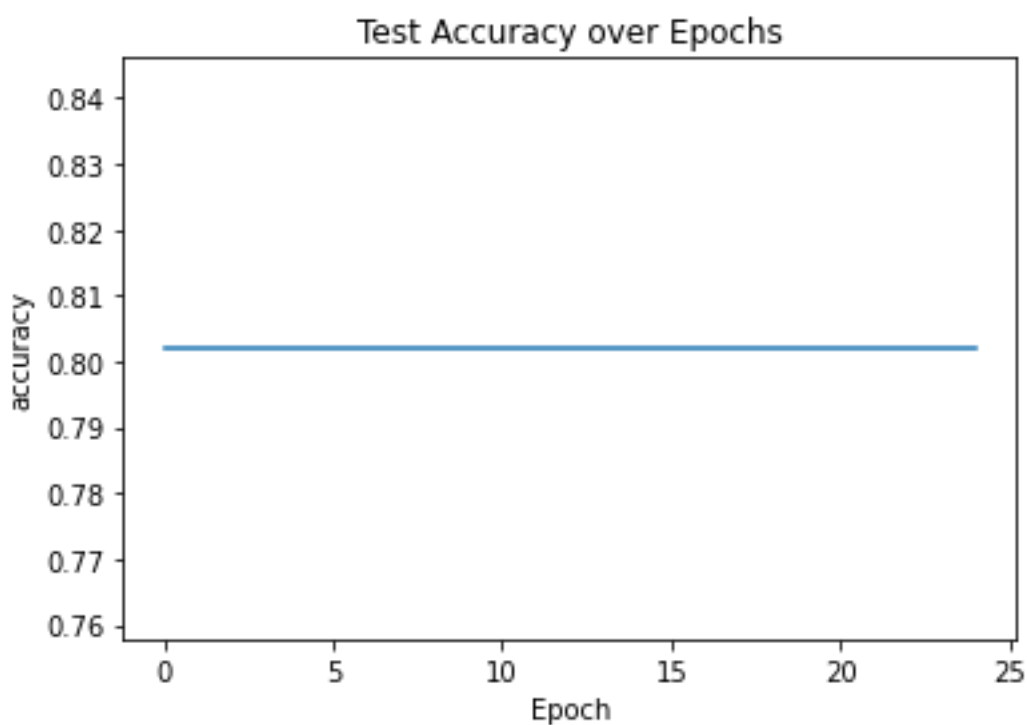
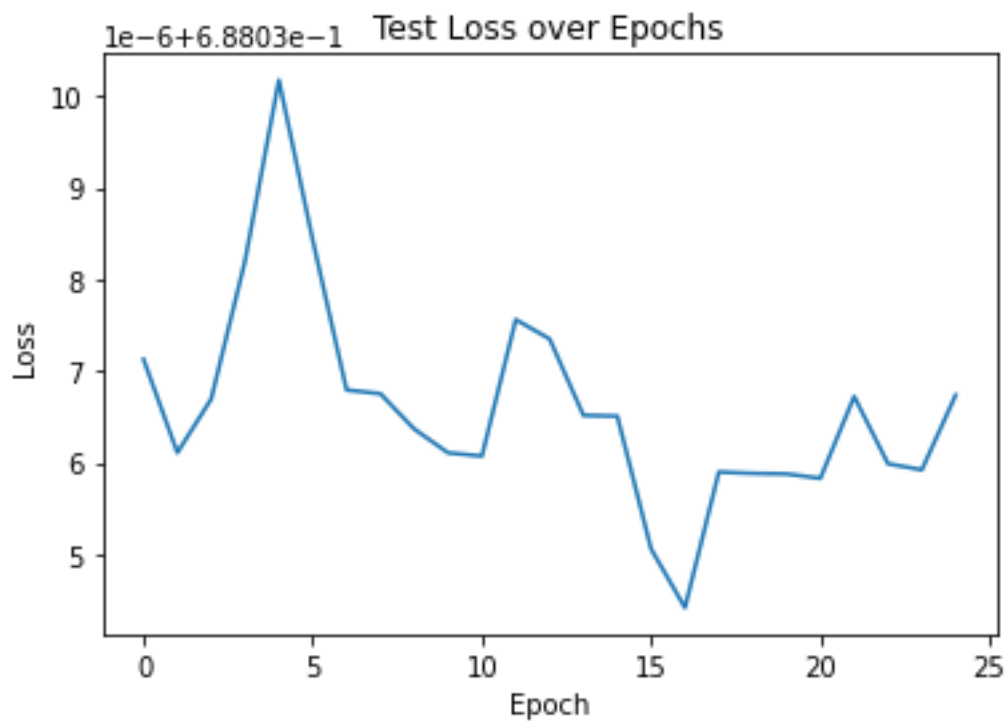




- چهار لایه با تانژانت هایپربولیک و تانژانت هایپربولیک و سیگموید و سیگموید.

```
train loss and accuracy
0.6880402749185512 0.9675
test loss and accuracy
0.6880367378380158 0.802
```



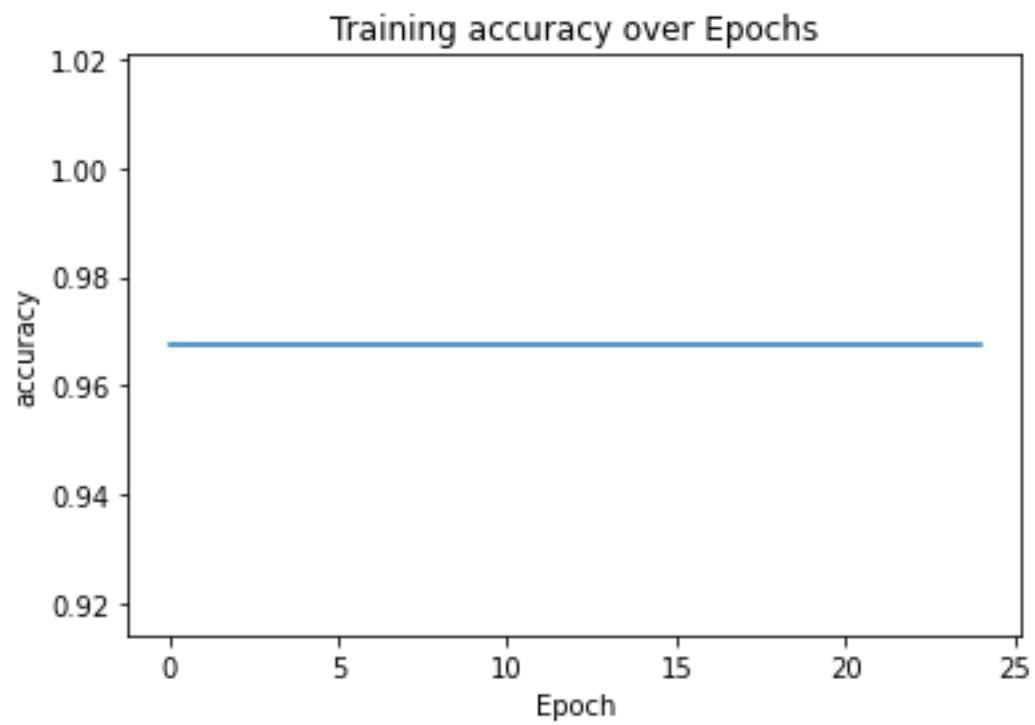
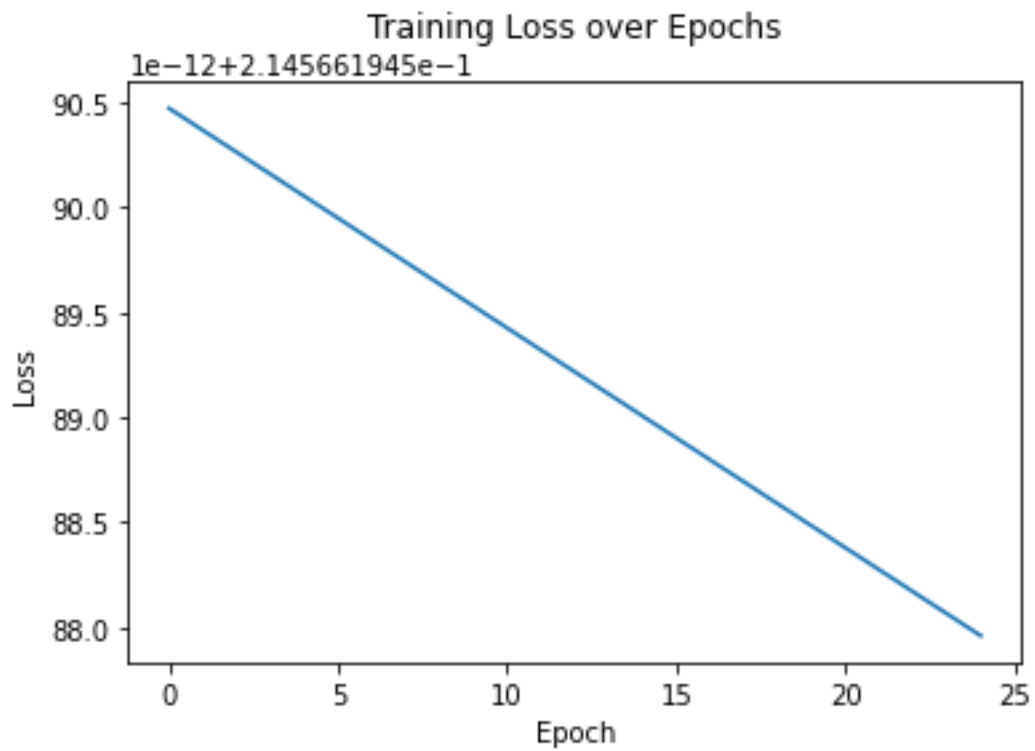


