

به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده مهندسی برق و کامپیوتر



یادگیری عمیق

تمرین دوم

محمد مهدی دارابی

810195083

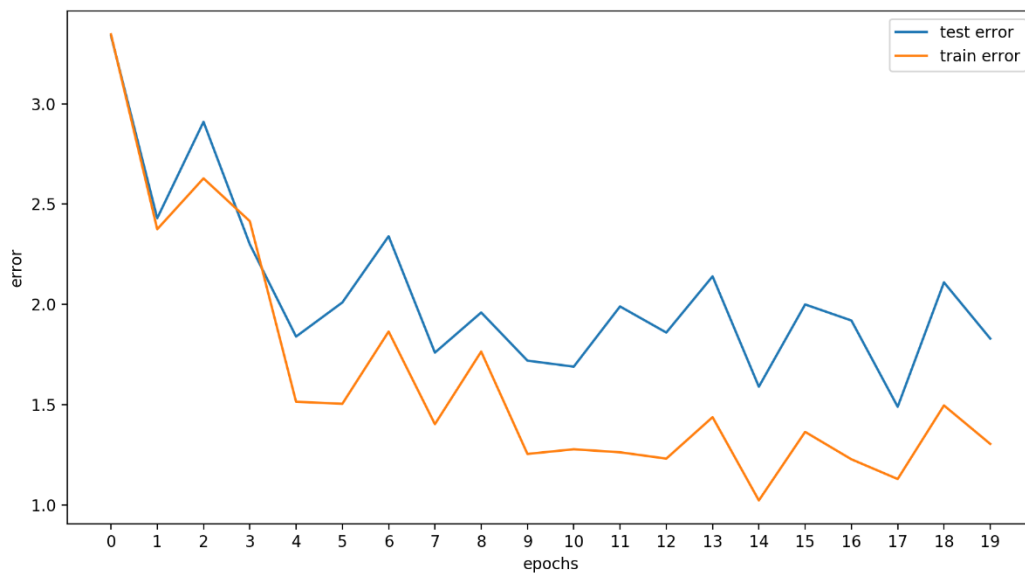
Contents

سوال 1	3
قسمت 1-1	3
قسمت 1-2	4
قسمت 1-3	5
قسمت 1-4	6
قسمت 1.5	8
قسمت 1.6	10
قسمت 1.7	11
قسمت 1.8	13
سوال 2	14
قسمت 2.1	14
قسمت 2.2	15
قسمت 2.3	16
قسمت 2.4	17
قسمت 2.5.1	19
قسمت 2.5.2	20
قسمت 2.5.3	21
قسمت 2.5.4	22
منابع	24

سوال 1

قسمت 1-1

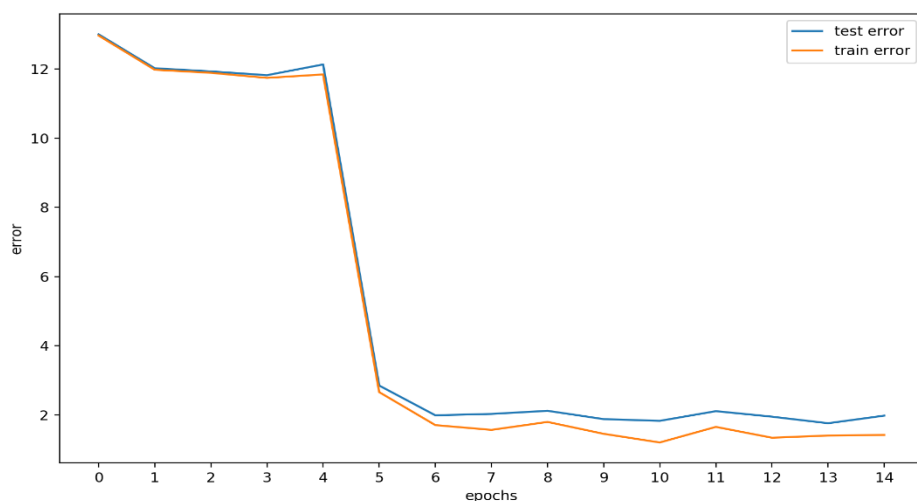
با توجه به نمودار خطای train و validation به دست آمده، می‌توان به این نتیجه رسید که تعداد 12 الی 14 اپوک برای یادگیری مناسب باشد زیرا در ادامه دقت مدل تنها دارای نوسان می‌شود.



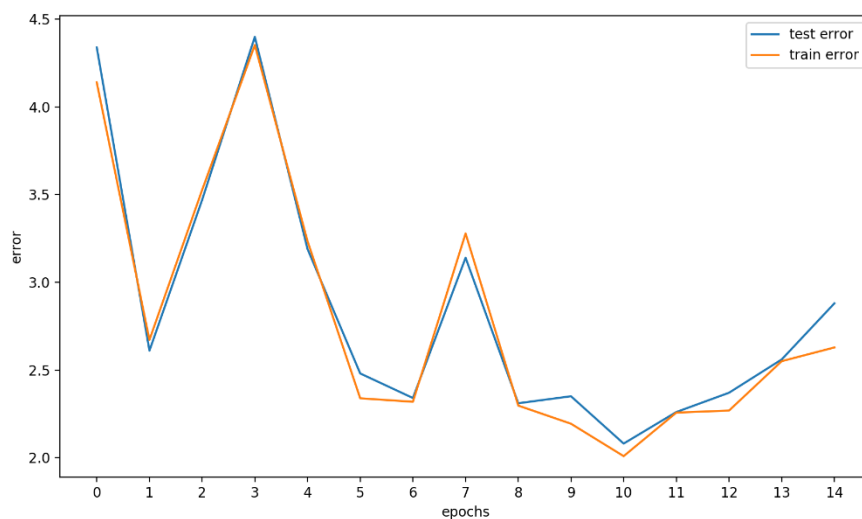
شکل 1 خطای داده‌های train و test

قسمت 1-2

نمودار خطای train و validation برای دو مدل با kernel size به اندازه‌ی 5x5 و 7x7 به ترتیب آمده است. همان‌طور که مشاهده می‌شود مدل 5 در 5 بدون نوسان و با تعداد اپوک کمتر به جواب مناسب می‌رسد اما مدل 7 در 7 پس از کاهش و رسیدن به جواب مناسب دچار نوسان می‌شود.



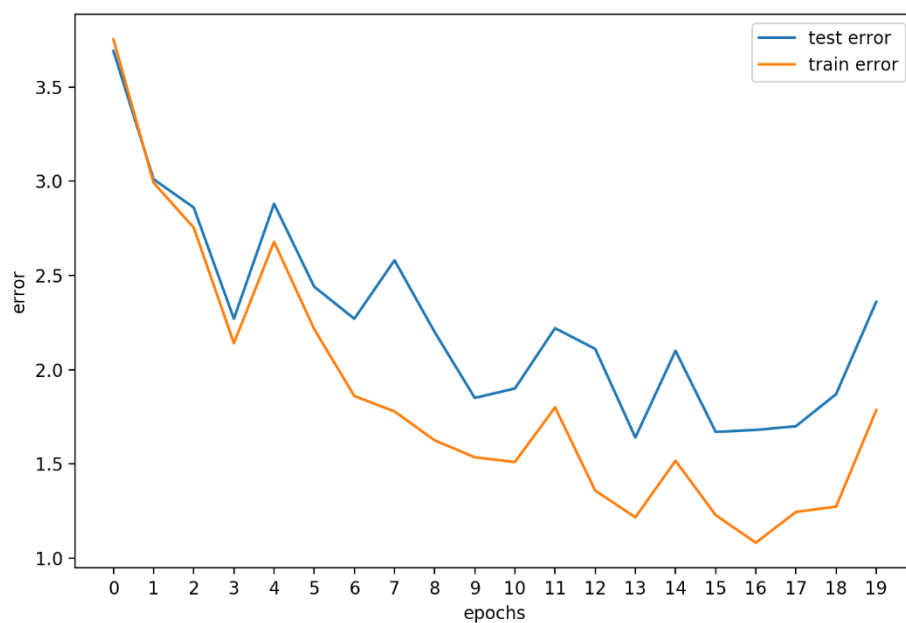
شکل 2- خطای داده‌ی train و test با kernel size 5 در 5



