



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

دانشکده مهندسی کامپیوتر و فناوری اطلاعات

گروه هوش مصنوعی

گزارش پروژه درس شبکه‌های عصبی-۱

مینا ترقی

۹۳۱۳۱۰۷۲

تاریخ تحویل: ۱۳۹۴/۷/۲۹

نیم‌سال اول ۹۵-۱۳۹۴

سوال ۱

طراحی سه پرسپترون برای دسته بندی داده های Iris با چهار ویژگی و سه کلاس.

شرایط آزمایش:

| پارامتر | مقدار |
|------------------|----------|
| تعداد ایپک | ۱۰۰۰ |
| نرخ یادگیری | ۰,۵ |
| درصد داده آموزشی | ٪۷۰ |
| درصد داده آزمون | ٪۳۰ |
| تابع فعالیت | سیگموئید |

برای این آزمایش برای این که داده های با کلاس های مختلف را مخلوط کنیم و مخلوط برای تمام دفعات اجرای آزمایش ثابت باشد، ابتدا با استفاده از متلب، ترکیب داده ها را به صورت تصادفی به هم زدیم (برنامه DataShuffler.m به همین منظور نوشته شده است).

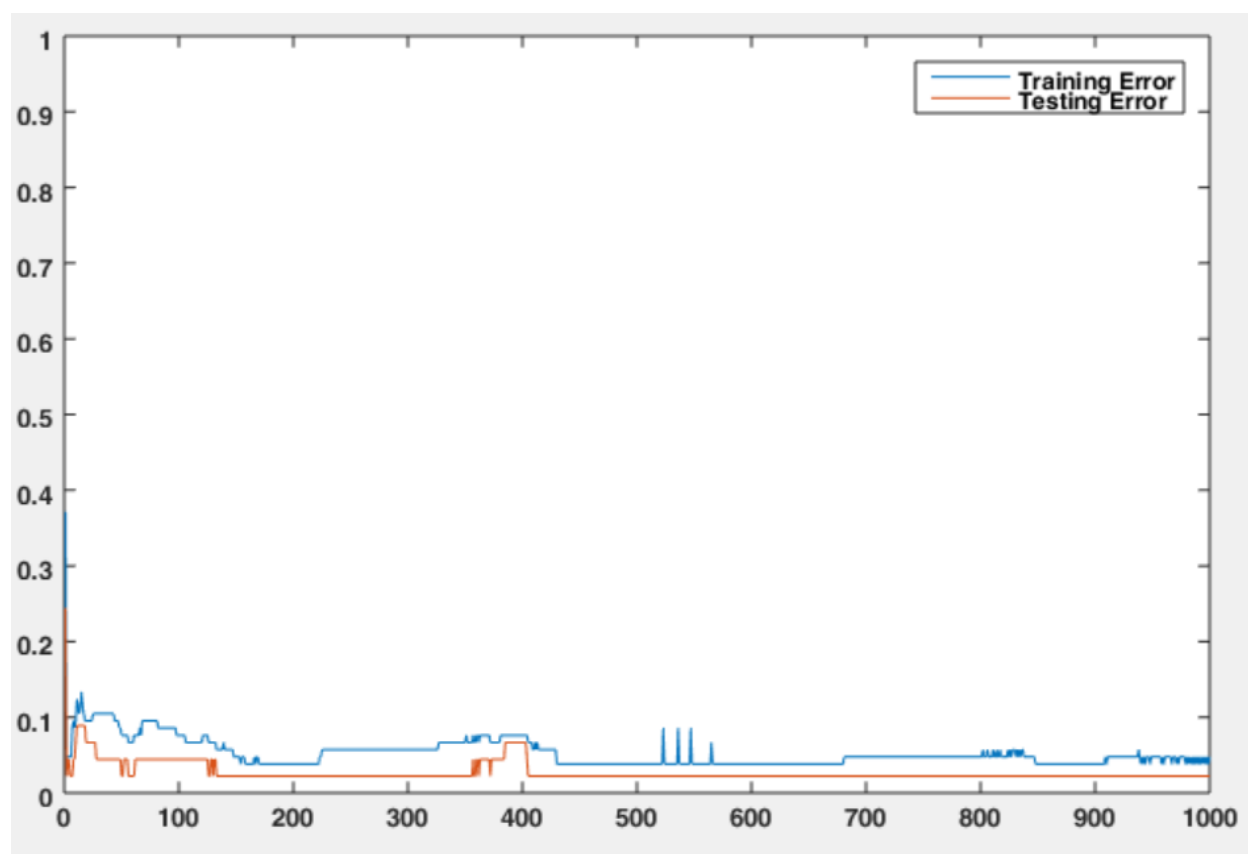
برای برنامه اصلی از فایل P_1_1.m استفاده کردیم.

