



دانشگاه صنعتی امیرکبیر  
(پلی تکنیک تهران)

دانشکده مهندسی کامپیوتر و فناوری اطلاعات

گروه هوش مصنوعی

# گزارش پروژه درس شبکه‌های عصبی-۳

مینا ترقی

۹۳۱۳۱۰۷۲

تاریخ تحویل: ۱۳۹۴/۸/۱۹

نیم‌سال اول ۹۵-۱۳۹۴

# سوال ۱

طراحی یک شبکه پرسپترونی چند لایه برای تخمین تابع.

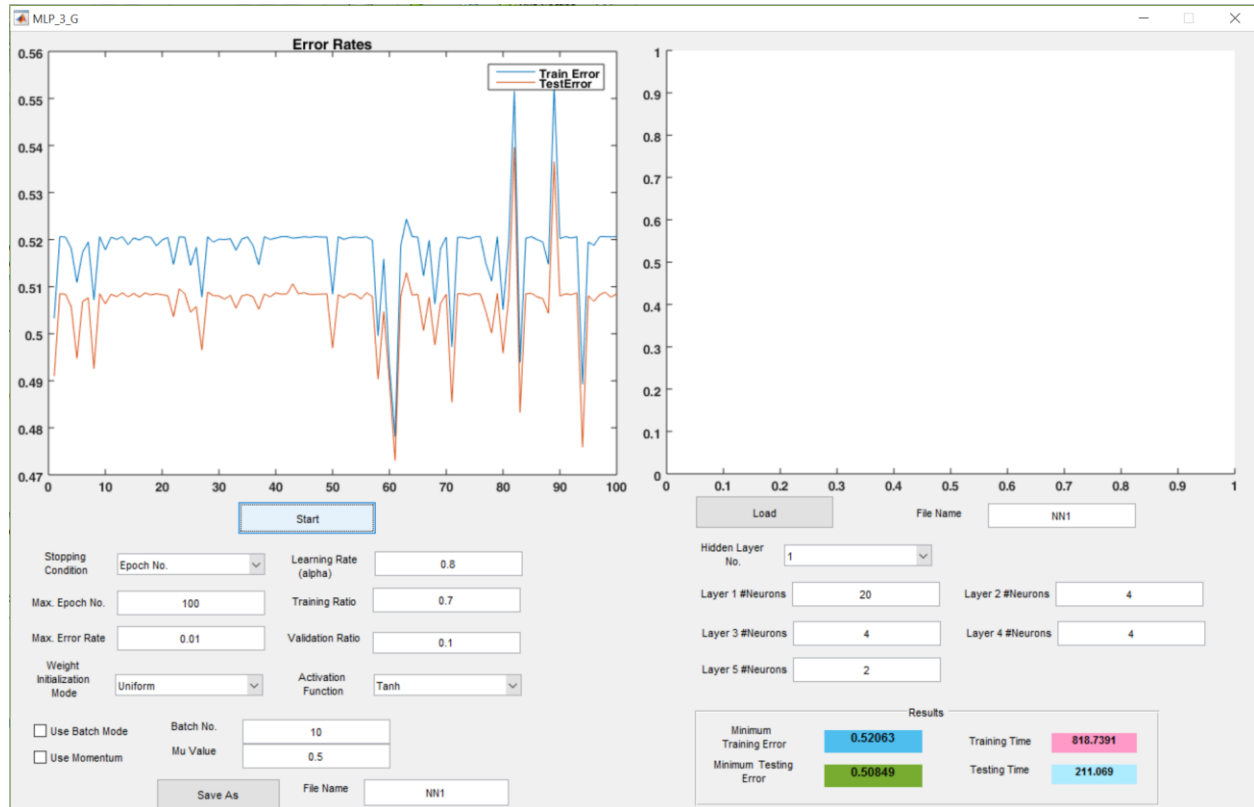
شرایط آزمایش:

پارامتر	مقدار
تعداد اپیک	۱۰۰
نرخ یادگیری	۰٫۸
درصد داده آموزشی	٪۷۰
درصد داده ارزیابی	٪۱۰
درصد داده آزمون	٪۲۰
تابع فعالیت	تانژانت هایپربولیک
مقادیر اولیه وزن ها	$U \sim [-۰٫۵, ۰٫۵]$
معماری شبکه	(۳۸۴، ۲۰، ۴، ۱)

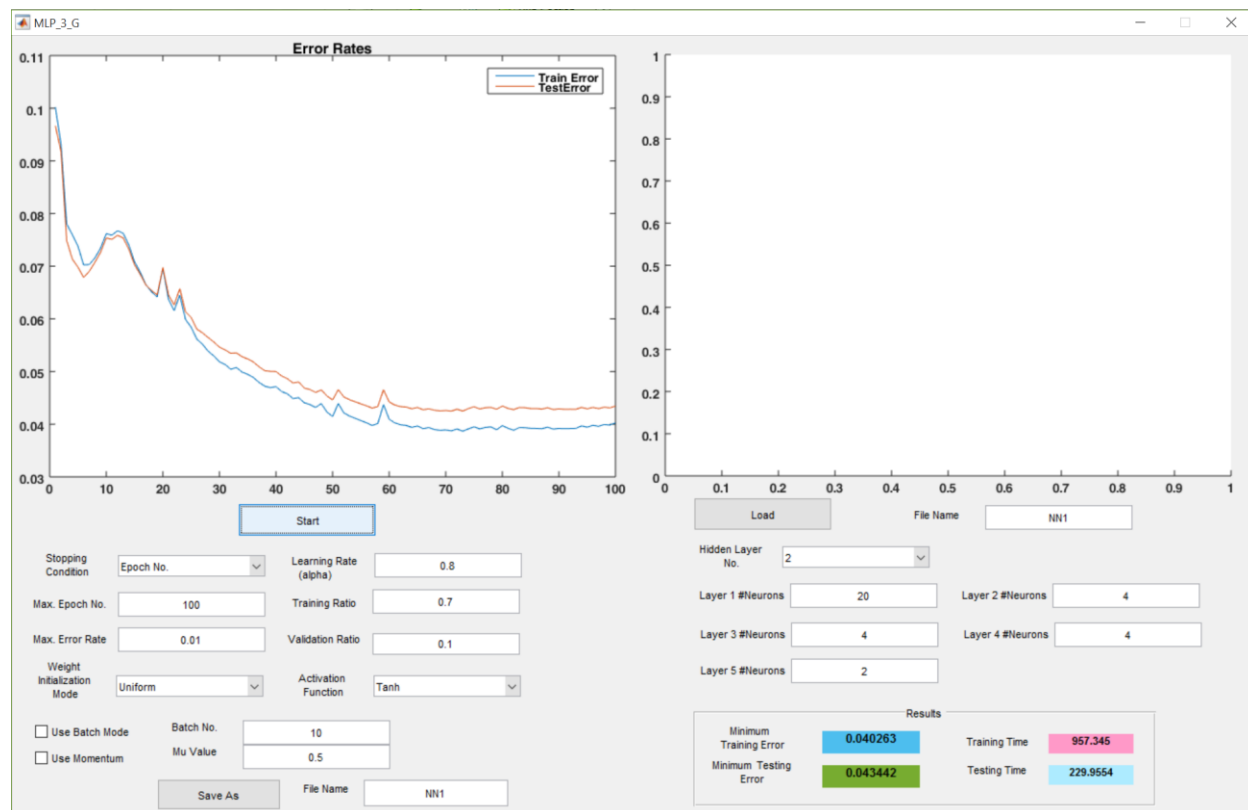
- ✓ برای این آزمایش و تمامی آزمایش‌های دیگر به بررسی داده‌ها پرداختیم. متوجه شدیم که ویژگی شماره ۱ بیماران، شماره شناسایی بیماران است که در واقع ویژگی محسوب نمی‌شود و در نتیجه ابتدا به صورت دستی ستون شماره یک تمامی داده‌ها را حذف کردیم و سپس تمامی آزمایش‌ها را بر روی داده جدید شامل ۳۸۴ ویژگی و ۱ خروجی انجام دادیم.
- ✓ مورد دیگر در مورد این مسئله بازه تغییرات خروجی مورد نظر بود که دارای تغییرات زیاد در بازه [۹۷/۴۸۹۱، ۱/۷۳۸۷] بود و برای استفاده از توابع فعال‌سازی که در تغییراتشان در بازه [۱-۱] است و در مقادیر بزرگ‌تر دامنه به ناحیه اشباع می‌روند، با استفاده از یک نگاشت خطی، خروجی‌های مورد نظر را به بازه [۰،۱] نگاشت کردیم.
- ✓ نکته دیگر این که با توجه به این که در زمان اجرای آزمایش‌های مختلف، شرایط آزمایش مانند شرایط فیزیکی و محیطی، اعم از دمای محیط، تعداد تردهای فعال سیستم، میزان حافظه در دسترس سیستم و ... یکسان نبود، نمی‌توانستیم مقایسه صحیحی بین زمان‌های ثبت‌شده انجام دهیم و در نتیجه در اکثر موارد از مقایسه و بررسی زمان‌های انجام آزمایش خودداری کردیم.
- ✓ برای این آزمایش با بررسی چندین معماری مختلف، شبکه با معماری (۳۸۴،۲۰،۴،۱) با دو لایه مخفی را حداقل شبکه مناسب یافته و نتایج آن را در بررسی موارد بررسی‌شده گزارش کردیم.