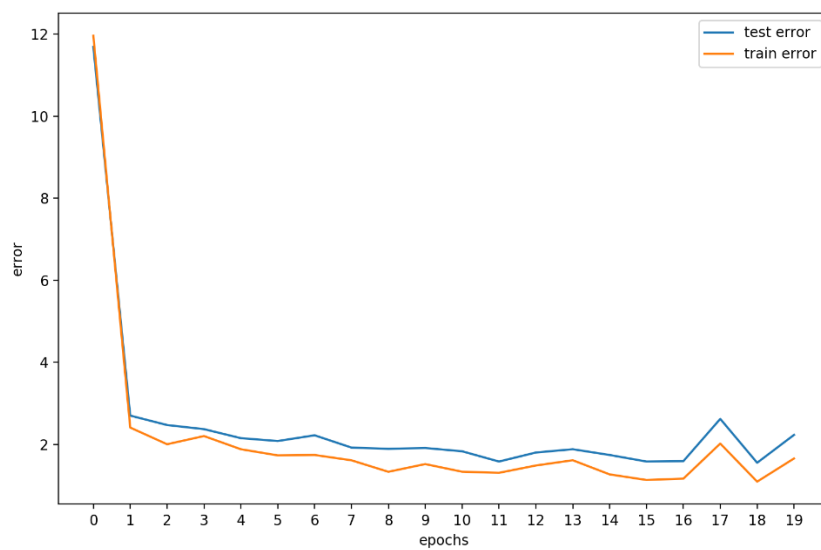
شکل 3- خطای داده‌ی train و test با kernel size 7 در 7

قسمت 1-3

با توجه به نمودارهای خطای نشان داده شده به نظر می‌رسد که **mean pooling** با اندازه 3 بسیار سریع‌تر از اندازه‌ی 2 به جواب می‌رسد و همچنین اندازه 2 با نوسان بیشتری به جواب نزدیک می‌شود.



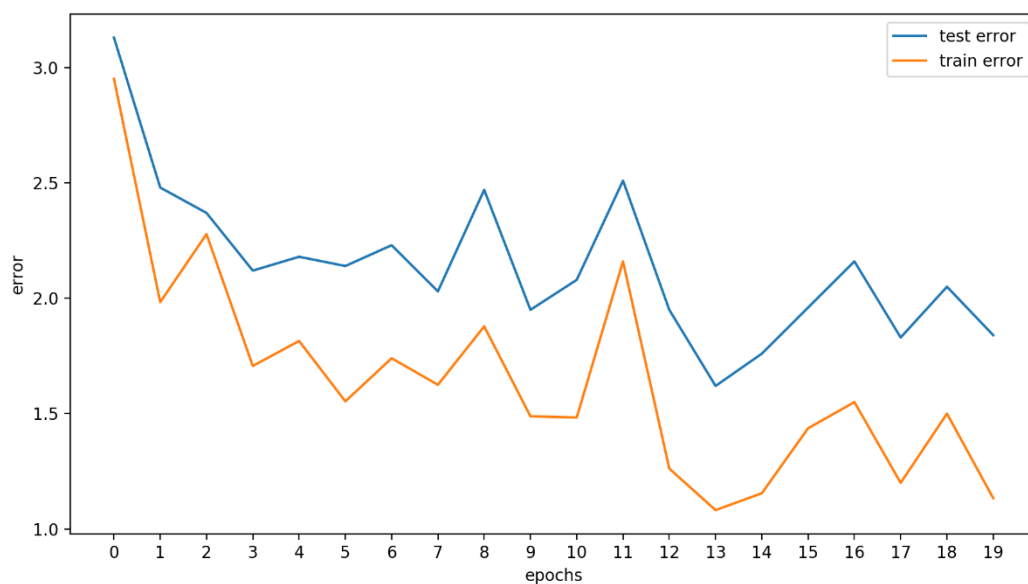
شکل 4 نمودار خطای *train* و *test* با استفاده از *mean pooling*



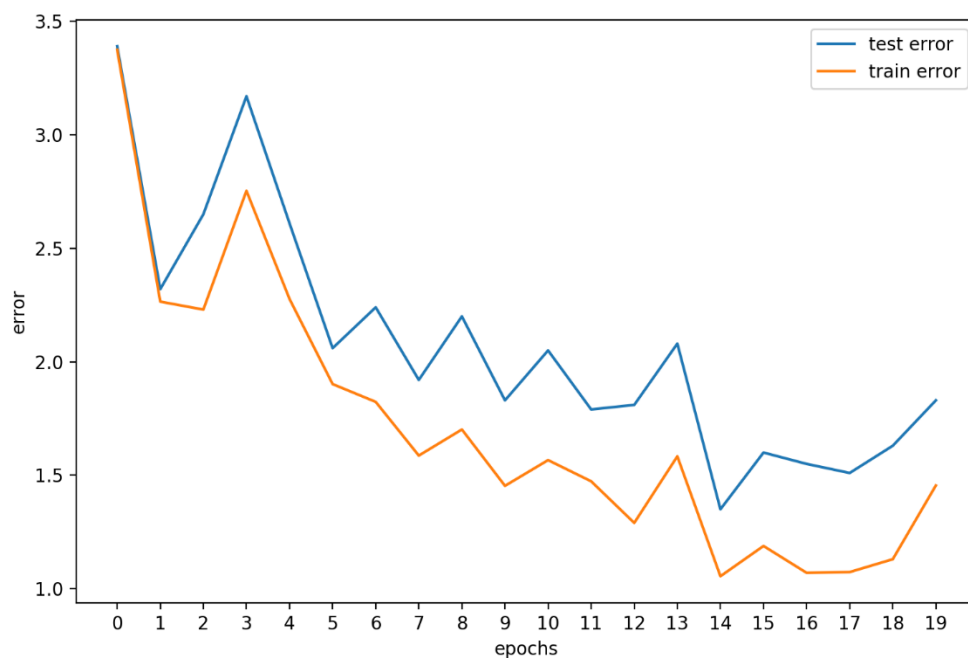
شکل 5 نمودار خطای *train* و *test* با استفاده از *mean pooling* با اندازه‌ی 3

قسمت 1-4

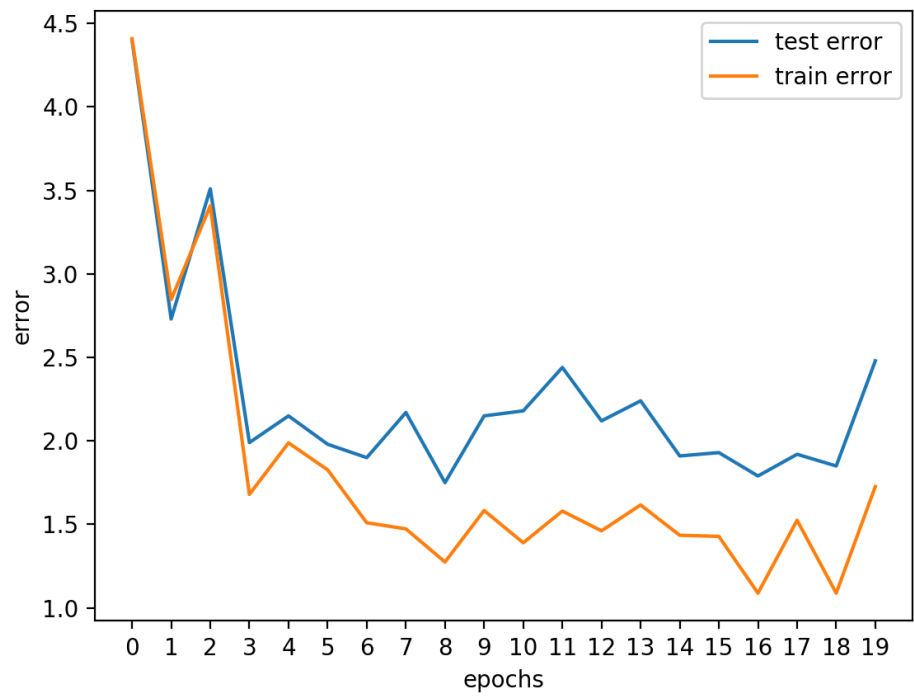
نمودار خطا با استفاده از dropout با مقادیر 0 و 0.25 و 0.5 به ترتیب زیر آمده است. دقت مدل‌های به دست آمده به ترتیب 98.7 و 98.9 و 98.7 است.