به نظر می‌رسد که در تمام موارد توابع فعالیت فرقی باهم نداشته و از ابتدا با دقت خوبی طبقه بندی انجام می‌شود.

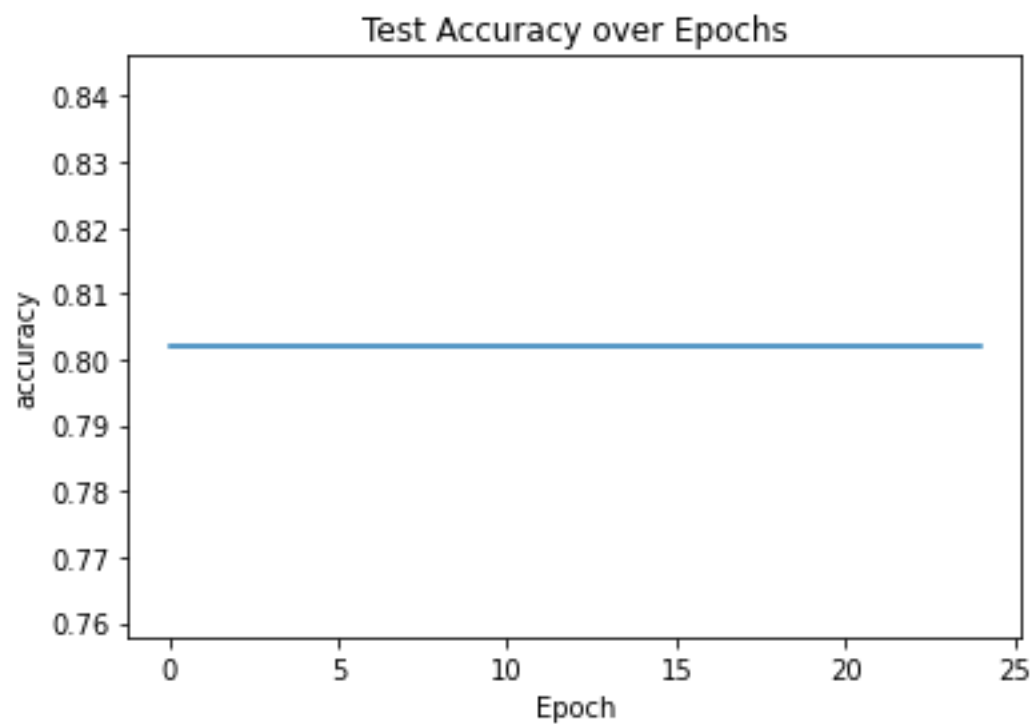
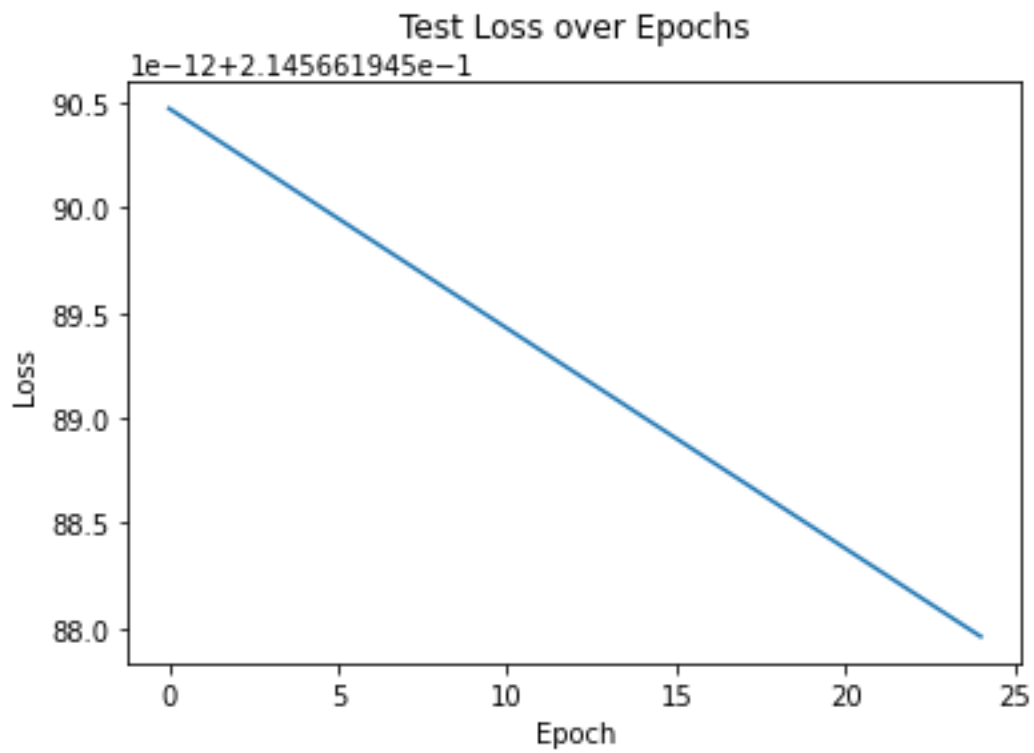
۲. بررسی تاثیر وزن دهی اولیه

- ابتدا وزن دهی تمام صفر

train loss and accuracy  
0.2145661945879601 0.9675  
test loss and accuracy  
0.2145661945879601 0.802