## نتیجه انجام آزمایش:

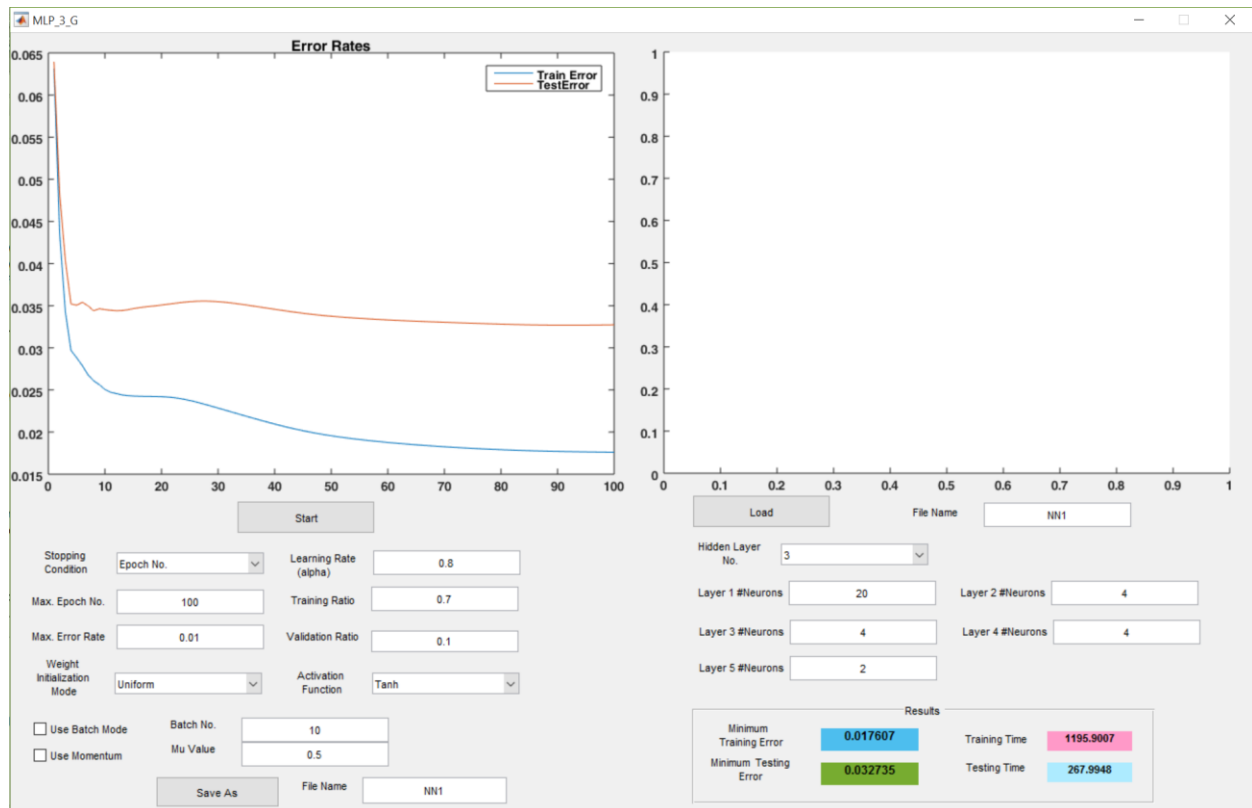
نتایج انجام آزمایش برای ۵ شبکه مختلف با معماری‌های مختلف به صورت زیر است.



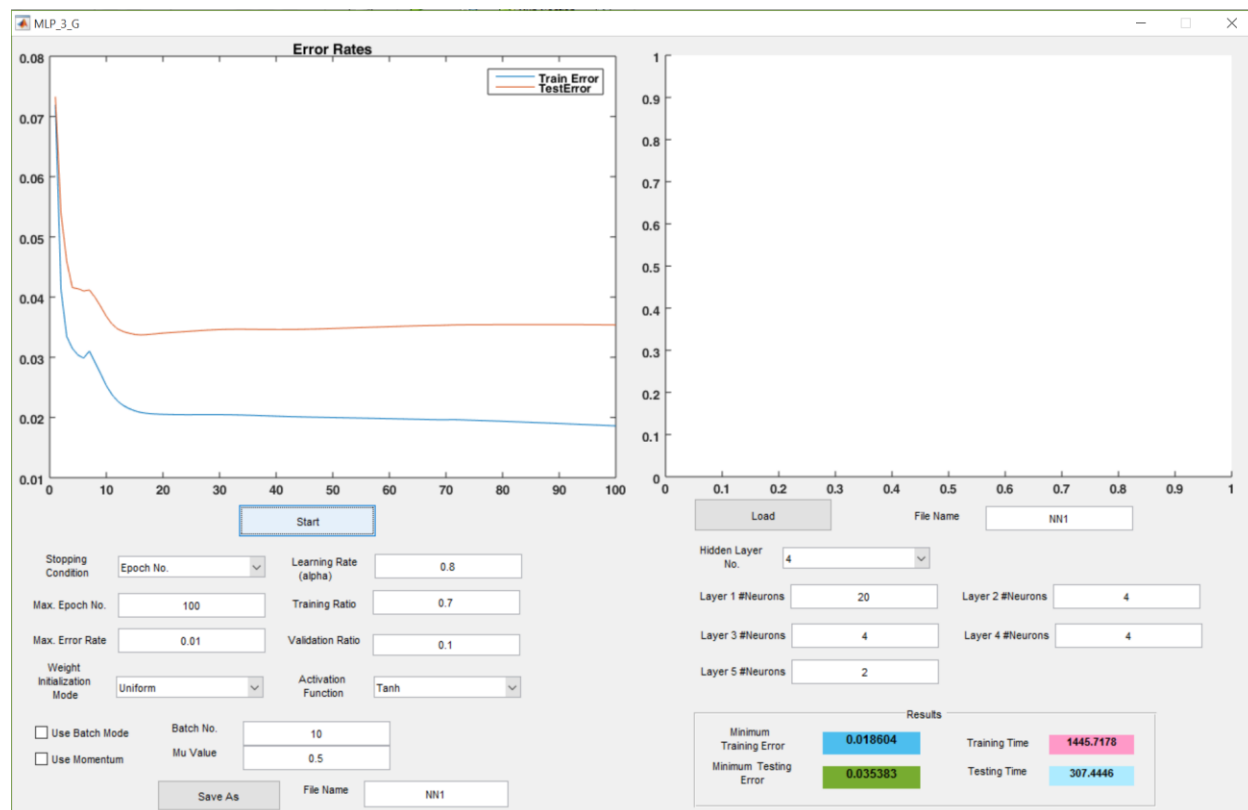
شکل ۱- نتایج شبکه با یک لایه مخفی - معماری (۳۸۴،۲۰،۱)