شکل 6 - نمودار خطای train و test با dropout صفر



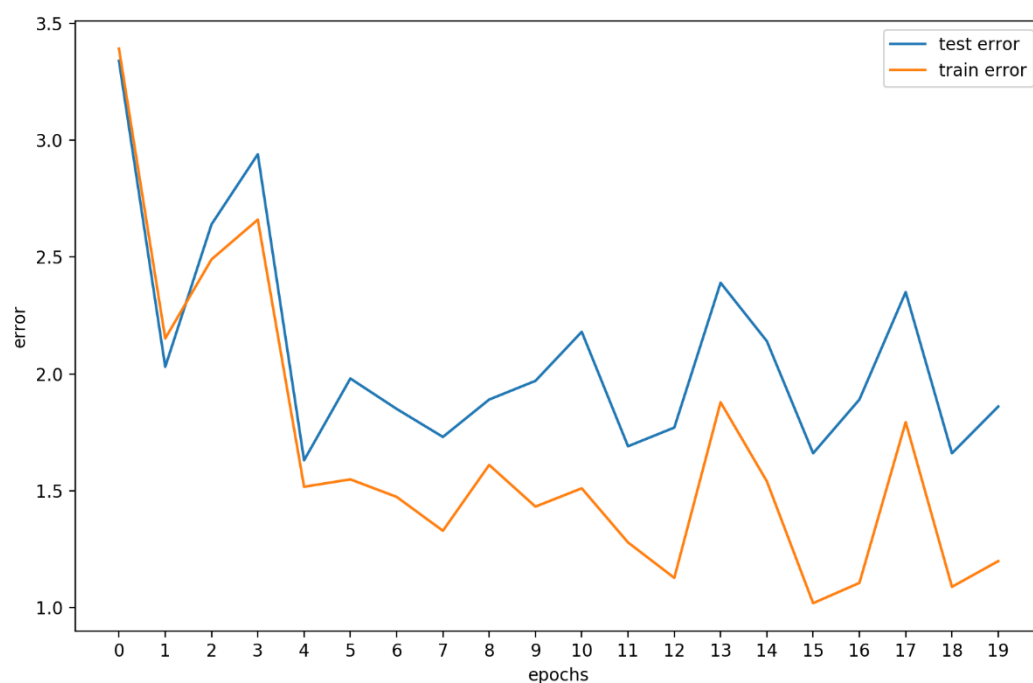
شکل 7 نمودار خطای train و test با dropout 0.25

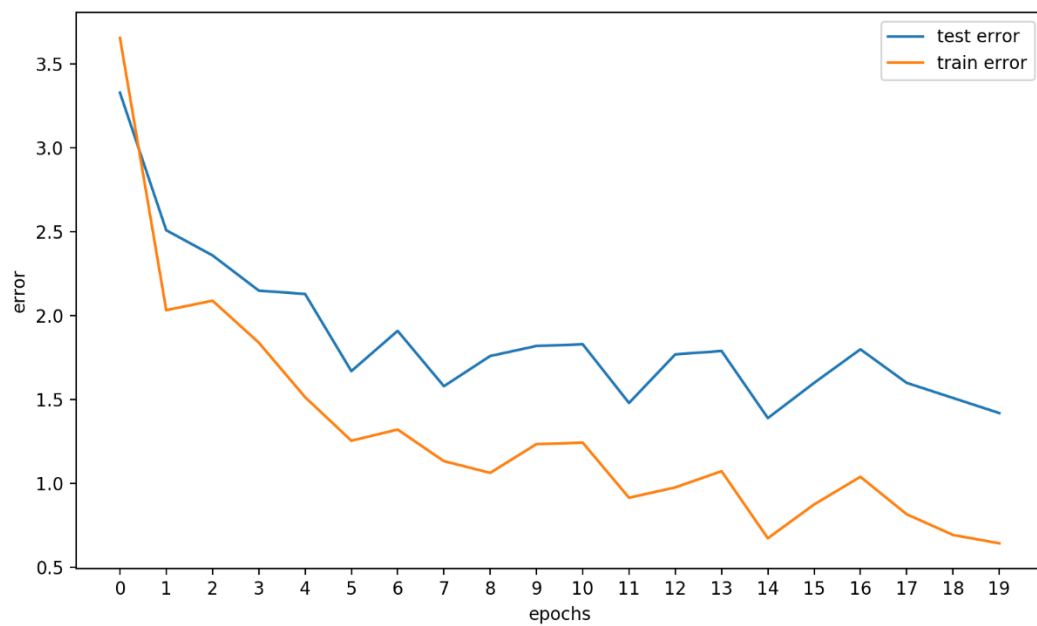
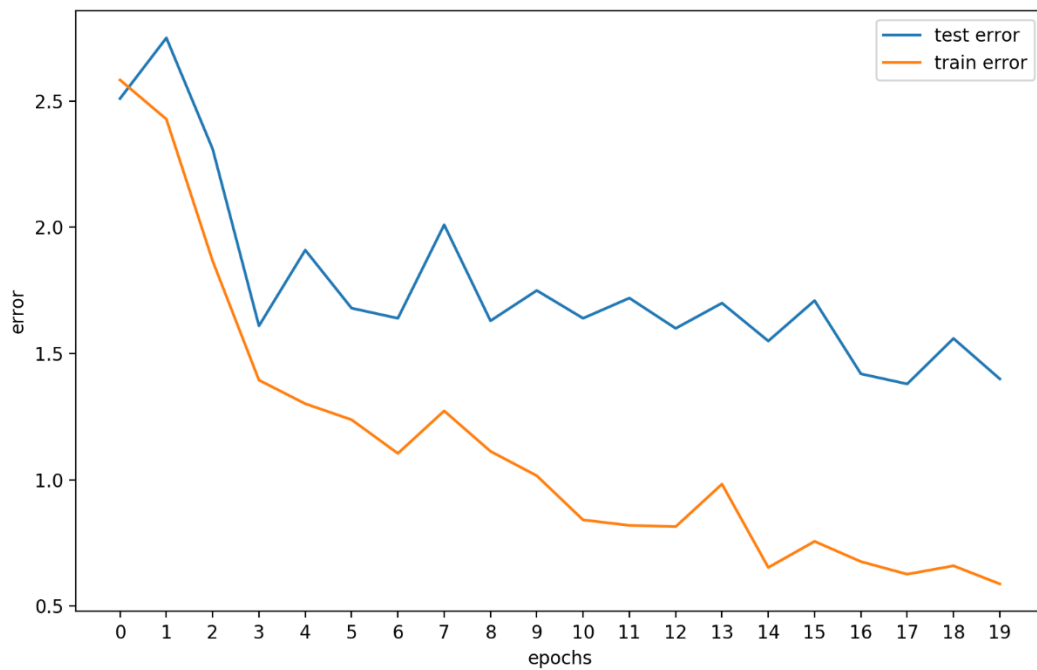