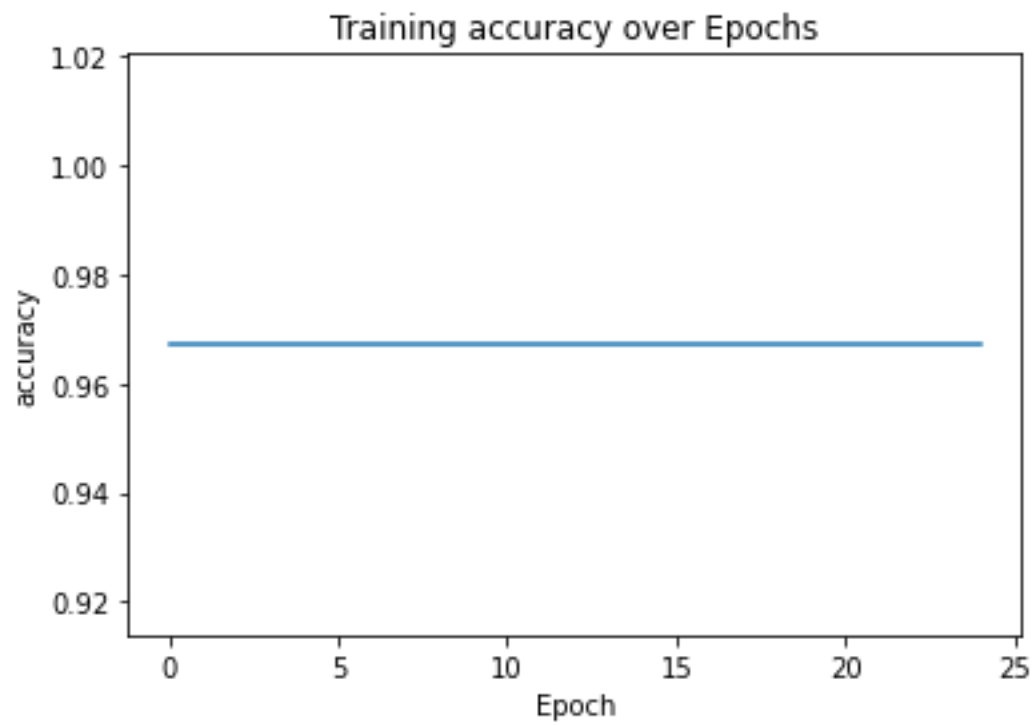
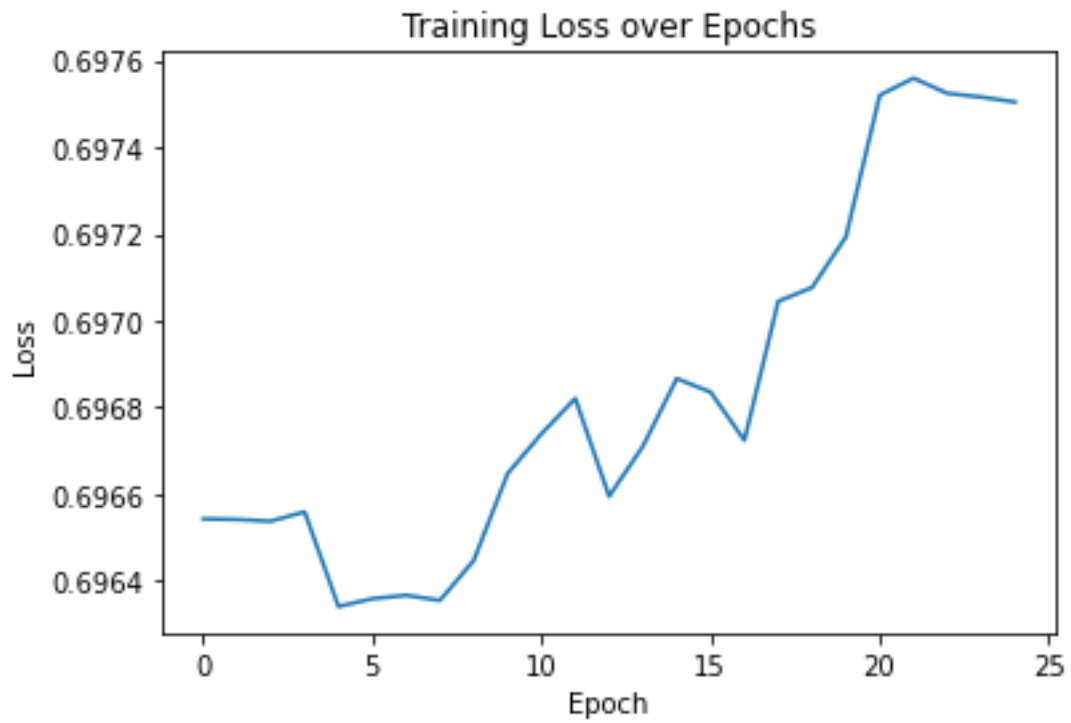


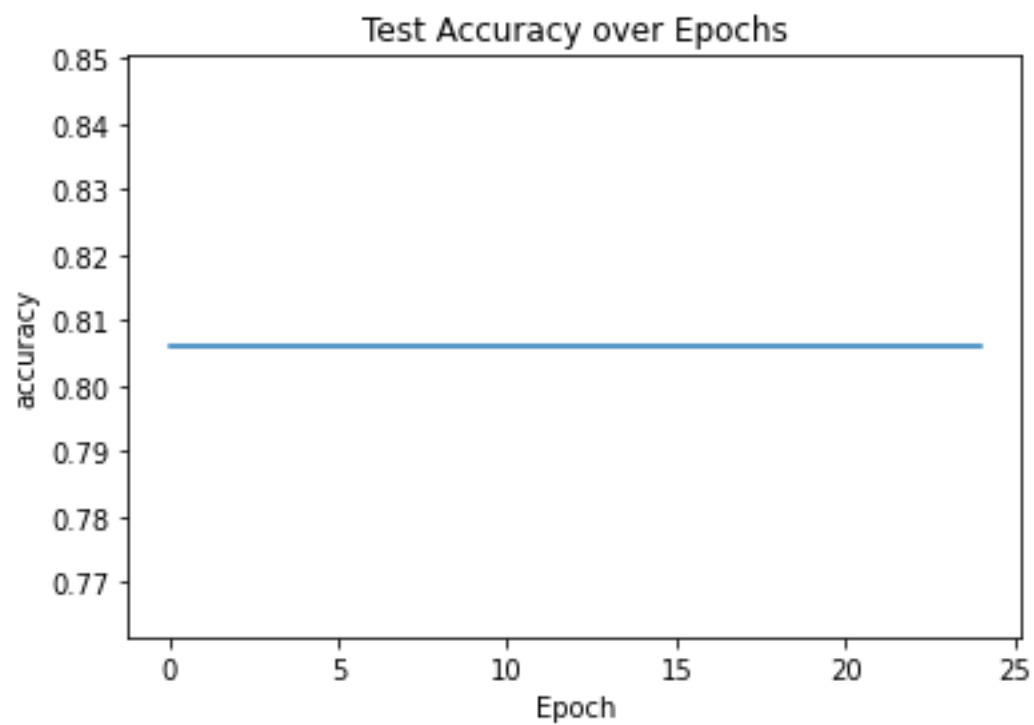
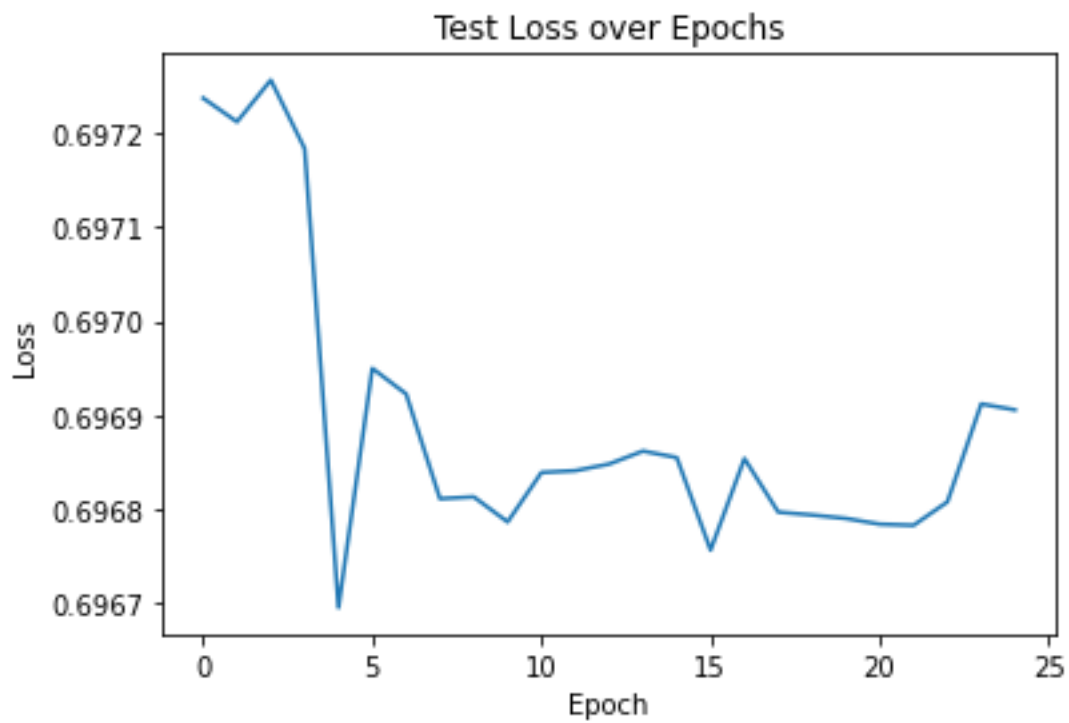