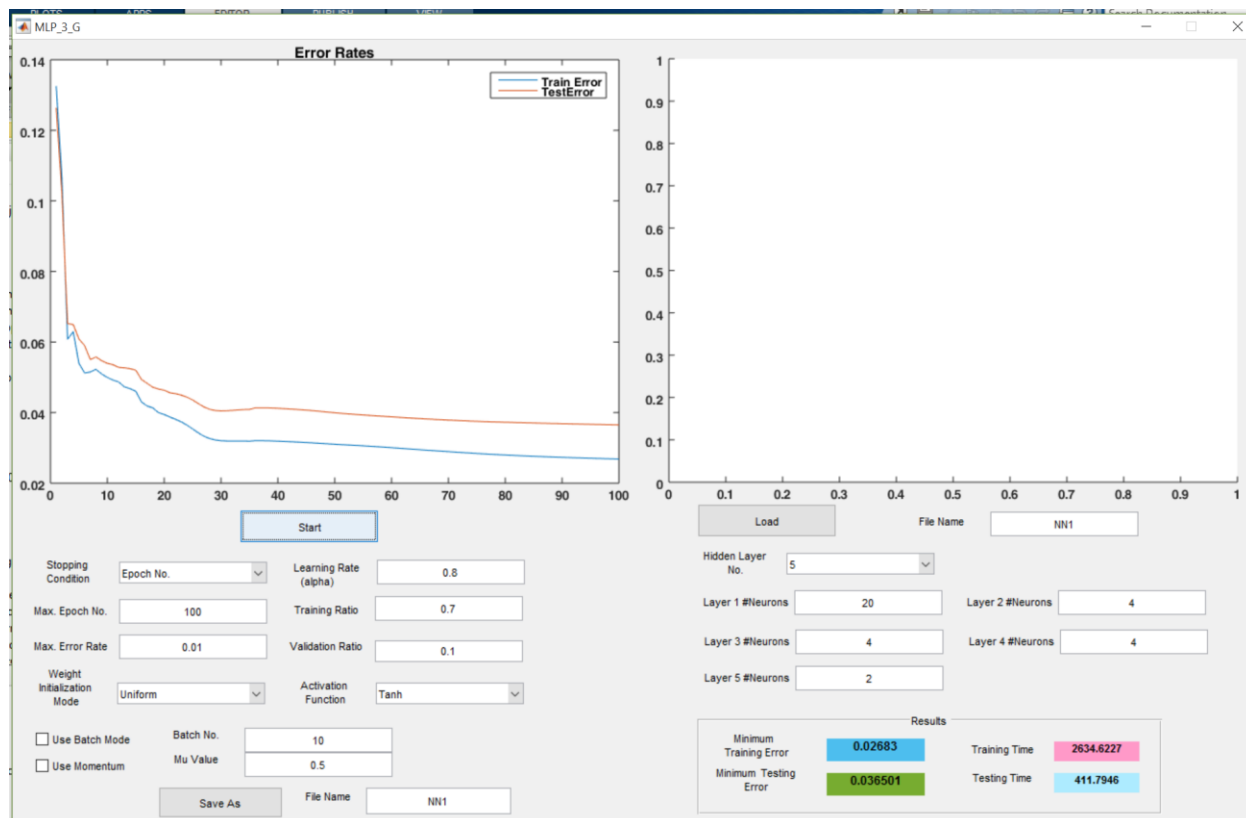
شکل ۲- نتایج شبکه با دو لایه مخفی - معماری (۳۸۴،۲۰،۴،۱)



شکل ۳- نتایج شبکه با سه لایه مخفی - معماری (۳۸۴،۲۰،۴،۴،۱)



شکل ۴ - نتایج شبکه با چهار لایه مخفی - معماری (۳۸۴،۲۰،۴،۴،۱)



شکل ۵- نتایج شبکه با ۵ لایه مخفی - معماری (۳۱۴،۲۰،۴،۴،۴،۲،۱)

خطای نهایی آموزش	خطای نهایی آزمون	معماری شبکه
۰،۵۲۶۰۳	۰،۵۰۸۴۹	۳ لایه
۰،۰۴۰۲۶۳	۰،۰۴۳۴۴۲	۴ لایه
۰،۰۱۷۶۰۷	۰،۰۳۲۷۳۵	۵ لایه
۰،۰۱۸۶۰۴	۰،۰۳۵۳۸۳	۶ لایه
۰،۰۲۶۸۳	۰،۰۳۶۵۰۱	۷ لایه



## نتیجه گیری:

با افزایش لایه های مخفی، در شرایط نسبتاً برابر، دقت تخمین شبکه بالاتر می رود. البته شیب این نمودار تغییر دقت به مرور زمان کم تر می شود، بدین معنا که افزودن لایه دومخ به لایه اول دقت را بسیار بهبود می بخد، در حالی که افزودن لایه چهارم به سه لایه، تاثیر به سزایی در افزایش دقت شبکه ندارد.

## سوال ۲

بررسی اثر توابع فعال سازی مختلف.