شکل 8 نمودار خطای $train$ و $test$ با 0.5 dropout

قسمت 1.5

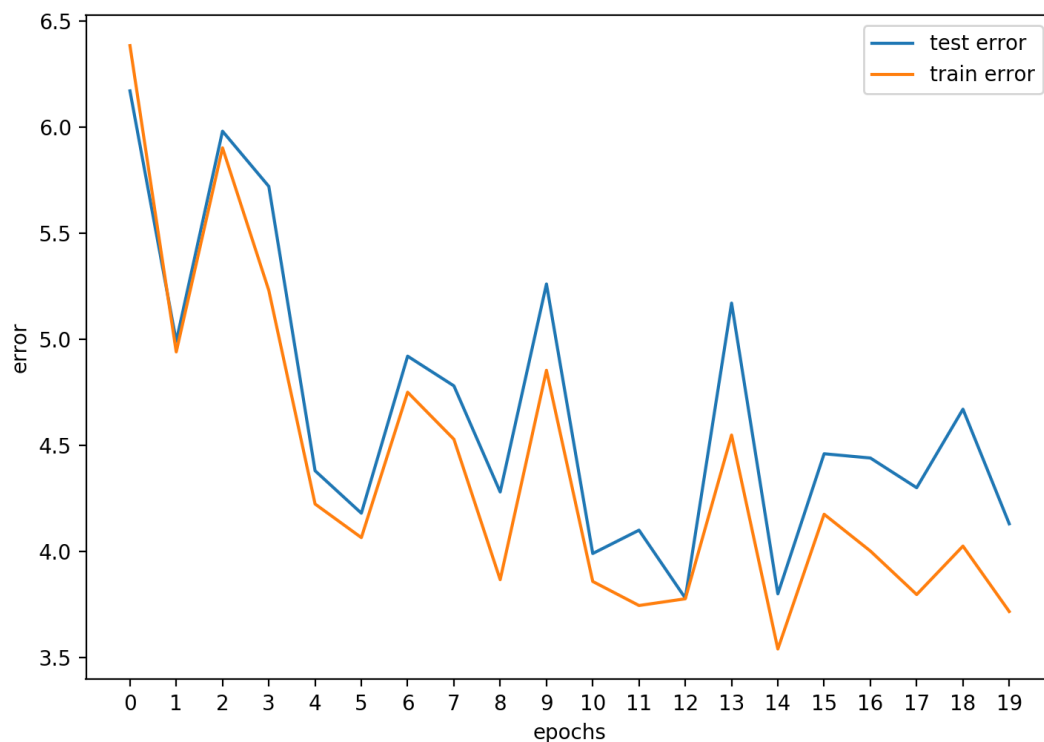
نمودار خطا با استفاده از batch normalization بعد از لایه اول و بعد از لایه‌ی دوم و بعد از هر دو لایه به ترتیب آمده است. زمان اجرا برای هنگامی که از batch normalization در لایه‌ی دوم استفاده می‌کنیم کمتر از دو حالت دیگر است و برابر با 1440 ثانیه است و زمان اجرای دو حالت دیگر تقریباً با یکدیگر برابر است. دقت اجرای مدل برای دو حالت آخر از حالت اول بهتر بوده است.





قسمت 1.6

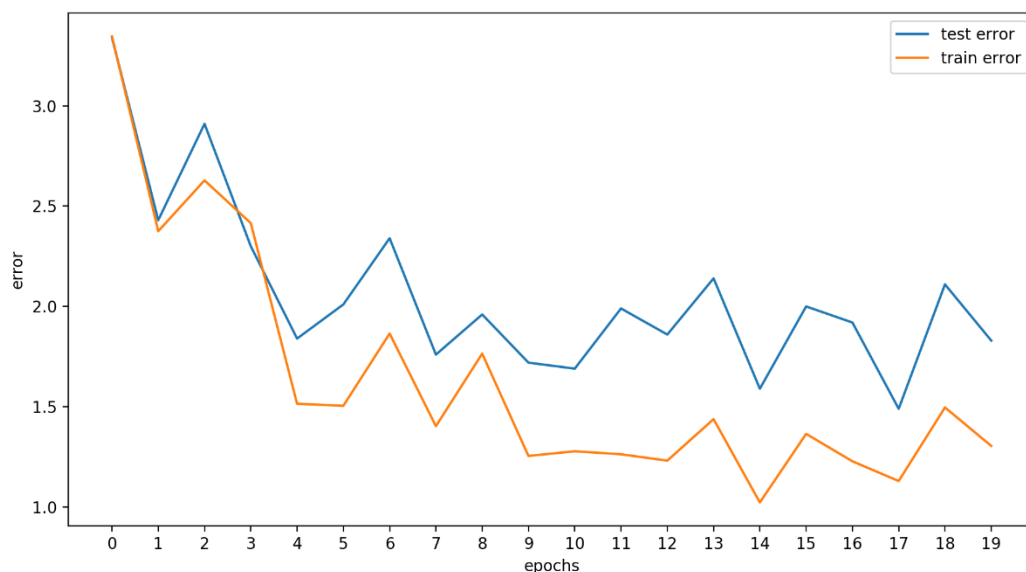
نمودار خطای مدل با لایه‌ی کانولوشن با سایز 5 در 5 در شکل زیر آمده است. همان طور که مشاهده می‌شود نتیجه‌ی به دست آمده از حالت اول بدتر شده است. این به این علت است که دو لایه با اندازه کم توانایی مدل‌سازی پیچیده‌تر نسبت به یک لایه بزرگ‌تر را دارند.



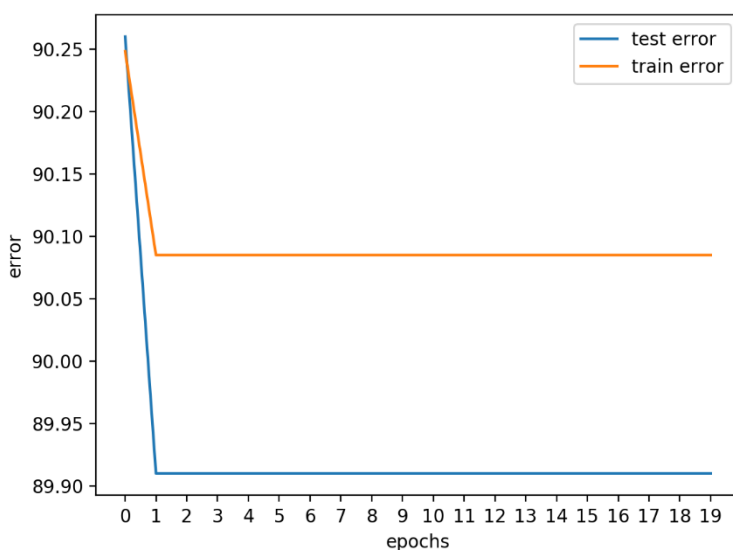
شکل 9 - نمودار خطای $train$ و $test$ با استفاده از یک لایه کانولوشن 5 در 5