نتیجه انجام آزمایش:

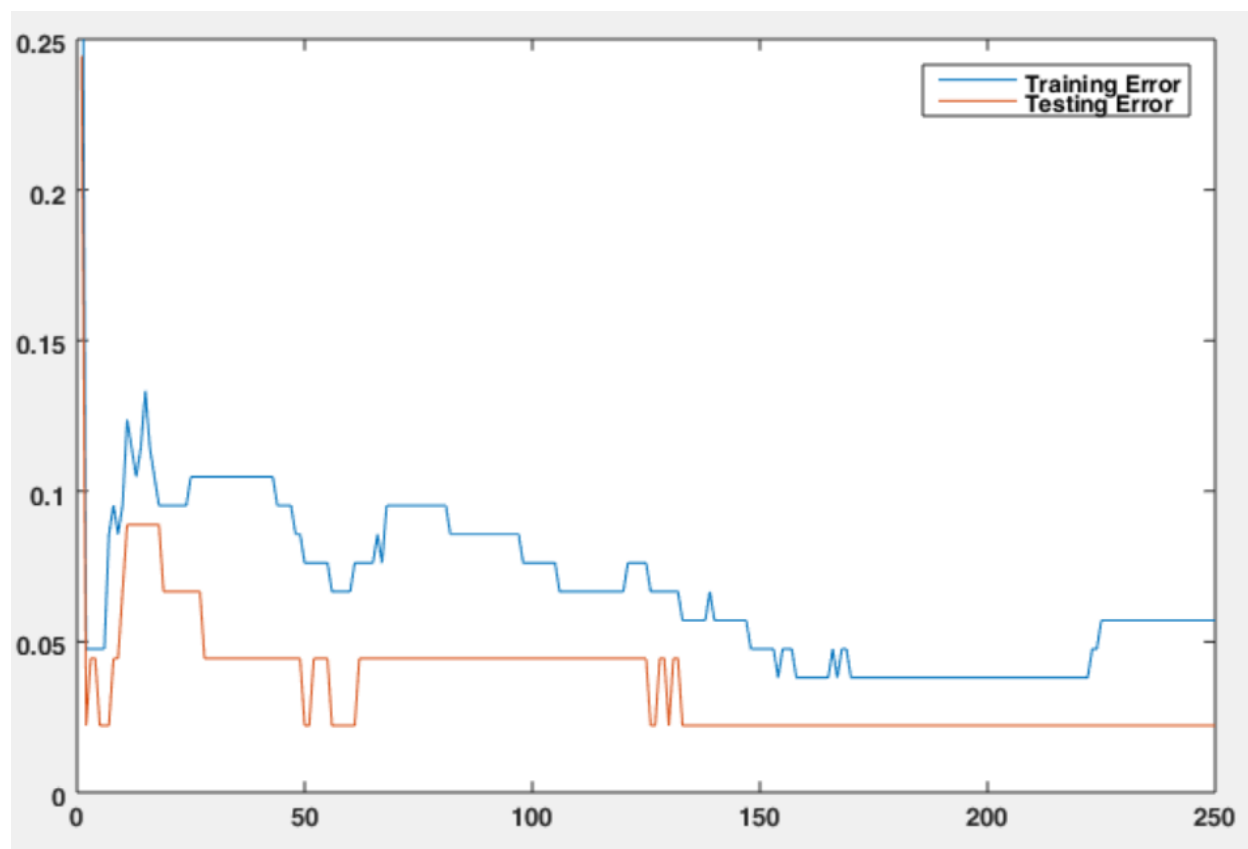
ابتدا سه پرسپترون را برای هر کلاس داده آموزش داده و سپس چهار نتیجه خطا را گزارش نمودیم. سه نمودار خطای اولیه برای دسته‌بندی کلاس‌های یک تا سه از دو کلاس دیگر و یک نمودار خطای دسته‌بندی کلی (سه کلاس).

برای دسته‌بندی سه کلاس با استفاده از تکنیک یکی در برابر همه (one versus all) به این صورت عمل کردیم که با اعمال هر نمونه آزمون به هر سه پرسپترون آموزش دیده، داده را به کلاسی اختصاص دادیم که مقدار خروجی پرسپترون برای آن نمونه از دو پرسپترون دیگر بزرگ‌تر (مثبت‌تر) بود.