### شرایط آزمایش:

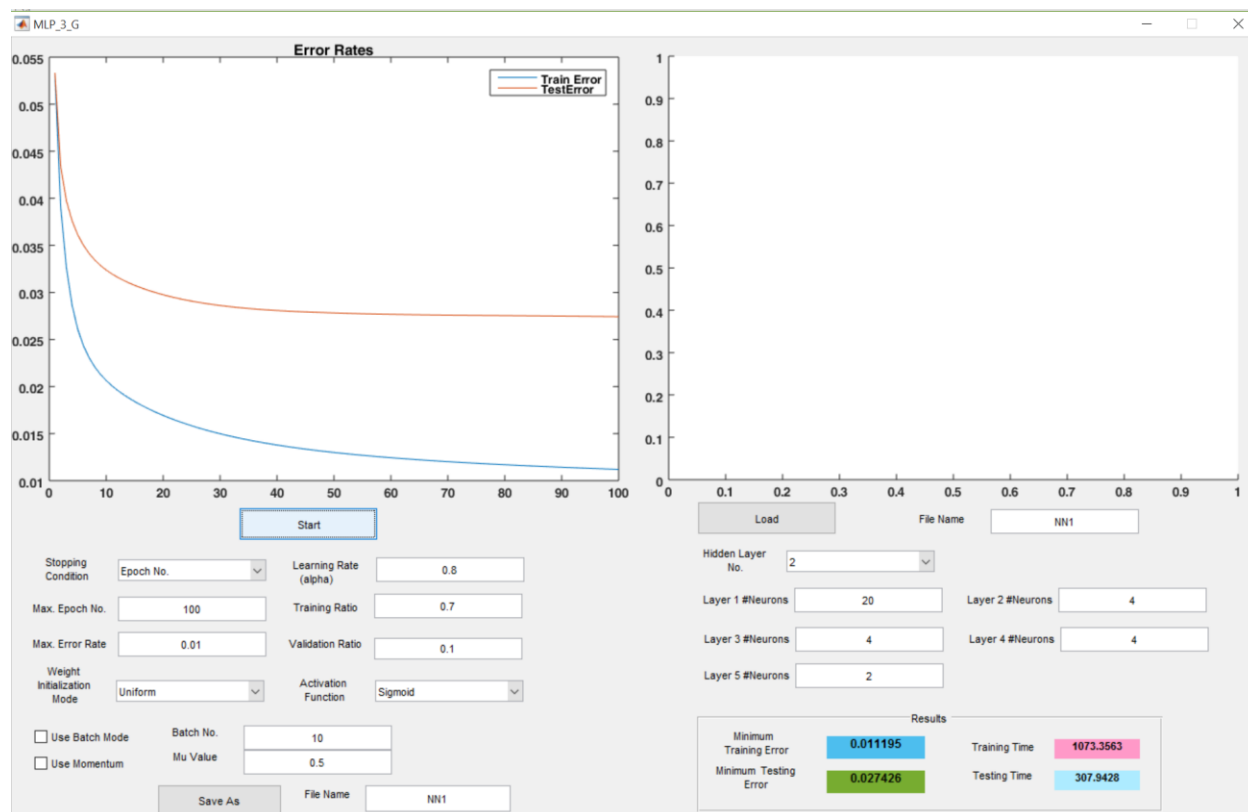
پارامتر	مقدار
تعداد ایپک	۱۰۰
نرخ یادگیری	۰٫۸
درصد داده آموزشی	٪۷۰
درصد داده ارزیابی	٪۱۰
درصد داده آزمون	٪۲۰

سیگموئید	تابع فعالیت
(۳۸۴،۲۰،۴،۱)	معماری شبکه

برای این آزمایش تابع فعال سازی را به سیگموئید تغییر دادیم و معماری گزیده سوال ۱ را برای شرایط بالا اجرا نمودیم.

## نتیجه انجام آزمایش:

پس از اجرای الگوریتم، نتایج برای تابع فعال سازی سیگموئید به صورت زیر بود که آن را با تابع فعال سازی تانژانت هایپربولیک و سایر شرایط یکسان مقایسه می کنیم و مشاهده می کنیم که به نظر می رسد تابع سیگموئید نتایج بهتری برای مسئله ما به دست می دهد.



شکل ۶- نتایج برای شبکه ۳ لایه با تابع فعال سازی سیگموئید

## نتیجه گیری:

می توانیم نتیجه بگیریم که تابع فعال سازی سیگموئید برای مسئله ما و شبکه طراحی شده با بهتر عمل می کند؛ در نتیجه، از این به بعد، باقی آزمایشات این پروژه را با تابع سیگموئید هم انجام می دهیم.

## سوال ۳

بررسی تاثیر مقادیر اولیه وزن ها و روش پیشنهادی ویدرو.

## شرایط آزمایش:

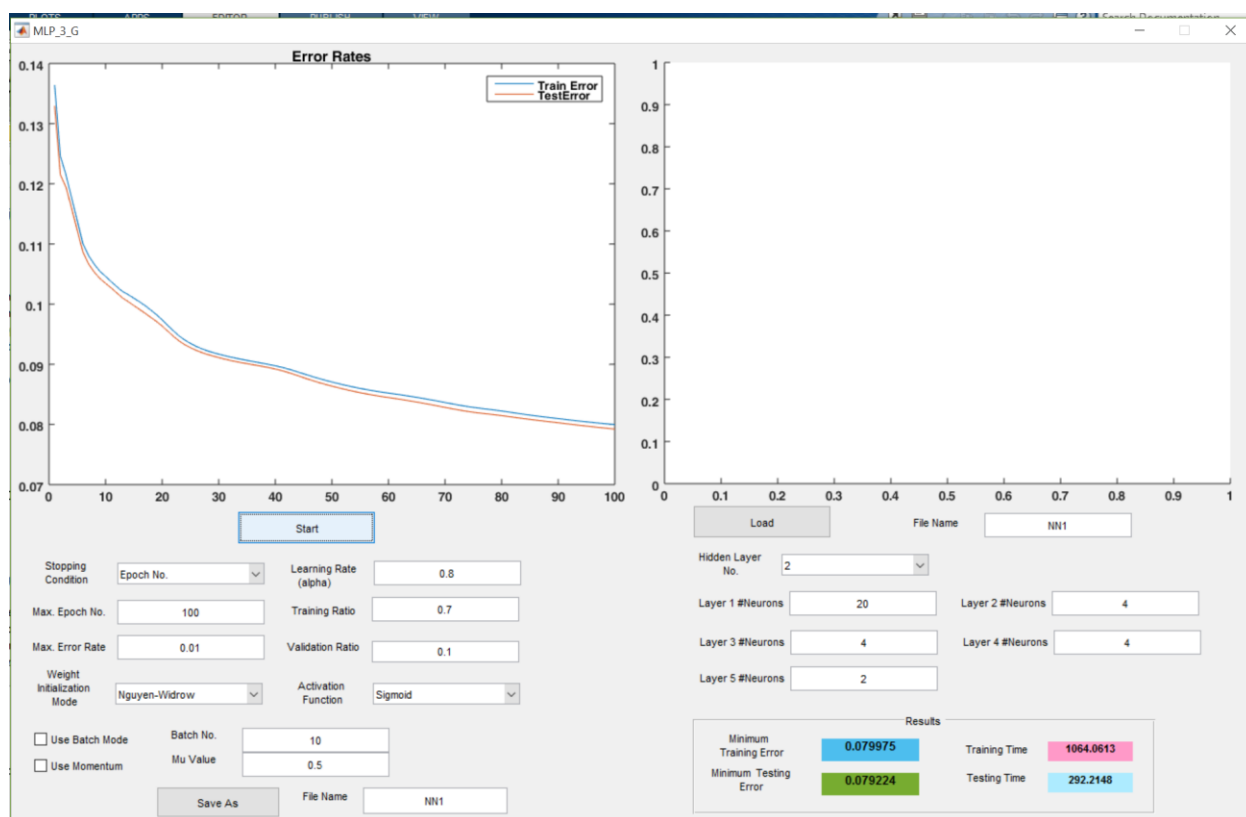
مقدار	پارامتر
۱۰۰	تعداد اپیک
۰٫۸	نرخ یادگیری
٪۷۰	درصد داده آموزشی
٪۱۰	درصد داده ارزیابی
۲۰٪	درصد داده آزمون
سیگموئید	تابع فعالیت