- وزن دهی مقادیر رندوم کوچک با توزیع نرمال

train loss and accuracy

0.6975041412420958 0.9671666666666666  
test loss and accuracy  
0.6969052942339031 0.806



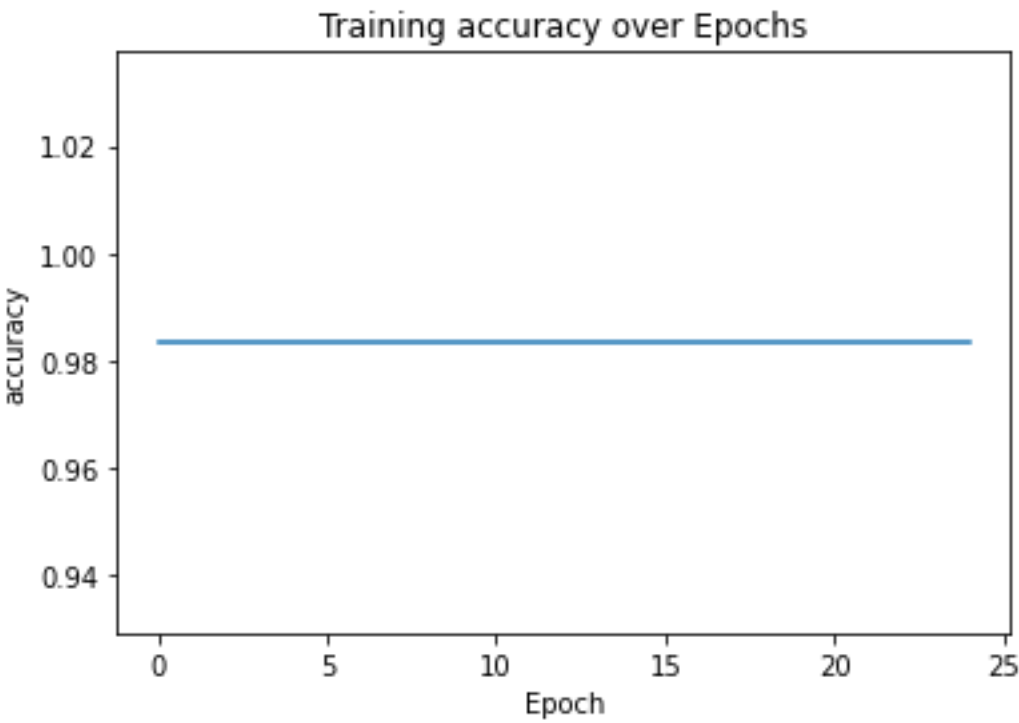
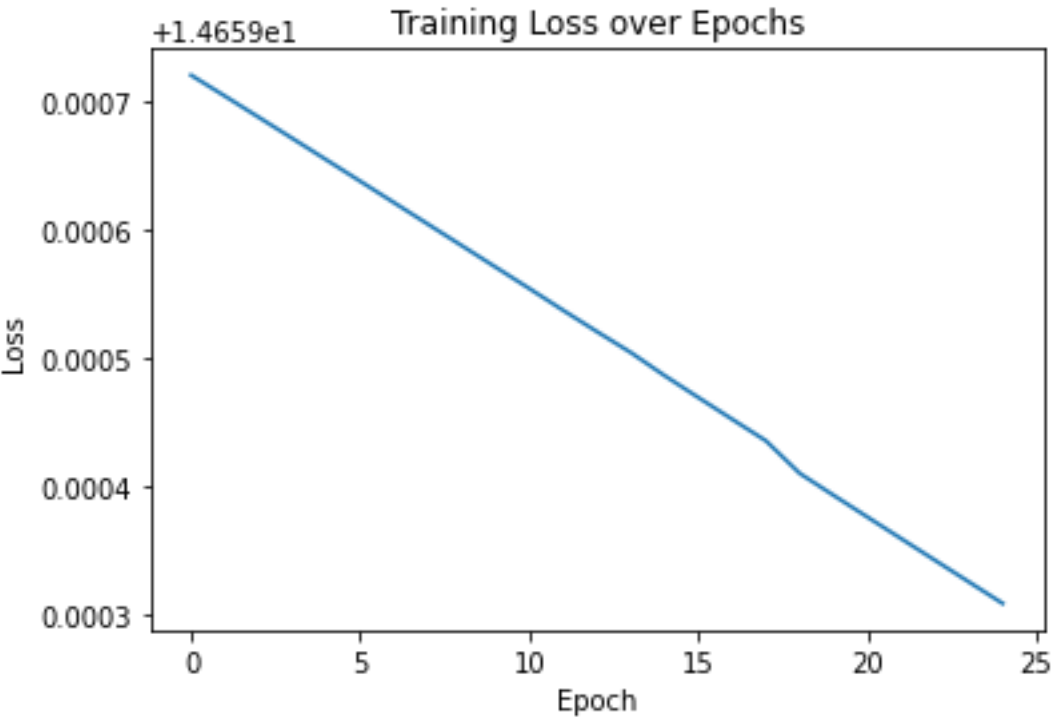