قسمت 1.7

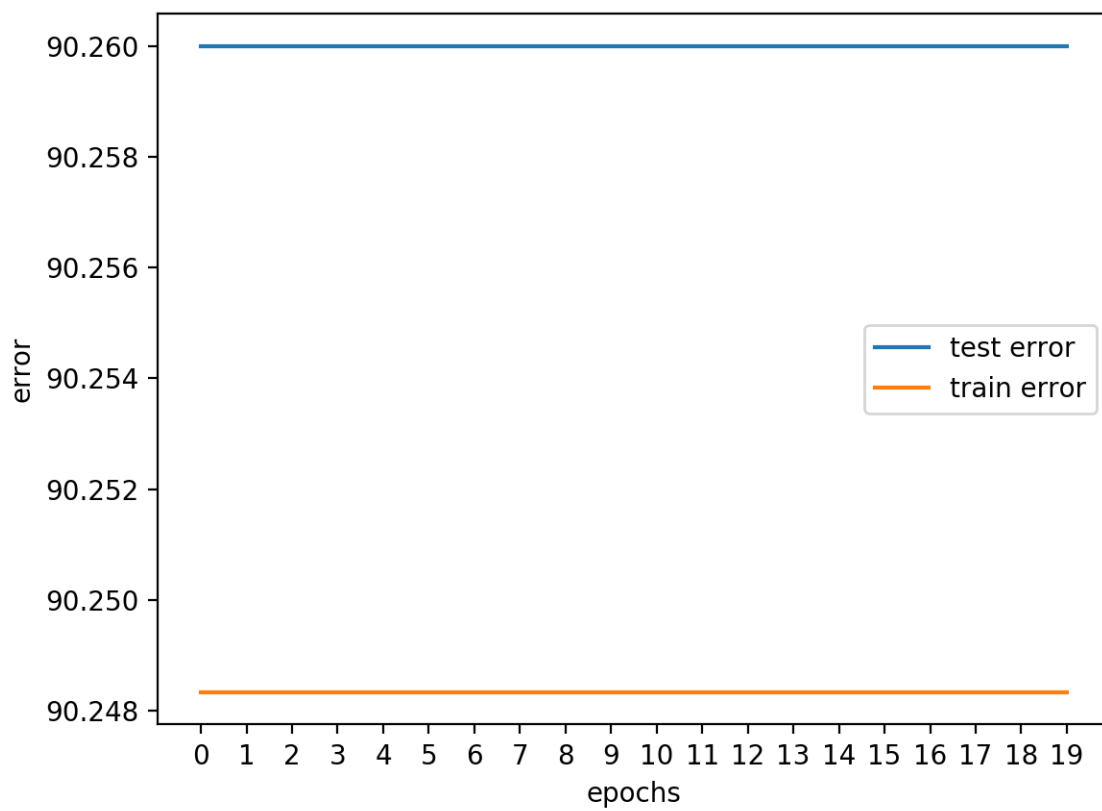
نمودار خطا برای مدل با مقادیر نرخ یادگیری 10^{-3} و 10^{-2} و 10^{-1} به ترتیب آمده است. هر چه نرخ یادگیری بزرگتر شود، توانایی مدل برای به دست آوردن جواب مناسب کم‌تر می‌شود و هنگامی که این نرخ برابر با 10^{-1} است، هیچ بهبودی نسبت به جواب رندم اولیه دیده نمی‌شود.



شکل 10 نمودار خطا با نرخ یادگیری یک هزارم



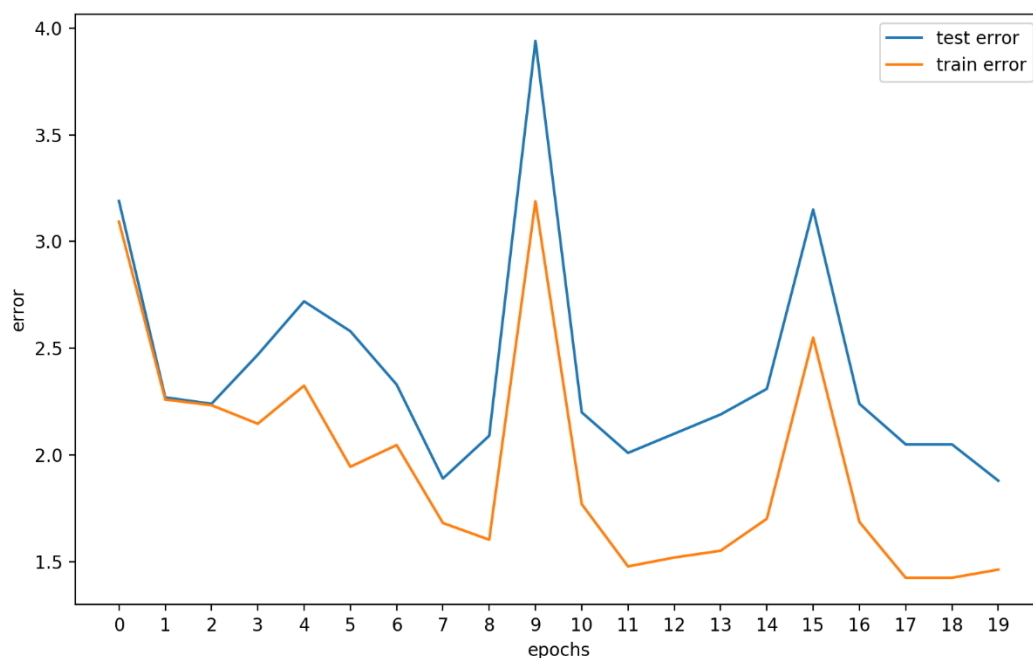
شکل 11 نمودار خطا با نرخ یادگیری یک صدم



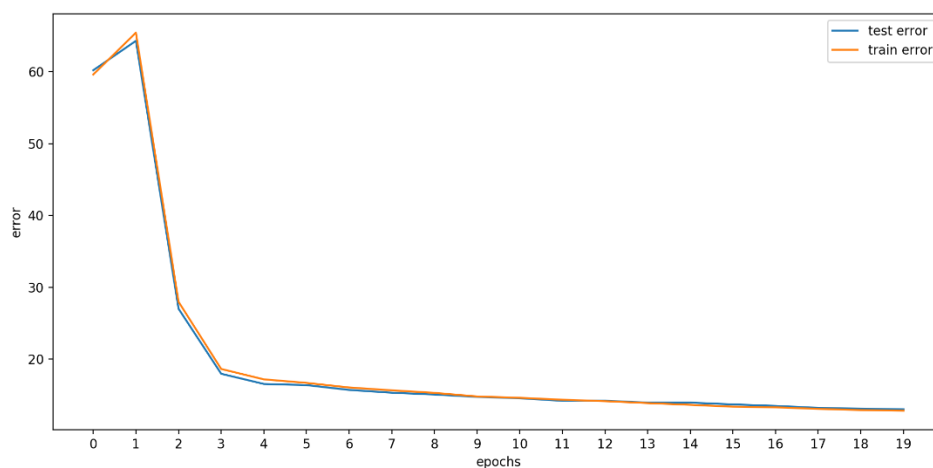
شکل 12 - نمودار خطا با نرخ یادگیری یک دهم

قسمت 1.8

نمودار خطا برای مدل با الگوریتم بهینه‌سازی adam و SGD به ترتیب زیر آمده است. برای الگوریتم adam درصد خطا به 1.3 رسیده است. اما برای الگوریتم SGD این مقدار برابر با 9.8 درصد است.