معماری شبکه	(۳۸۴،۲۰،۴،۱)
-------------	--------------

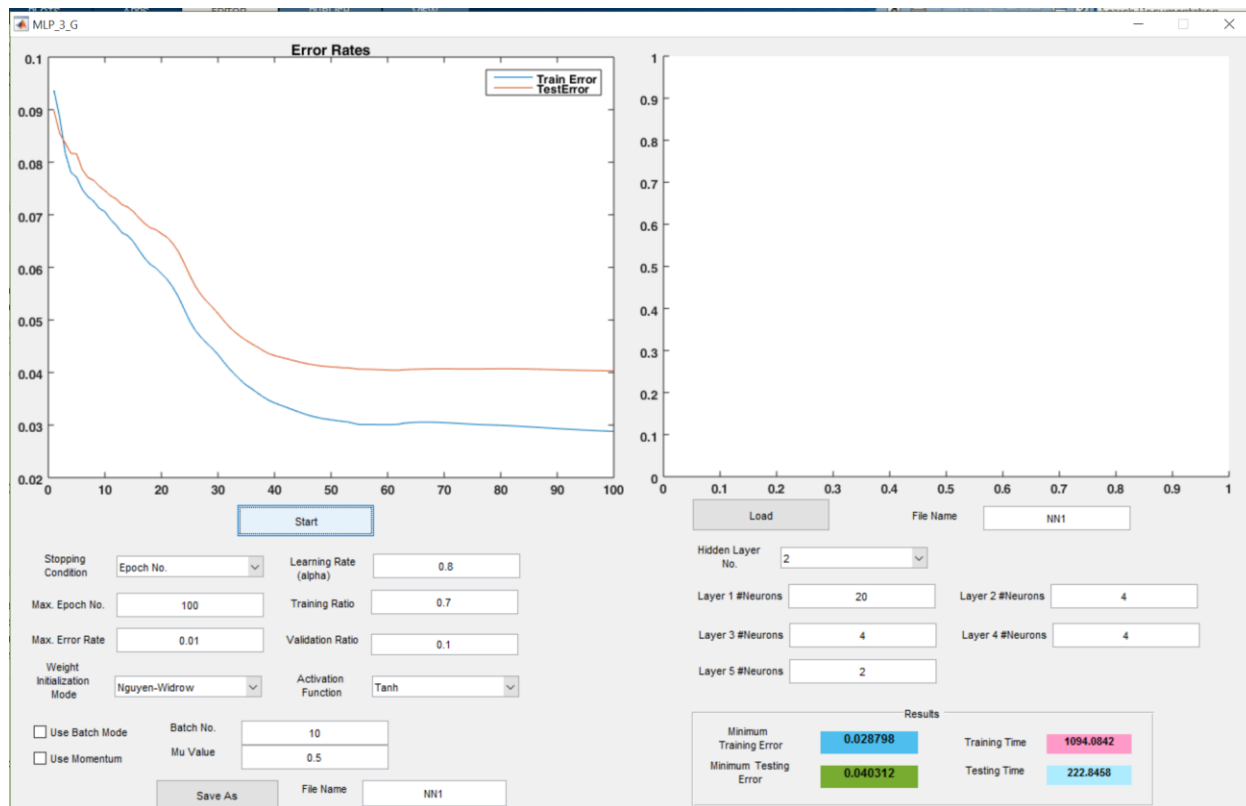
در این آزمایش با استفاده از تابعی که در فایل InitNW.m نوشتیم، وزن‌های اولیه مقادیر را بنا بر روش Nguyen-Widrow تعیین کردیم.

## نتیجه انجام آزمایش:

همان‌طور که در شکل زیر پیداست، به نظر می‌رسد که با مقداردهی اولیه وزن‌های شبکه به روش Nguyen-Widrow شبکه و استفاده از تابع تانژانت‌هایپربولیک شبکه به نتایج بهتری می‌رسد، در حالی که با شرایط مشابه، شبکه با تابع فعال‌سازی سیگموئیدی دارای افت عملکرد می‌شود.



شکل ۷- نتایج شبکه ۳ لایه با تابع فعال‌سازی سیگموئید و روش مقداردهی اولیه Nguyen-Widrow



شکل ۸- نتایج شبکه ۳ لایه باتابع فعال سازی  $Tanh$  و روش مقداردهی اولیه Nguyen-Widrow

## نتیجه گیری:

به نظر می‌رسد که برای این مسئله، استفاده از تابع تانژانت هایپربولیک به همراه روش مقداردهی اولیه Nguyen-Widrow بهتر از سایر گزینه‌هاست.

## سوال ۴

بررسی الگوریتم پس انتشار خطا با پیاده سازی دسته ای و غیردسته ای.

## شرایط آزمایش:

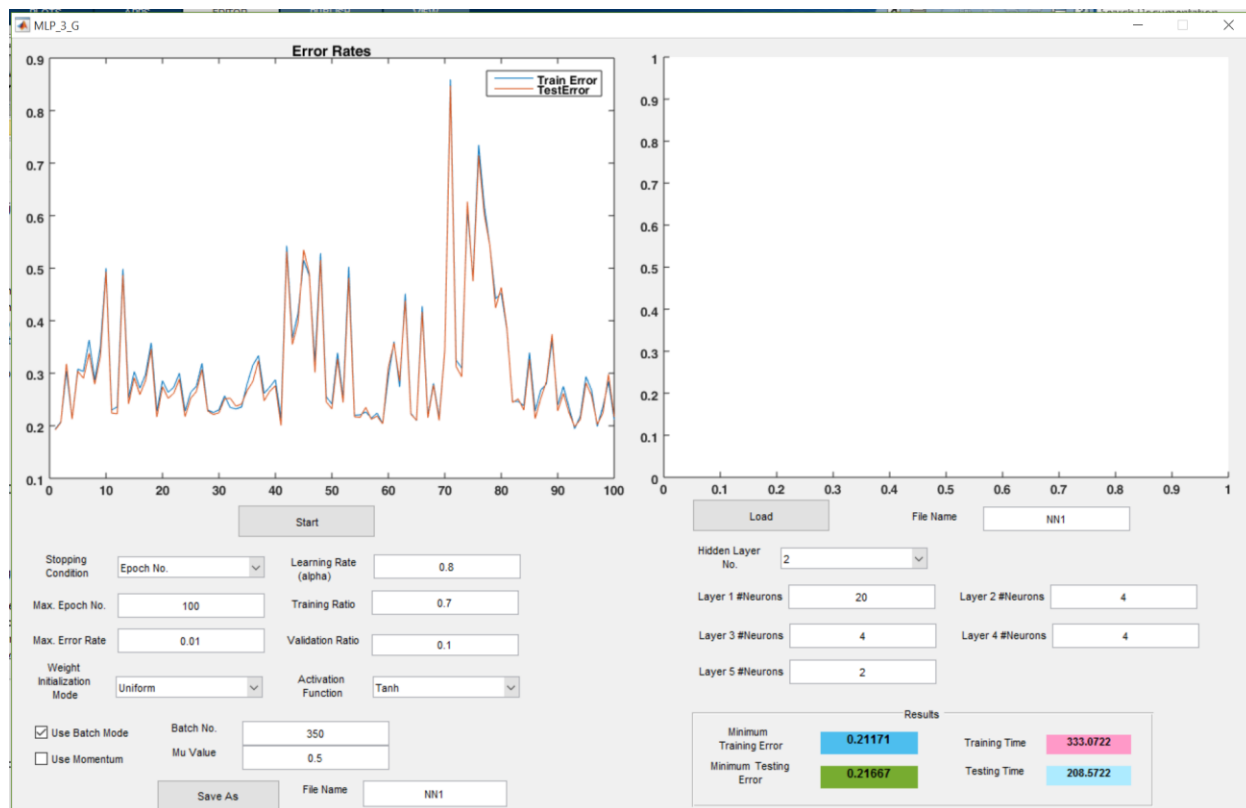
مقدار	پارامتر
-------	---------

تعداد ایپک	۱۰۰
نرخ یادگیری	۰٫۸
درصد داده آموزشی	٪۷۰
درصد داده ارزیابی	٪۱۰
درصد داده آزمون	٪۲۰
تابع فعالیت	تانژانت هایپربولیک
معماری شبکه	(۳۸۴،۲۰،۴،۱)
اندازه دسته برای بهروزرسانی وزن‌ها	هر ایپک، ۵۰ نمونه، ۳۵۰ نمونه