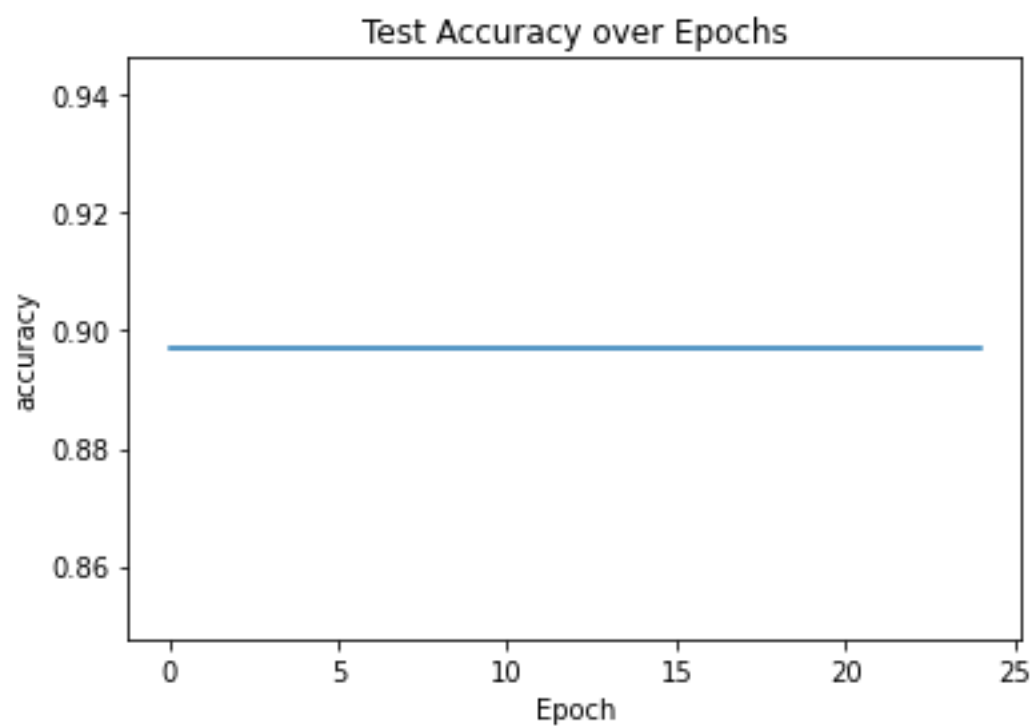
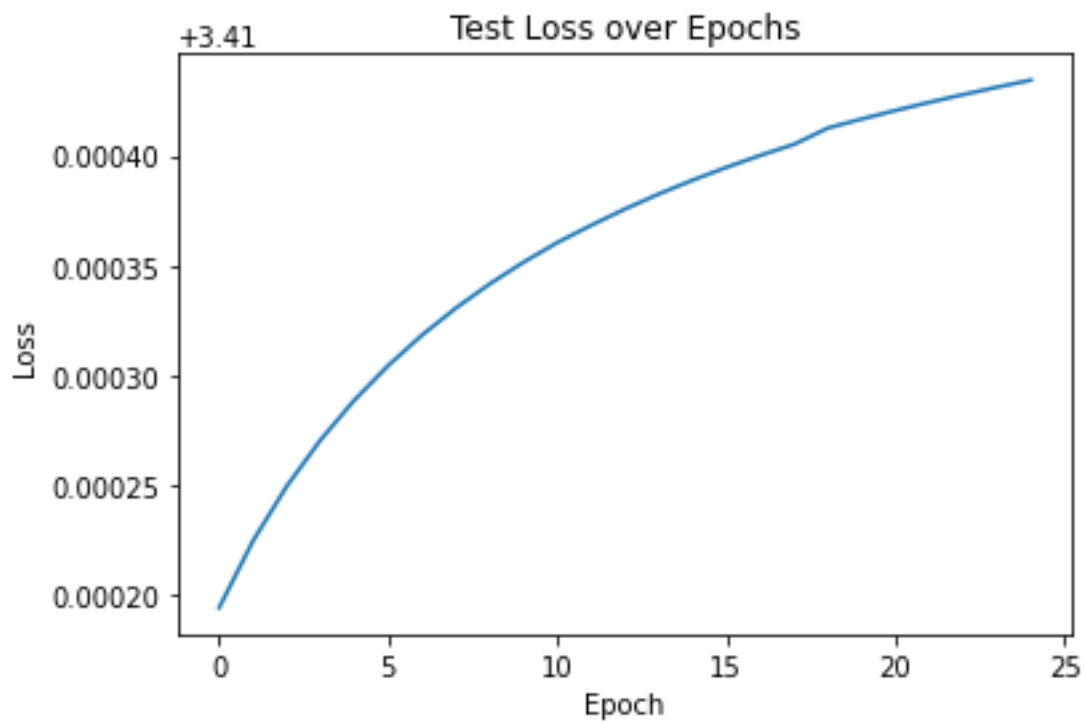
- مقادیر رندوم بزرگ با توزیع نرمال

train loss and accuracy

14.659309018910125 0.9834166666666667

test loss and accuracy  
3.4104343571486644 0.897

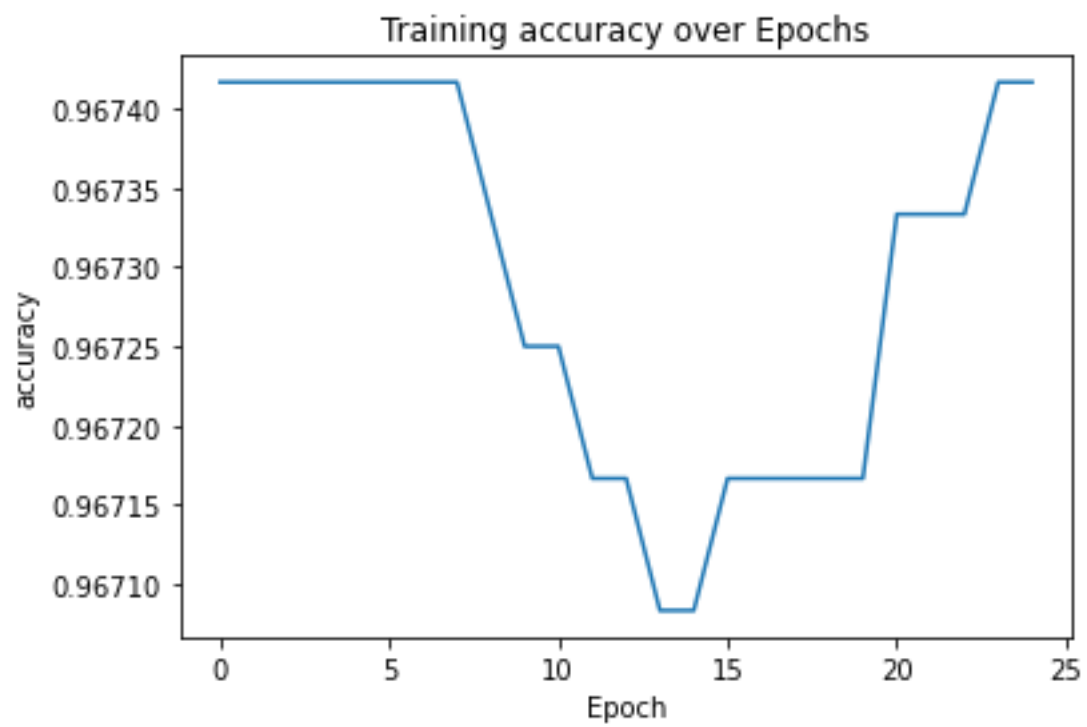
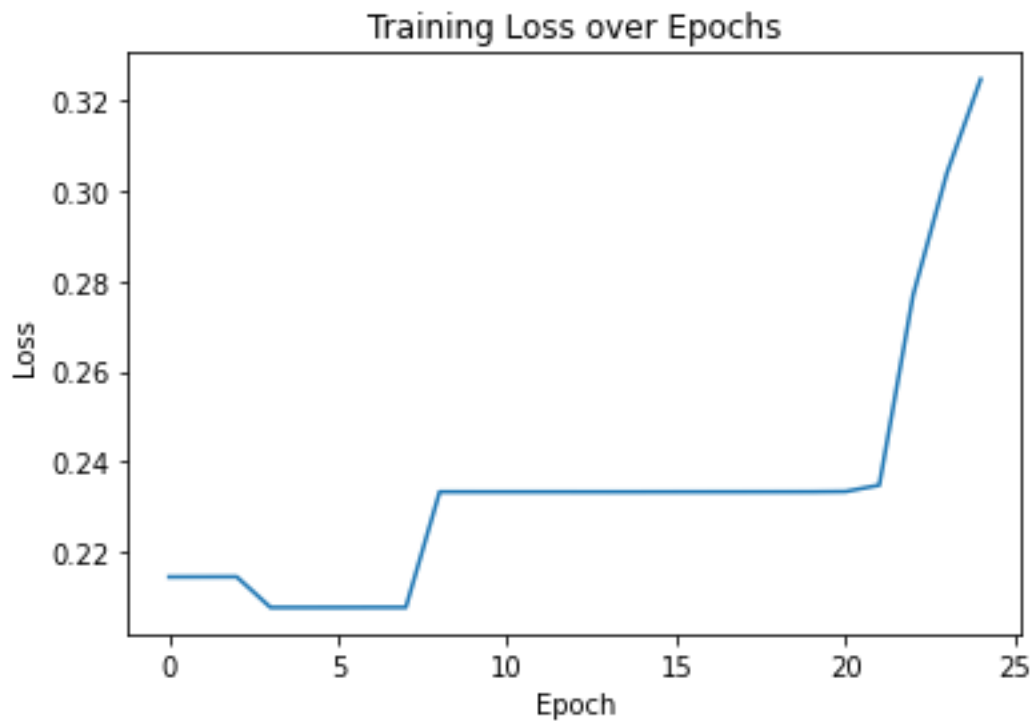