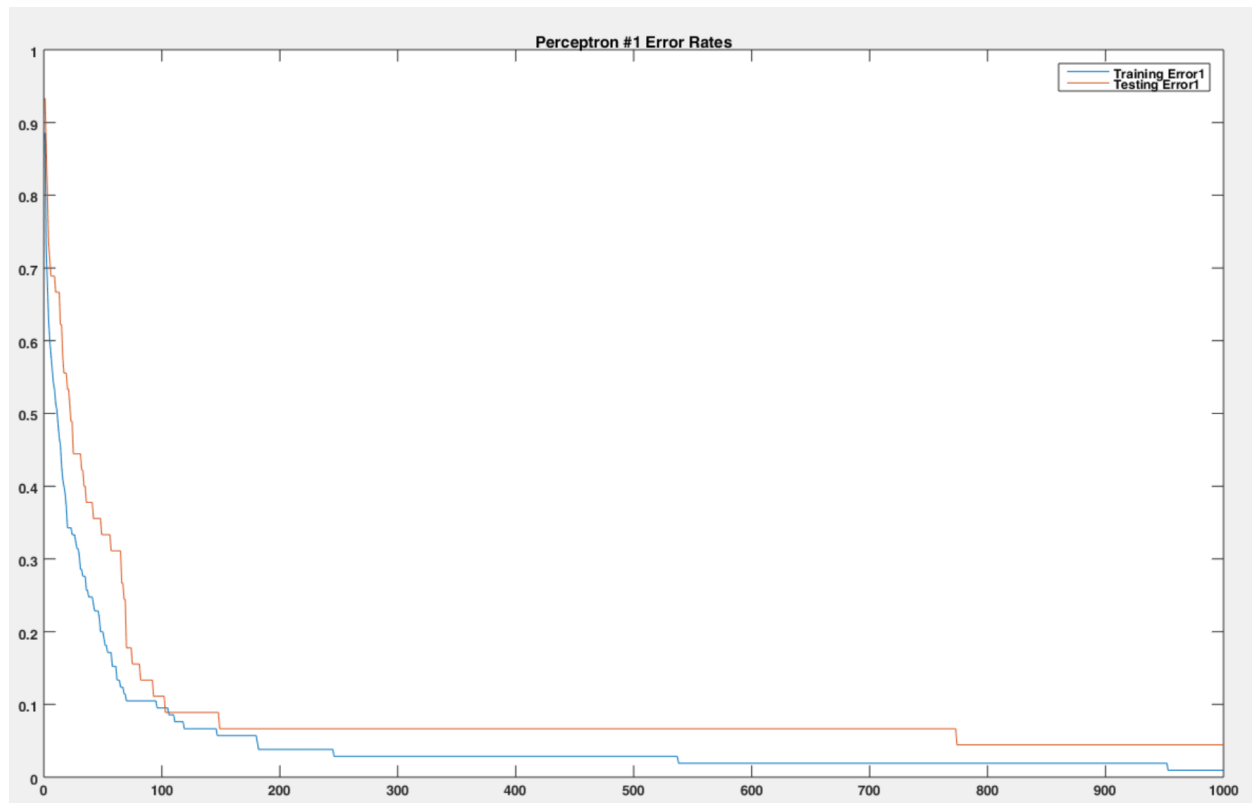
جداول زیر نتایج را در یک با اجرای برنامه برای شرایط آزمایش ذکر شده در قسمت قبل نشان می‌دهند.