شکل 13 نمودار خطا برای مدل با الگوریتم بهینه‌سازی adam

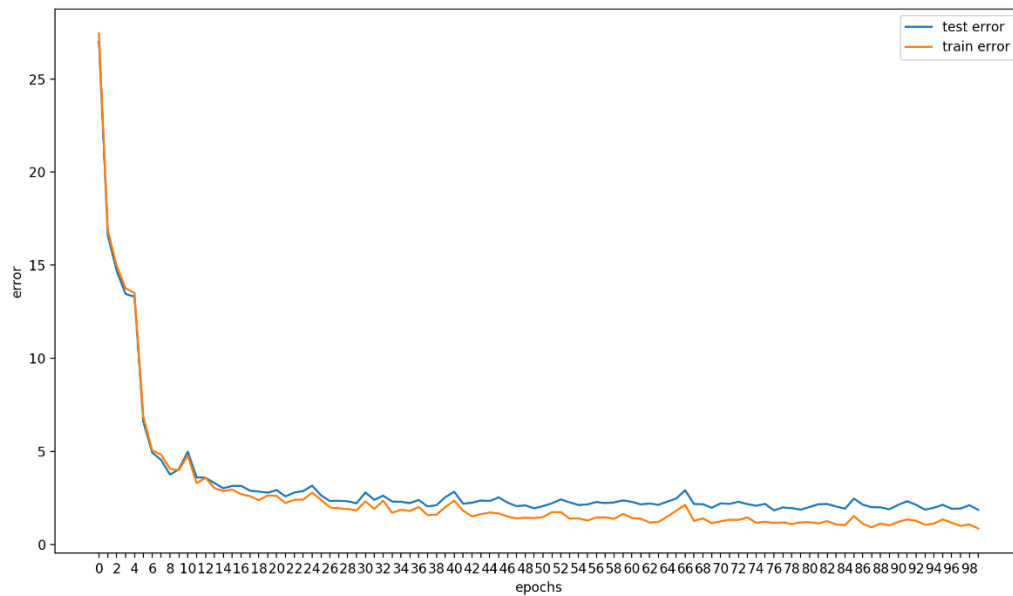


شکل 14 نمودار خطا برای مدل با الگوریتم بهینه‌سازی SGD

سوال 2

قسمت 2.1

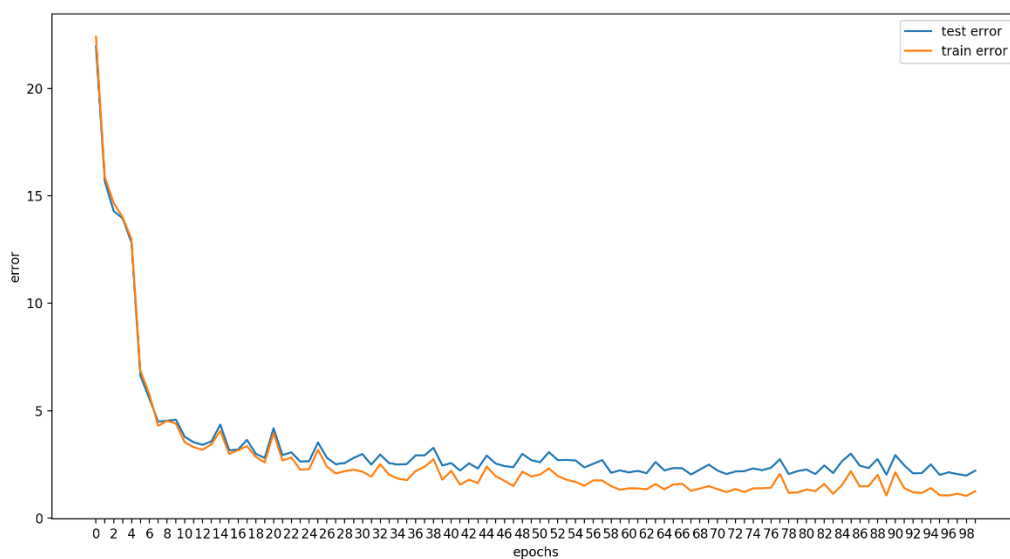
نمودار خطا برای در ادامه آمده است. دقت به دست آمده با این مدل برابر با 98 درصد شده است. همان طور که مشاهده می شود مدل در اپوک بیستم ام تقریبا به جواب مناسبی رسیده است.



شکل 15 نمودار خطا برای مدل با نرون *lstm* دو طرفه

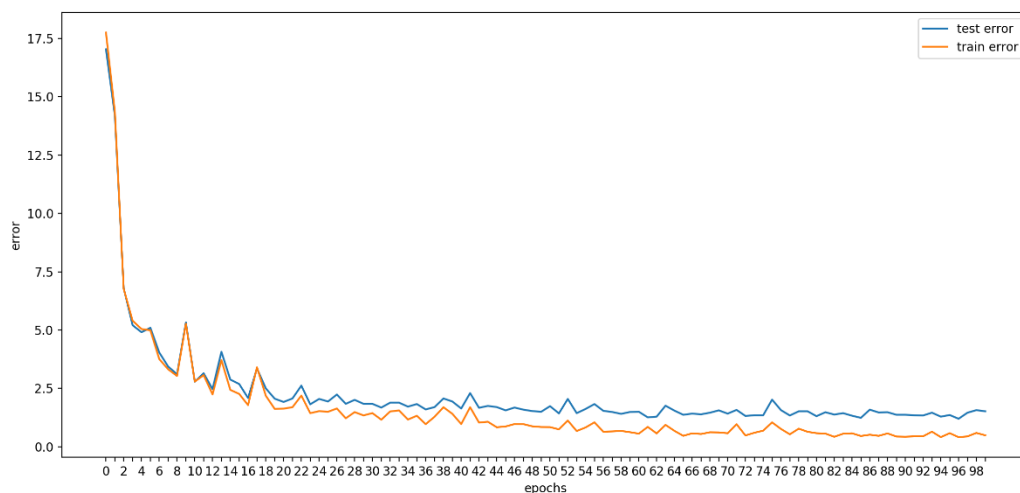
قسمت 2.2

نمودار خطا برای مدل با نرون lstm یک طرفه در شکل زیر آمده است. دقت مدل در این حالت برابر با 97 درصد بوده است. اما به نظر می‌رسد که تفاوتی بین دو مدل ندارد.



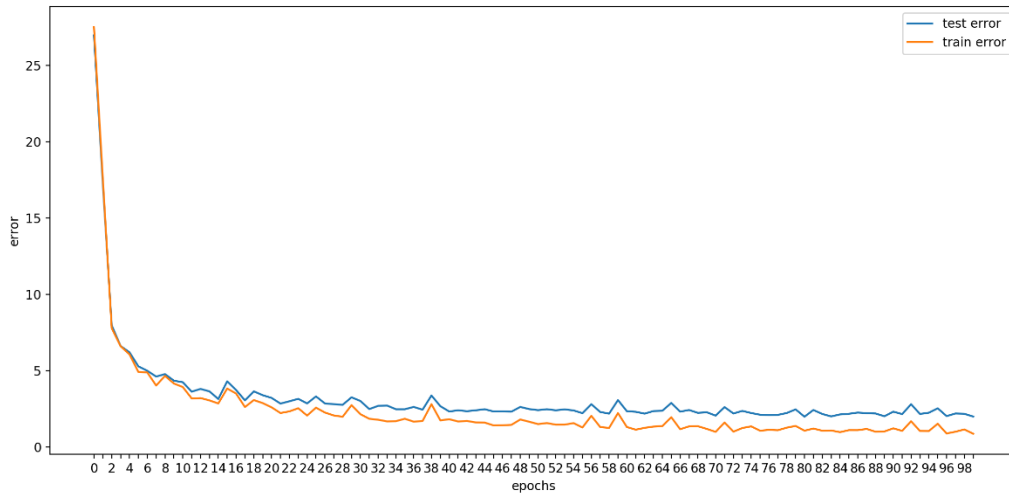
قسمت 2.3

نمودار خطا برای یک مدل با دو لایه نرون lstm در شکل زیر آمده است. دقت مدل به 99.2 رسیده است. که از مدل‌های قبلی بهتر است.