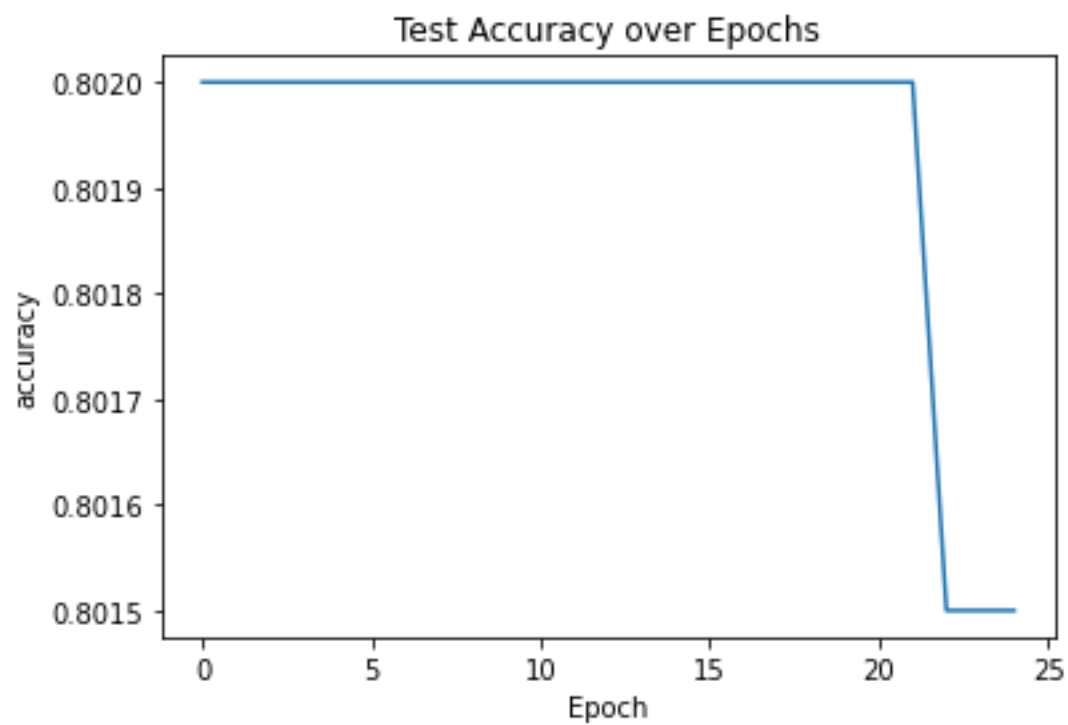
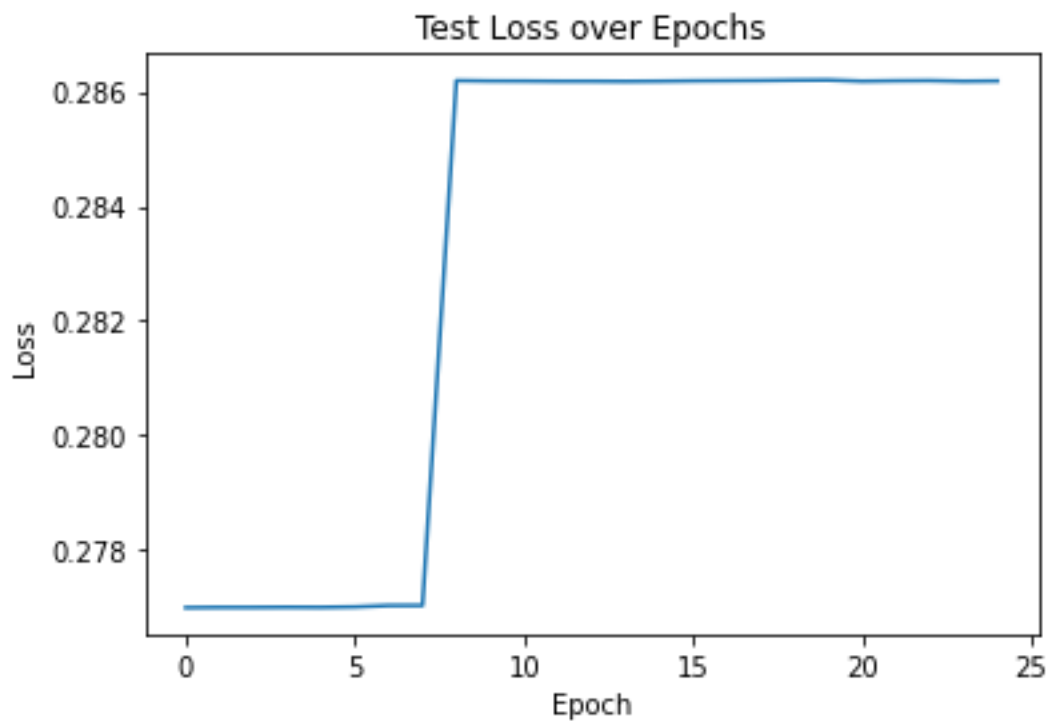