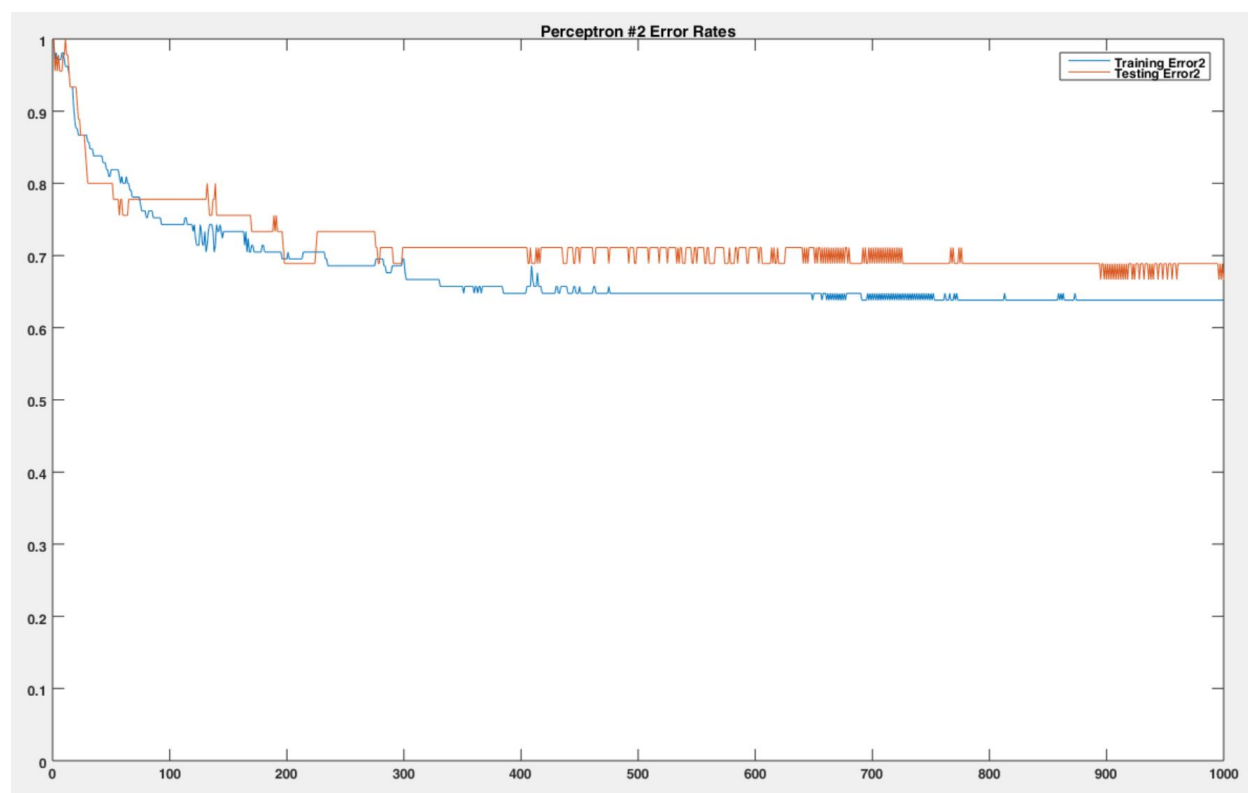
شکل ۱- نمودار خطای کل برای سوال یک (پرسپترون) نرخ یادگیری ۰.۵.



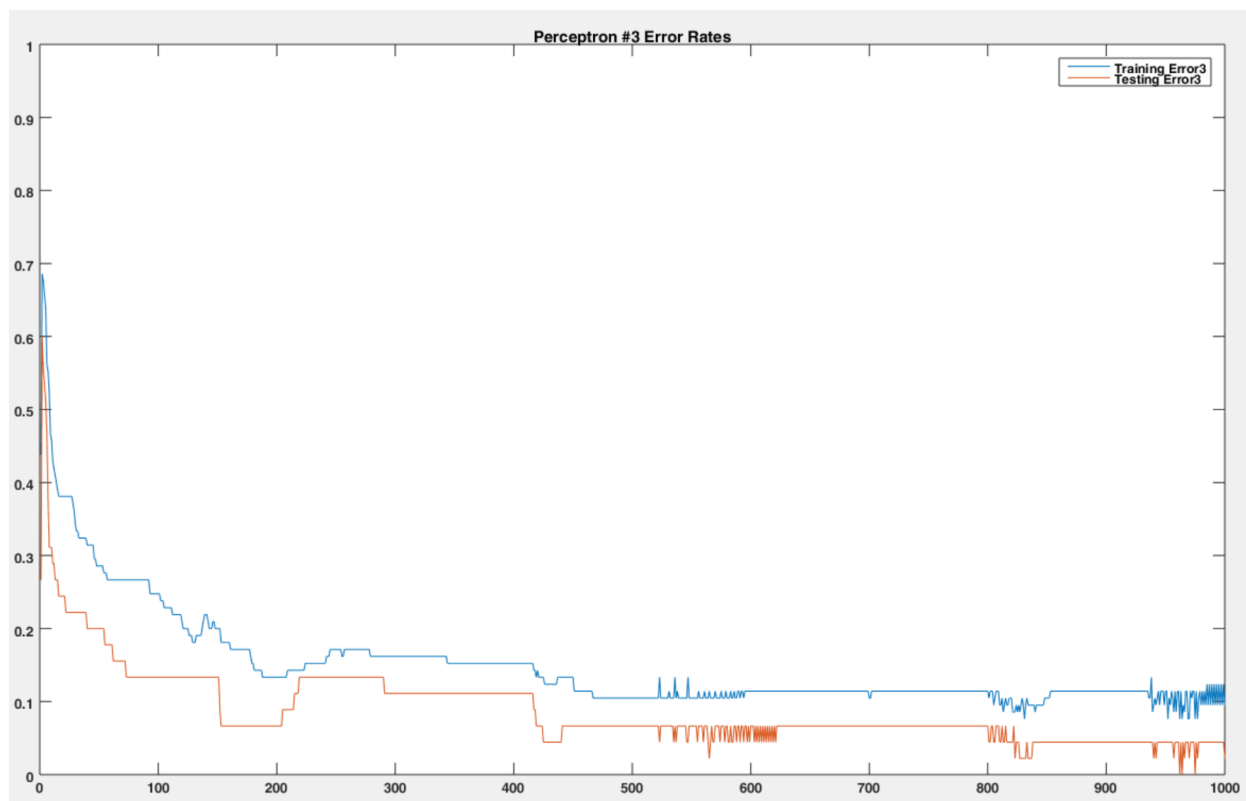
شکل ۲- بزرگ‌نمایی خطای ۲۵۰ اپیک اول برای سوال ۱ (پرسپترون)