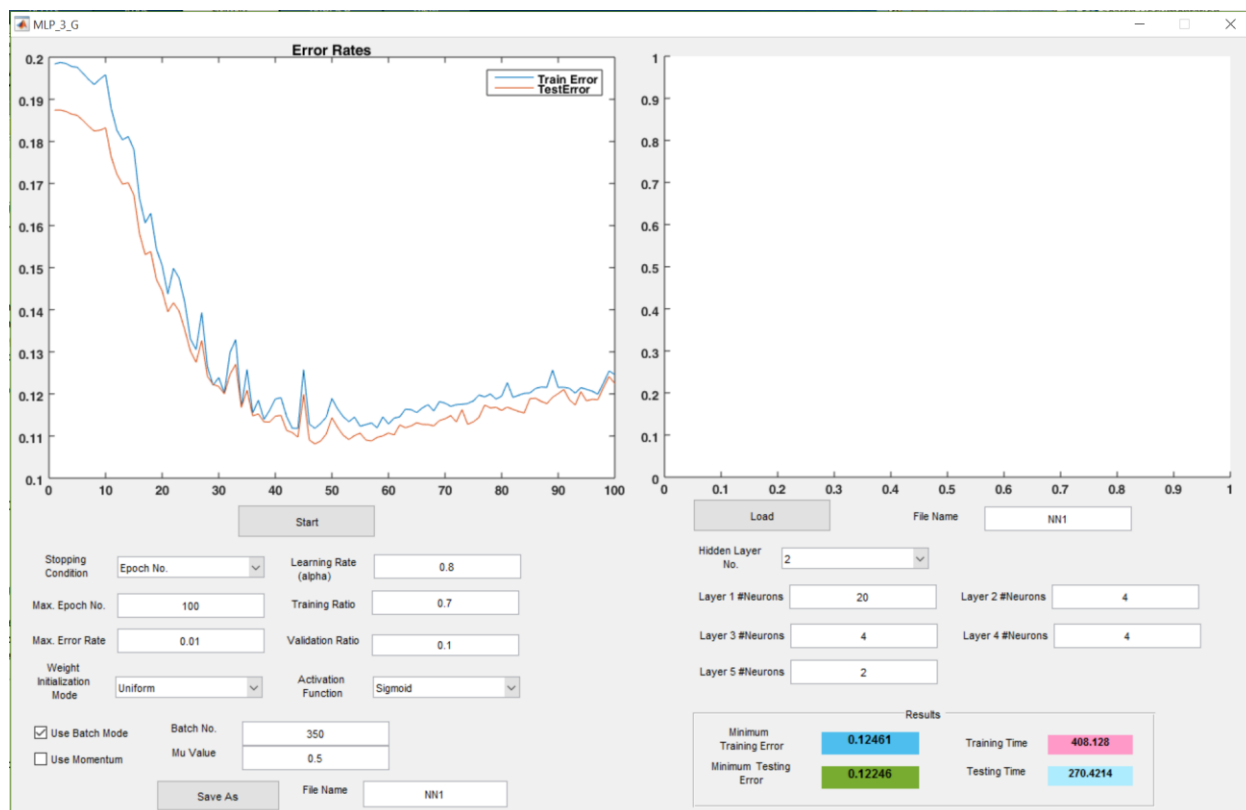
در این آزمایش، وزن‌ها را با استفاده از مقدار خطای دسته‌ای بهروزرسانی کردیم.

## نتیجه انجام آزمایش:

همان‌طور که مشاهده می‌شود، در بهروزرسانی دسته‌ای وزن‌ها



شکل ۹- نتایج شبکه ۳ لایه باتابع فعال سازی  $Tanh$  و روش به روزرسانی دسته ای وزن ها (هر ۳۵۰ نمونه)

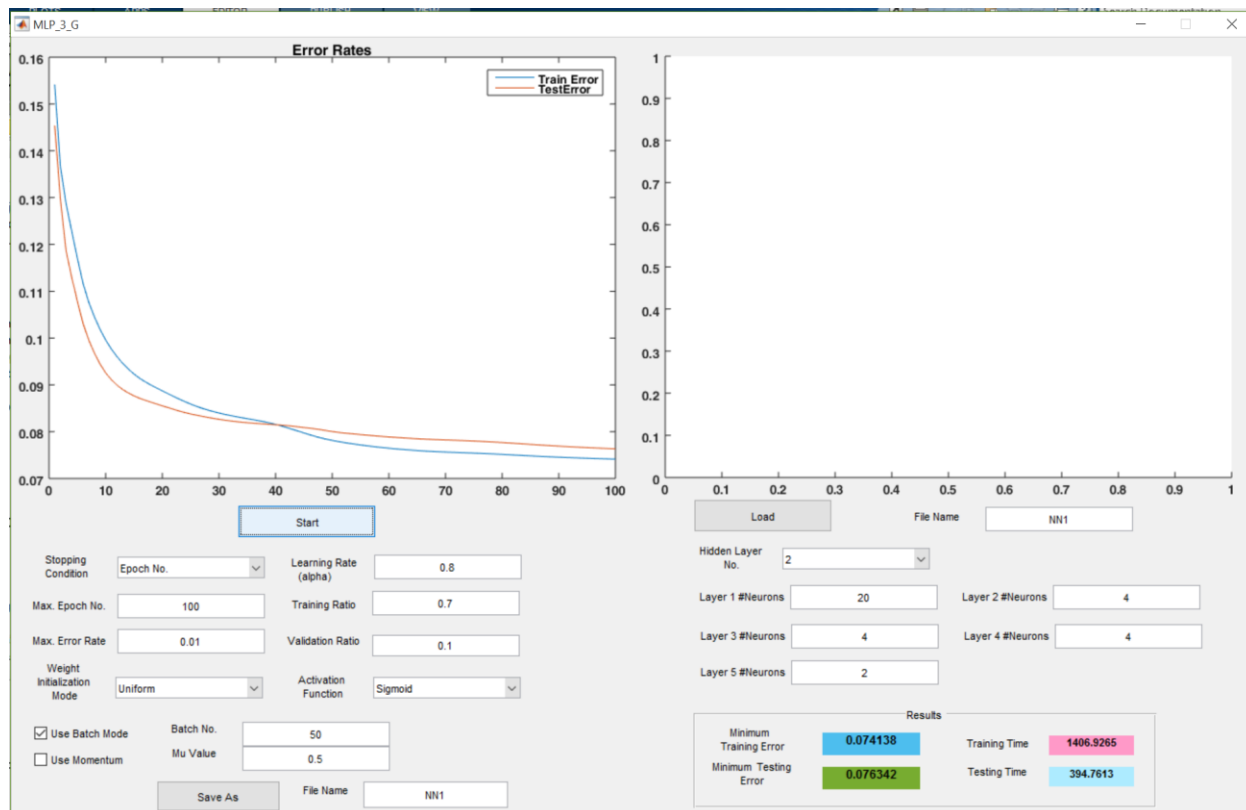


شکل ۱۰ - نتایج شبکه ۳ لایه باتابع فعال سازی سیگموئید و روش به روزرسانی دسته ای وزن ها (هر ۳۵۰ نمونه)