• وزن دهی Xavier

train loss and accuracy  
0.32466002397578714 0.9674166666666667

test loss and accuracy  
0.28618426171393097 0.8015

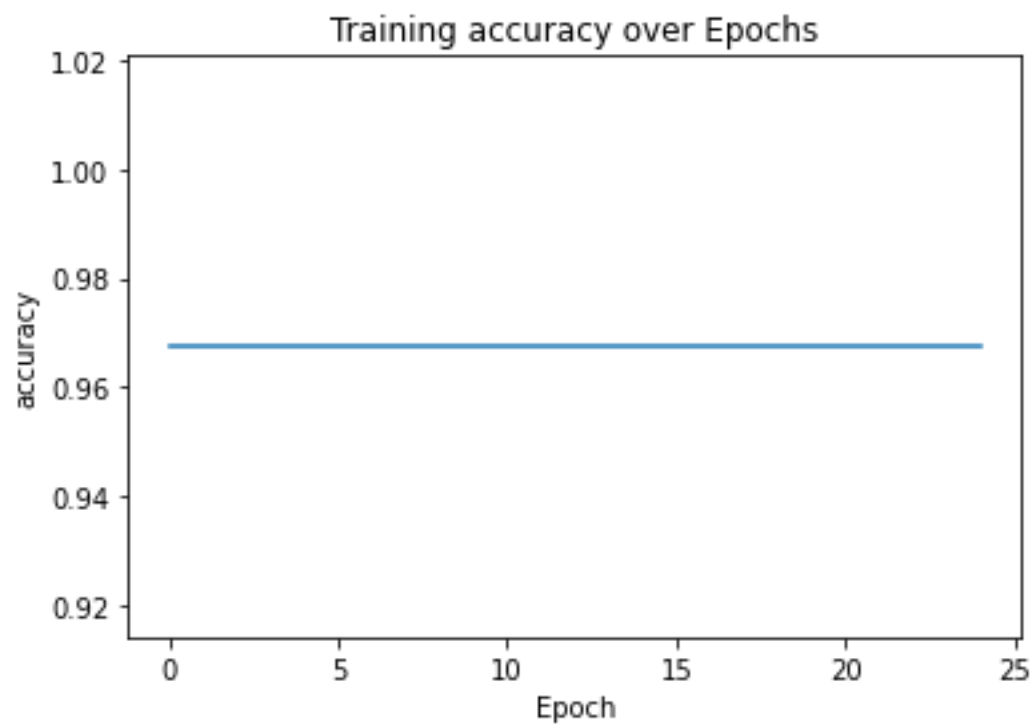
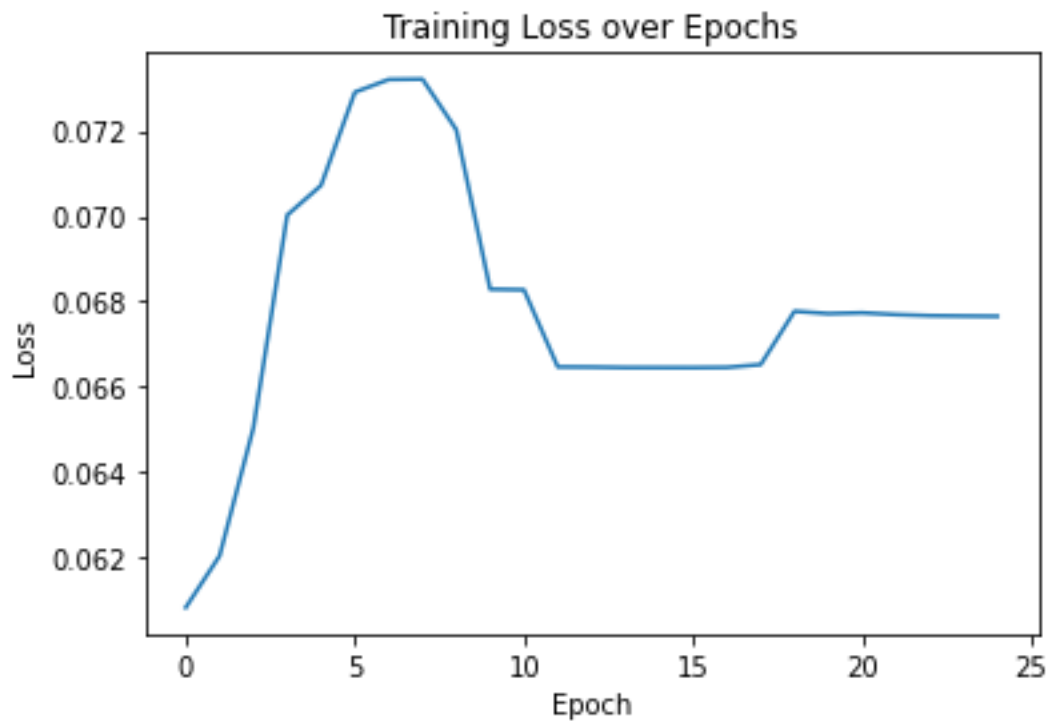




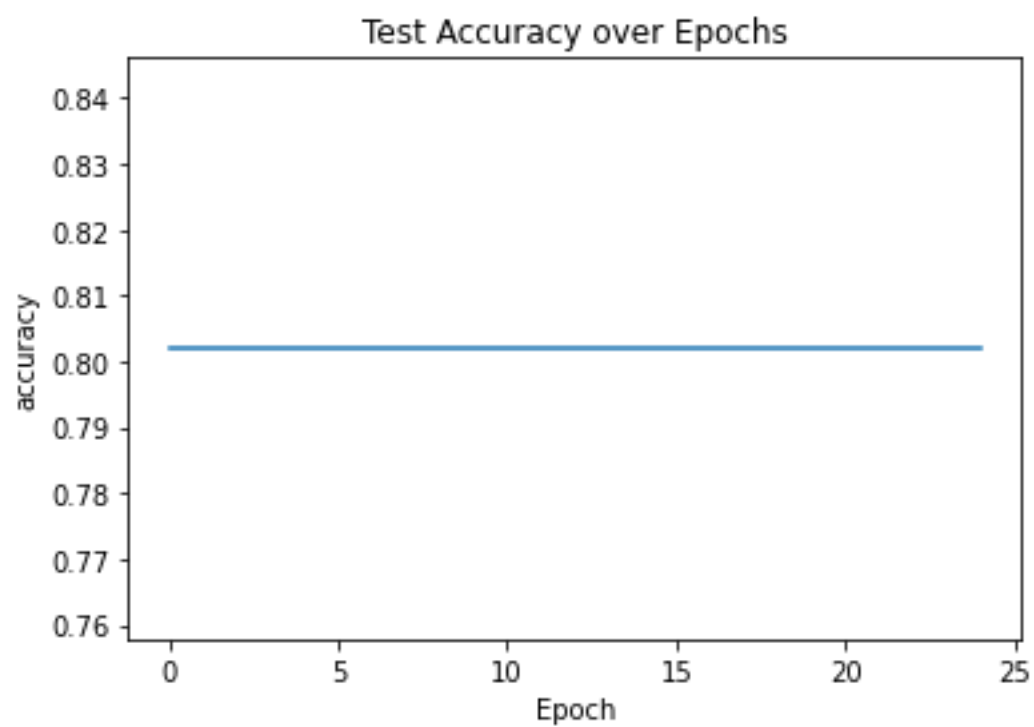
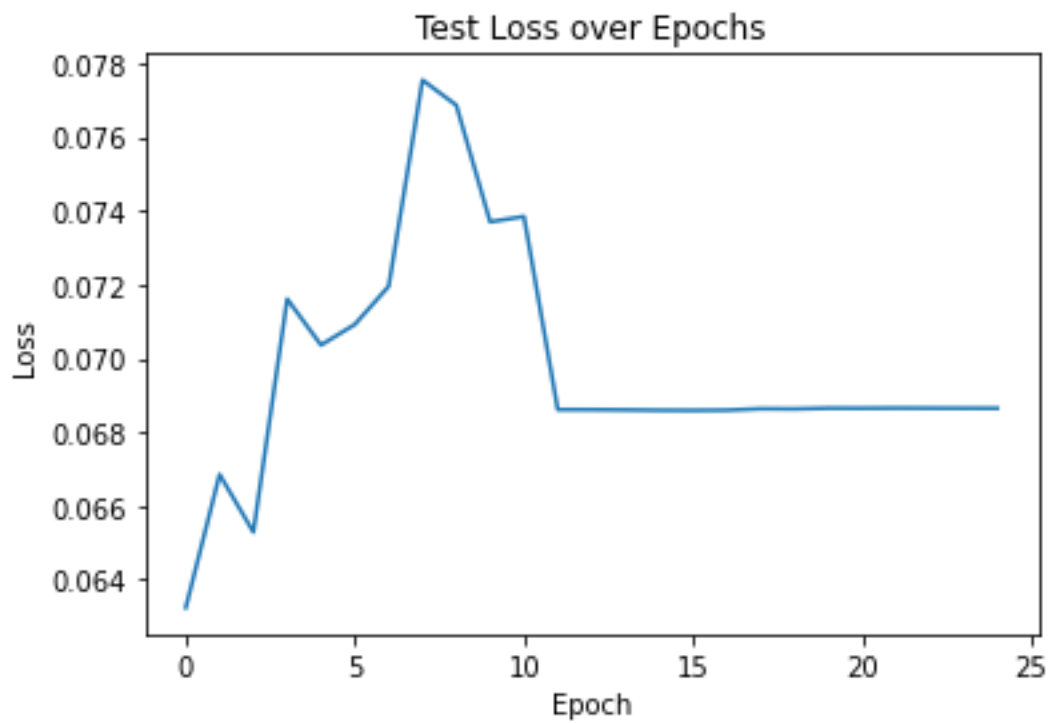
• وزن دهی He