شکل ۳- خطای پرسپترون شمای یک برای دسته‌بندی داده‌های دسته یک



شکل ۴- خطای پرسپترون شماره ۲ برای دسته‌بندی داده‌های دسته دو

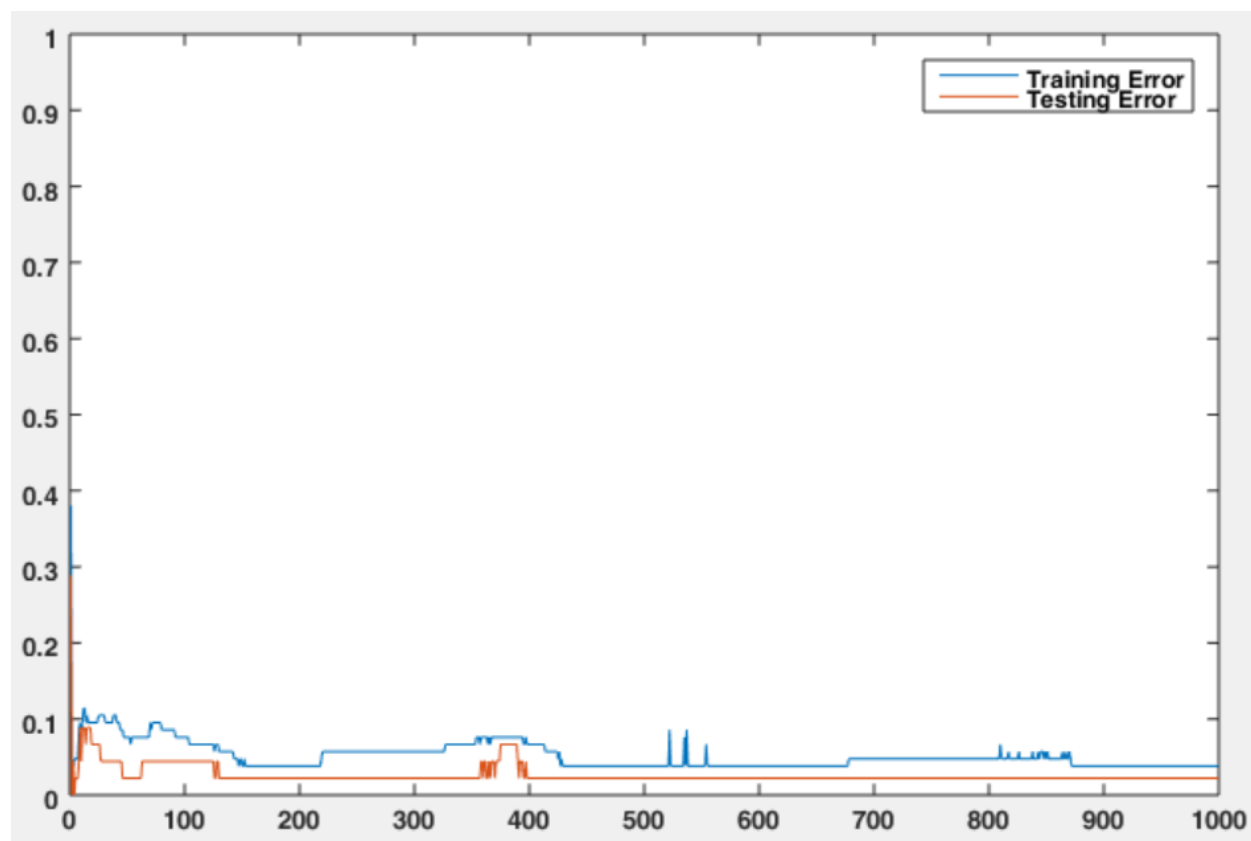


شکل ۵- خطای پرسپترون شماره ۳ برای دسته‌بندی داده‌های دسته سه