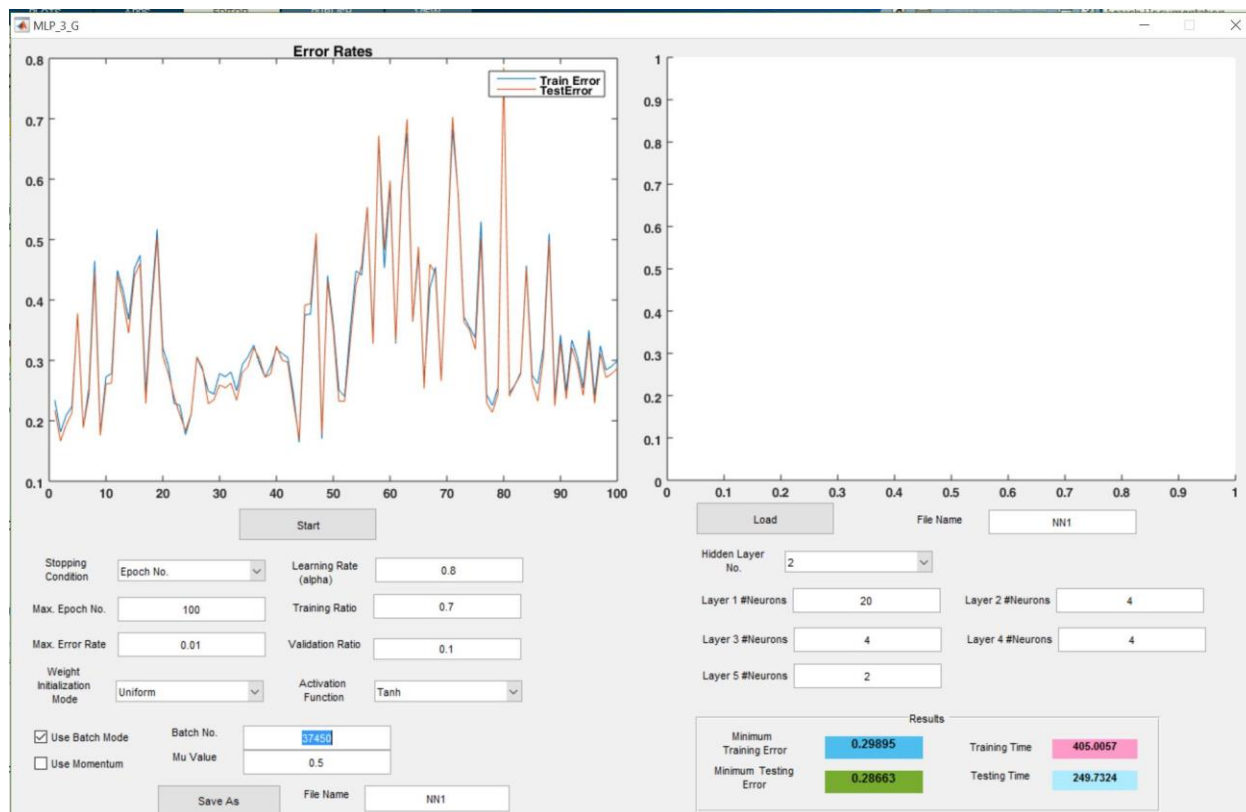




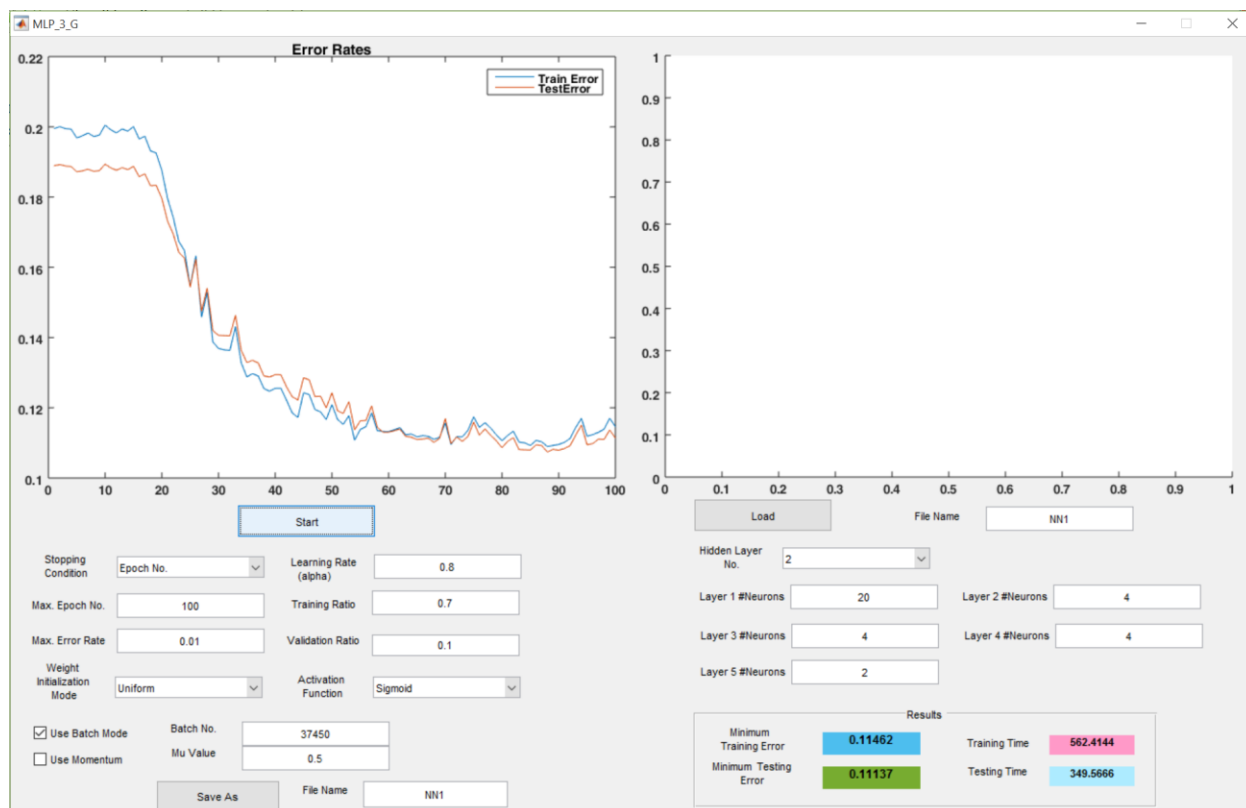
شکل ۱۱- نتایج شبکه ۳ لایه باتابع فعال سازی  $Tanh$  و روش به روزرسانی دسته ای وزن ها (هر ۵۰ نمونه)



شکل ۱۲- نتایج شبکه ۳ لایه باتابع فعال سازی سیگموئید و روش به روزرسانی دسته ای وزن ها (هر ۵۰ نمونه)



شکل ۱۳ - نتایج شبکه ۳ لایه باتابع فعال سازی  $Tanh$  و روش به روزرسانی دسته ای وزن ها (هر اپیک)



شکل ۱۴ - نتایج شبکه ۳ لایه باتابع فعال سازی سیگموئید و روش به روزرسانی دسته ای وزن ها (هر اپیک)

## نتیجه گیری:

همان طور که انتظار داشتیم با به روز رسانی دسته ای وزن ها، در تعداد اپیک های مساوی با شرایط آزمایش قبل، سرعت آموزش شبکه افزایش چشم گیری یافت، اما با کاهش دقت روبه رو هستیم.