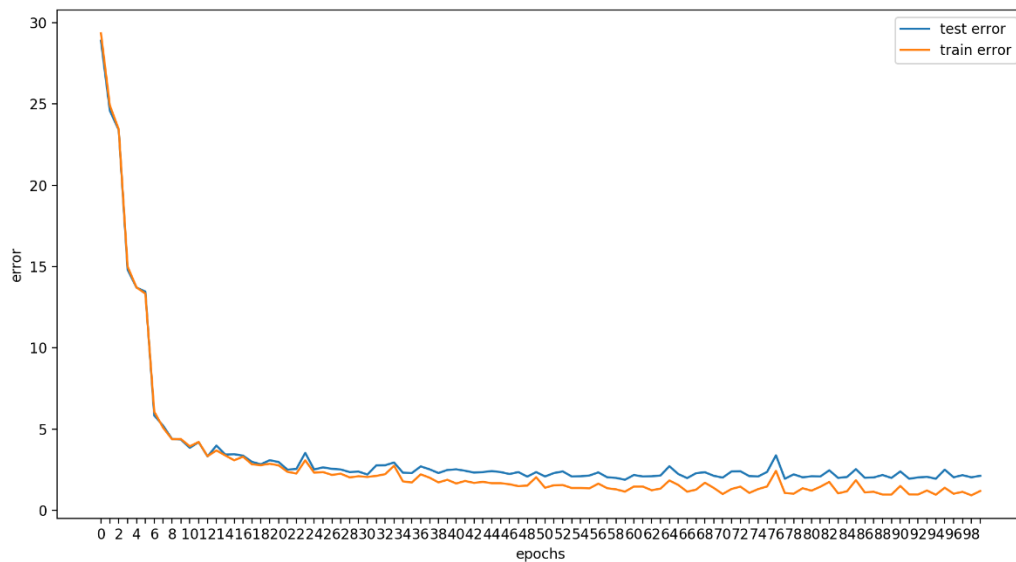


قسمت 2.4

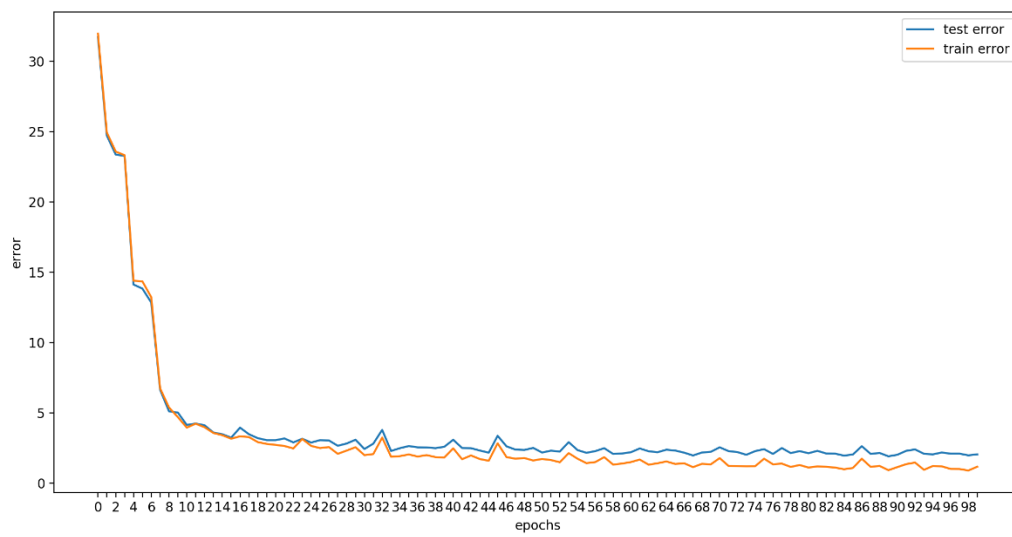
نمودار خطا برای مدل با عملگر dropout با مقادیر 10، 30 و 50 به ترتیب در ادامه آمده است. دقت مدل به ترتیب برابر با 98، 97 و 98 درصد شده است. تفاوت قابل ملاحظه‌ای بین مدل‌های مختلف مشاهده نمی‌شود.



شکل 16 نمودار خطای مدل با عملگر dropout با مقدار 0.1



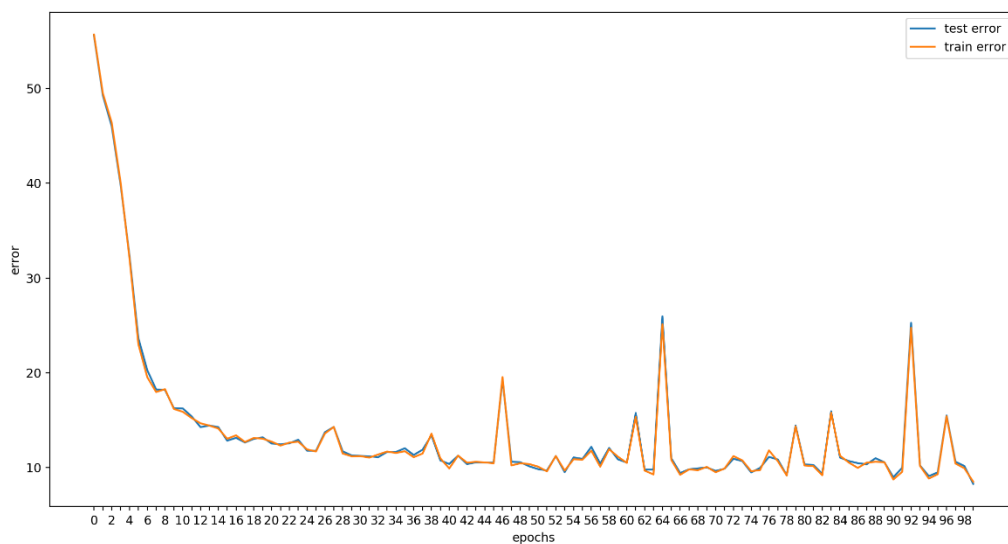
شکل 17 نمودار خطای مدل با عملگر dropout با مقدار 0.3



شکل 18 نمودار خطای مدل با عملگر dropout با مقدار 0.5

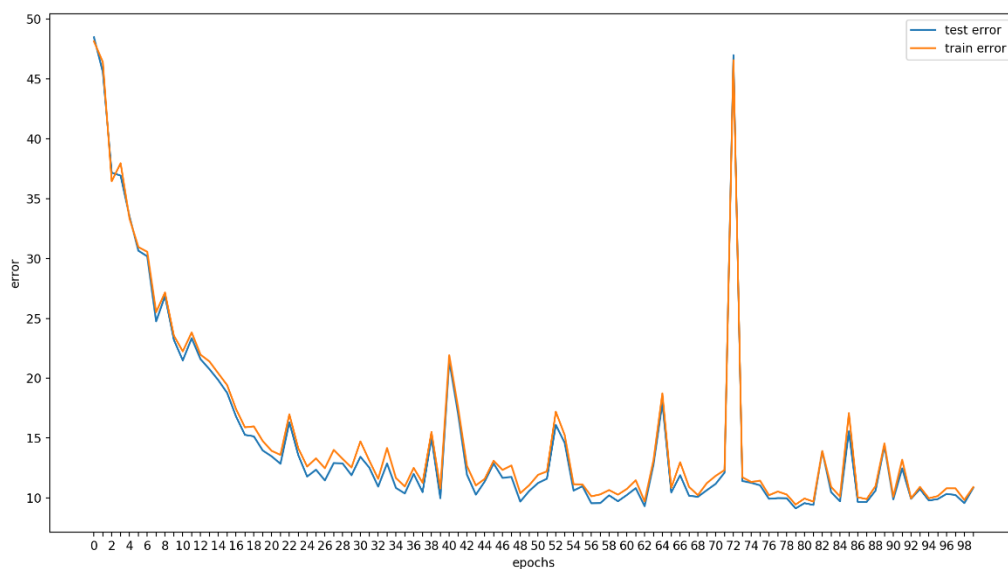
قسمت 2.5.1

نمودار خطا برای در ادامه آمده است. دقت به دست آمده با این مدل برابر با 91 درصد شده است. همان طور که مشاهده می شود مدل در اپوک شصتم تقریباً به جواب مناسبی رسیده است.