train loss and accuracy  
0.06765118336644273 0.9675

test loss and accuracy  
0.0686485542961549 0.802







۳. بررسی بهینه سازها

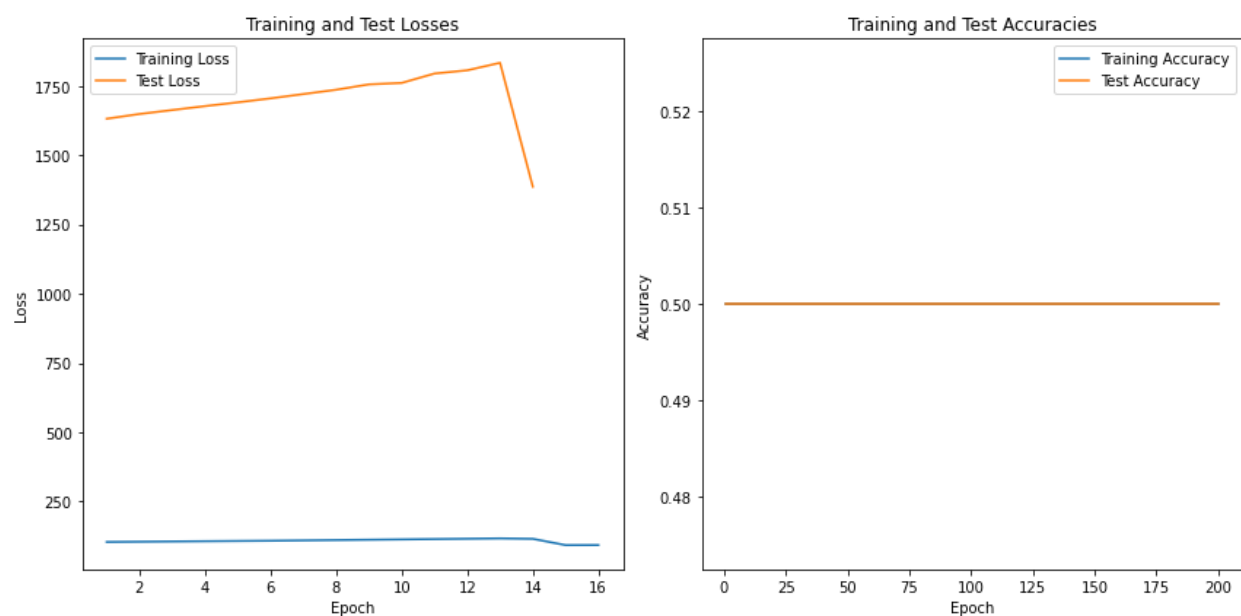
• SGD

که البته در چند آزمایش شاهد افزایش میزان ضرر بودیم و عدم یادگیری شبکه و در نهایت هم با یک شبکه ۴ لایه و اندازه بچ ۳۲ هم به نتیجه مثبتی نرسید.



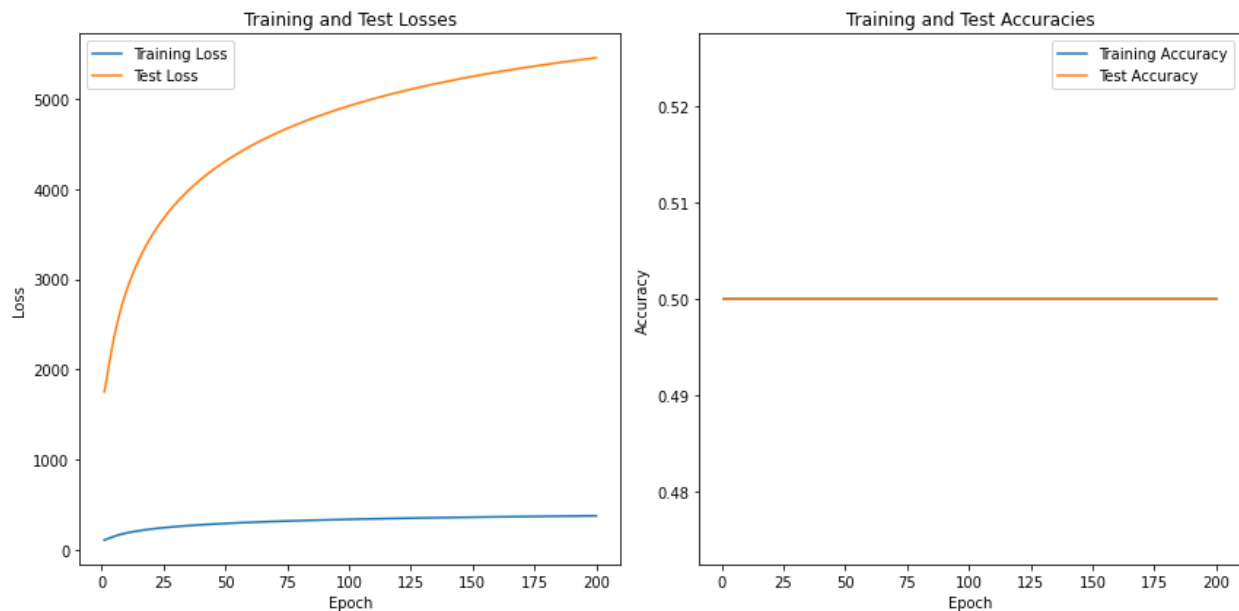
### SGD with momentum •

متأسفانه بدلیل پیچیدگی کدها و پیچیدگی شبکه و ناپدید شدن گرادیان‌ها و وجود مقادیر ناچیز، همواره به مقادیر Nan در برخی جاها و در نتیجه در loss رسیدیم که مانع از یادگیری شبکه می‌شود.



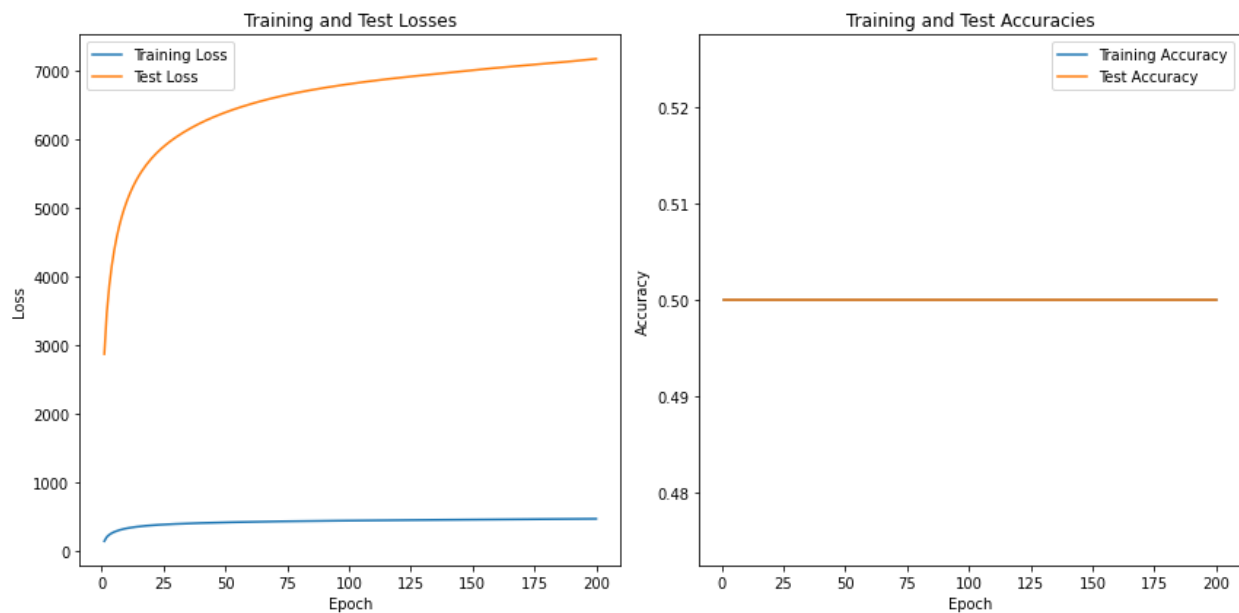
مقدار ضرر از اپوک ۱۶ به بعد نمایش داده نشده چون Nan بدست آمد.

## RMS prop •



ضرر همواره بیشتر شد و یادگیری انجام نشد.

## Adam •



در این حالت هم یادگیری انجام نشد.

۴. بررسی تاثیر یک بهینه ساز با و بدون L2 Regularization

این حالت هم به نتیجه‌ای نرسید.