در مورد به کارگیری هم زمان روش به روزرسانی دسته ای وزن ها و تابع فعال سازی تانژانت هایپربولیک، شبکه دارای نوسانات شدید در خروجی و افت محسوس دقت تخمین می شود.

مشاهده می کنیم که با کاهش اندازه دسته ها به ۵۰ نمونه، تابع سیگموئید هم نتایج بهتری به نمایش می گذارد و هم منحنی خطای هموارتری را گزارش می کند.

در مورد به روز رسانی در هر ایپک با تابع فعال سازی سیگموئید، تفاوت چندانی با اندازه دسته ۳۵۰ تایی مشاهده نمی شود که نشان می دهد یا نباید اندازه دسته ها را از حدی (مثلا ۵۰ برای این مسئله) بیش تر گرفت، یا اگر از حدی گذشت، بهتر است اندازه را بالا برده و در هر ایپک به روز رسانی را انجام دهیم تا بار محاسباتی کاهش یابد.

## سوال ۵

بررسی الگوریتم پس انتشار خطا با استفاده از مومنتم و بدون استفاده از مومنتم.

### شرایط آزمایش:

پارامتر	مقدار
تعداد ایپک	۱۰۰
نرخ یادگیری	۰٫۸
درصد داده آموزشی	٪۷۰
درصد داده ارزیابی	٪۱۰
درصد داده آزمون	٪۲۰
تابع فعالیت	سیگموئید
معماری شبکه	(۳۸۴،۲۰،۴،۱)

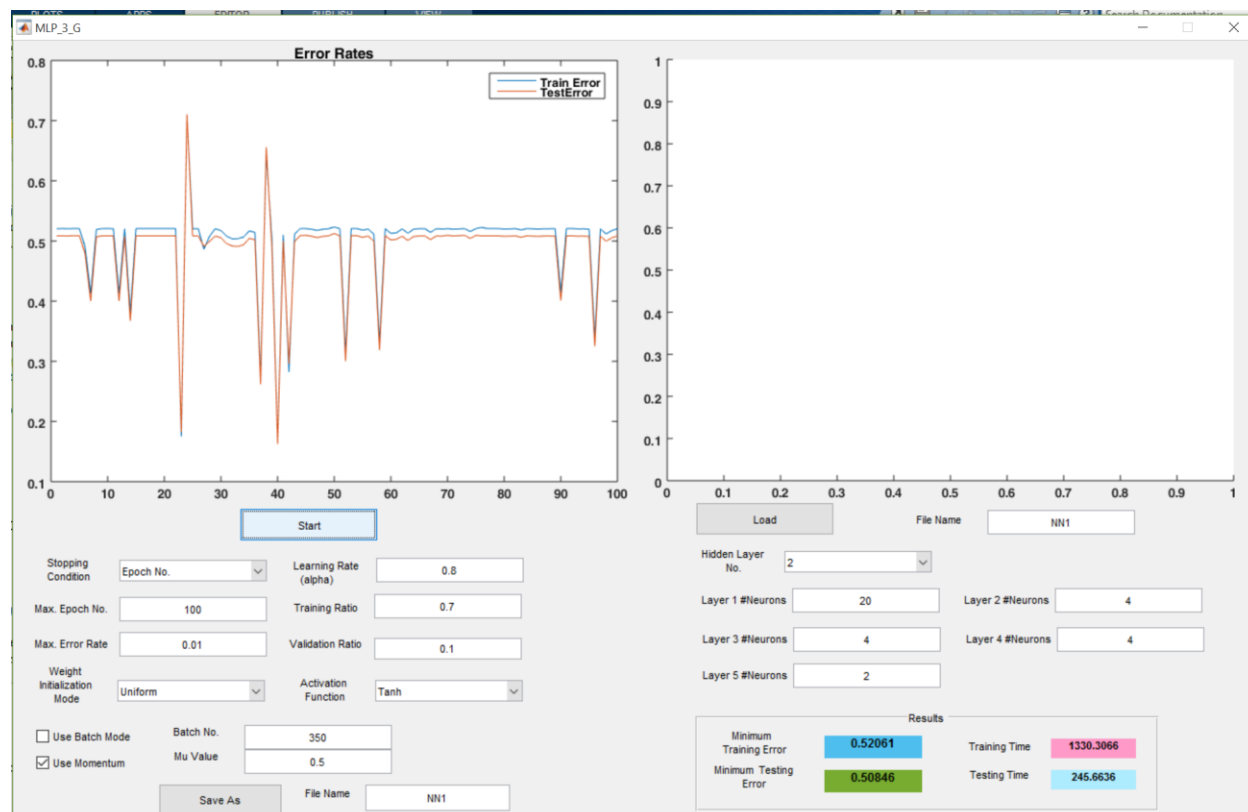
مقدار $\mu$	۰,۵
-------------	-----

در این آزمایش، برای تطبیق بهتر به روزرسانی وزن‌ها، وزن‌ها را با استفاده از مومنتم به روزرسانی کردیم.

## نتیجه انجام آزمایش:

همان‌گونه که می‌بینیم، دقت الگوریتم در سمت جواب بهینه کاهش یافته است. دوباره، در مورد تابع فعال‌سازی تانژانت‌هایپربولیک با افت عملکرد شدید روبرو هستیم؛ تا حدی که الگوریتم در حد تخمین تصادفی عمل میکند و دارای دقت حدود ۵۰٪ است. اما در مورد شبکه با تابع فعال‌سازی سیگموئید، شبکه پس از رسیدن به مینیمم خطای خود، دوباره سیر صعودی در پیش گرفته و دارای خطای بیشتر می‌شود!

مینیمم خطای شبکه با تابع فعال‌سازی سیگموئید و استفاده از مومنتم در به روزرسانی وزن‌ها، بسیار بیشتر از نمونه مشابه بدون استفاده از مومنتم در به روزرسانی وزن‌هاست.



شکل ۱۵- نتایج شبکه ۳ لایه با تابع فعال‌سازی Tanh و استفاده از مومنتم برای به روزرسانی وزن‌ها



شکل ۱۶- نتایج شبکه ۳ لایه باتابع فعال سازی سیگموئید و استفاده از مومنتم برای به روزرسانی وزن ها

## نتیجه گیری:

استفاده از مومنتم هم زمان با استفاده از تابع فعال سازی تانژانت هایپربولیک، عملکرد شبکه را با اختلال شدید مواجه می کند.

استفاده از مومنتم هم زمان با استفاده از تابع فعال سازی سیگموئید، باعث افت دقت تخمین شبکه می شود.