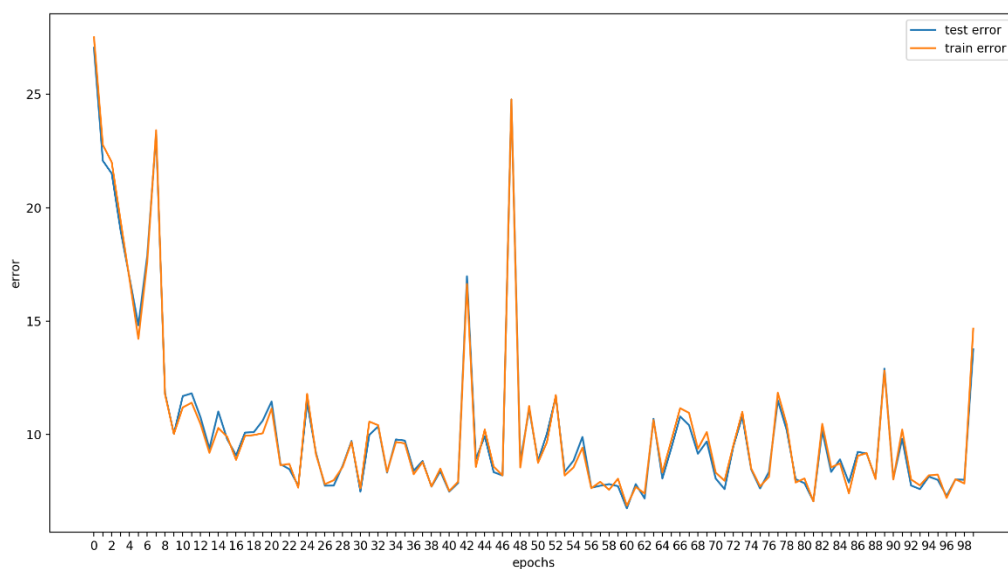
قسمت 2.5.2

نمودار خطا برای مدل یک طرفه در شکل زیر آمده است. دقت این مدل با مدل قبل فرق چندانی نکرده است. اما با حالتی که از lstm استفاده می کردیم بسیار تفاوت دارد و ضعیف تر شده است.



قسمت 2.5.3

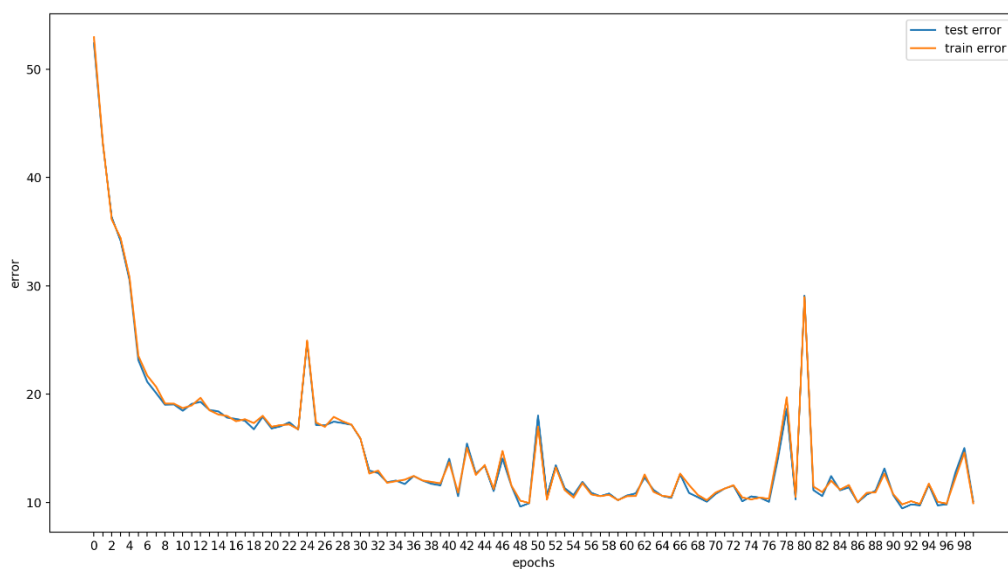
نمودار خطا برای مدل با دو لایه در شکل زیر نشان داده شده است. این مدل نسبت به مدل‌های قبلی نتایج بهتری داشته است و دارای دقت 86 درصد است. اما نسبت به حالت lstm ضعیف‌تر عمل کرده است.

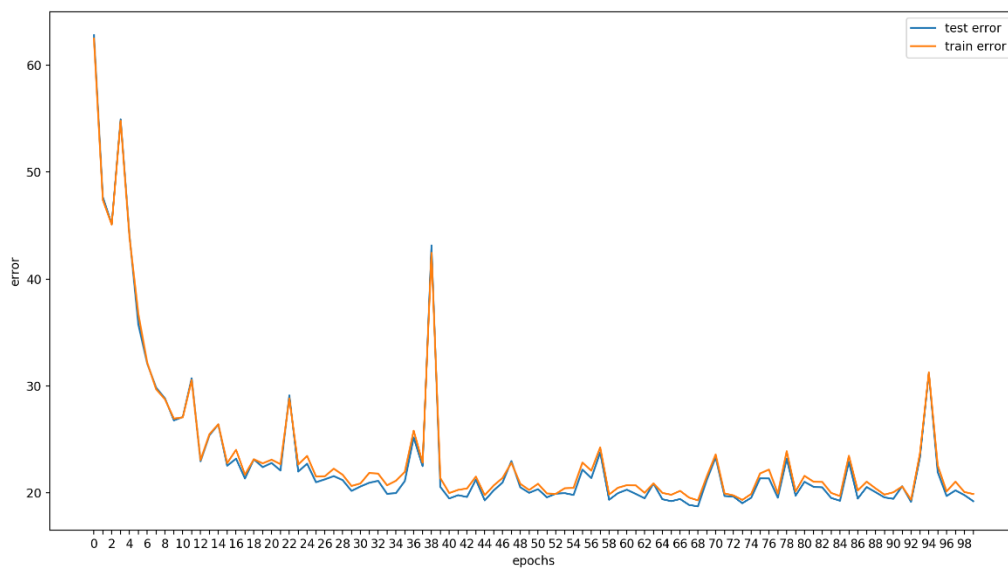
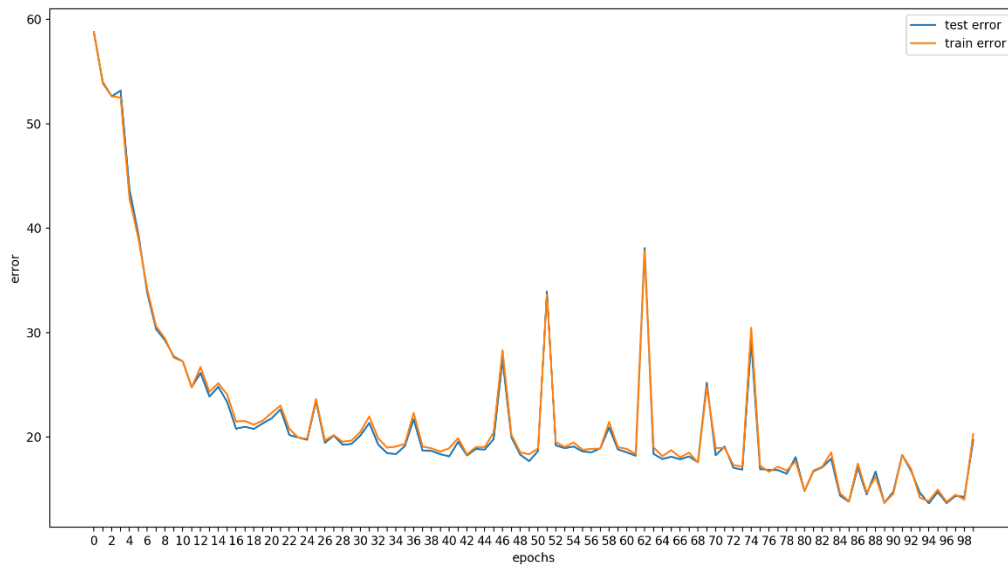


شکل 19 نمودار خطا برای مدل دو لایه

قسمت 2.5.4

نمودار خطای مدل با عملگر dropout با سه مقدار 10، 30 و 50 به ترتیب در ادامه آمده است. دقت مدل‌های به ترتیب برابر با 89 و 80 و 80 درصد است.





<http://pytorch.org/tutorials/>

<https://github.com/pytorch/examples/tree/master/mnist>