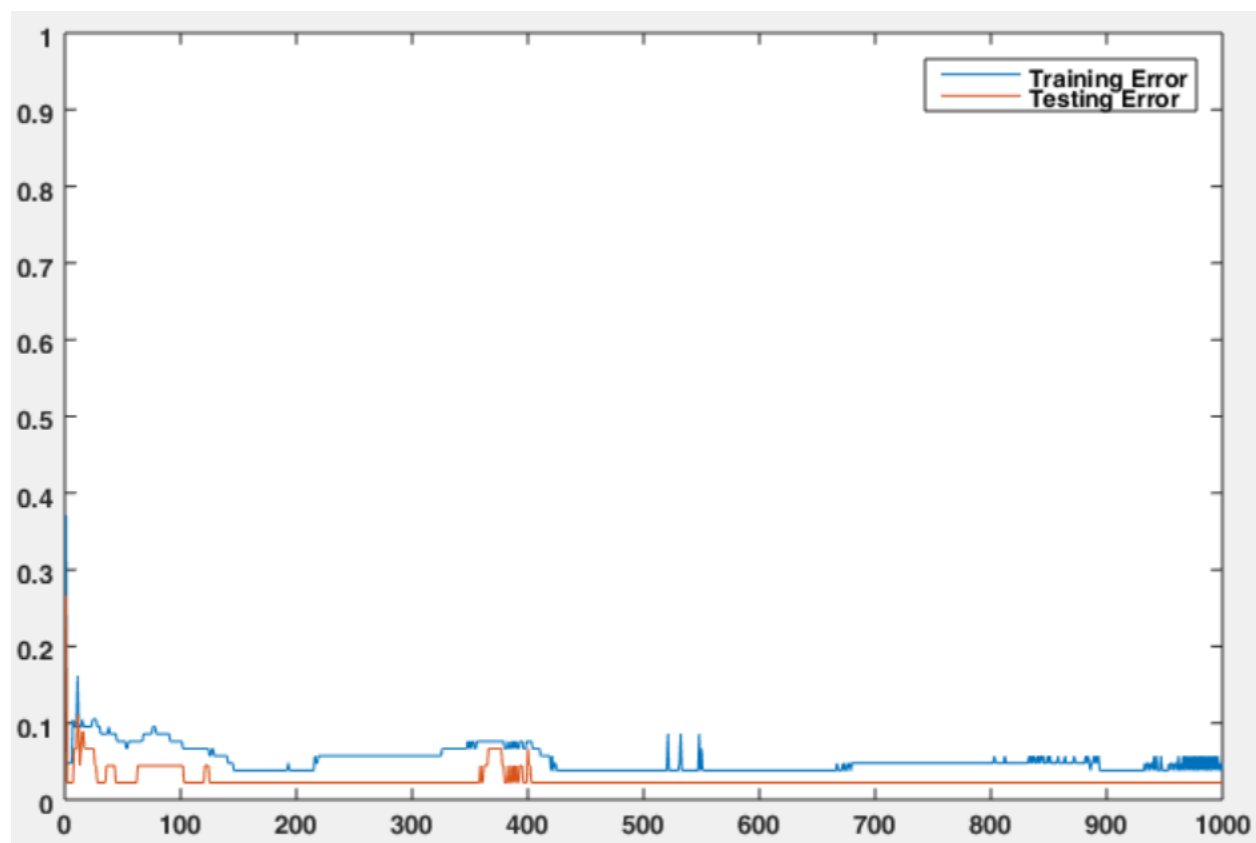
همان‌طور که از شکل‌ها پیداست، خطا برای پرسپترون شماره ۲ و پرسپترون شماره ۳ به صفر نمی‌رسد، اما برای پرسپترون شماره ۱ به نزدیکی صفر می‌رسد.

خطا در دسته‌بندی کلی دارای نوسانات شدید است.

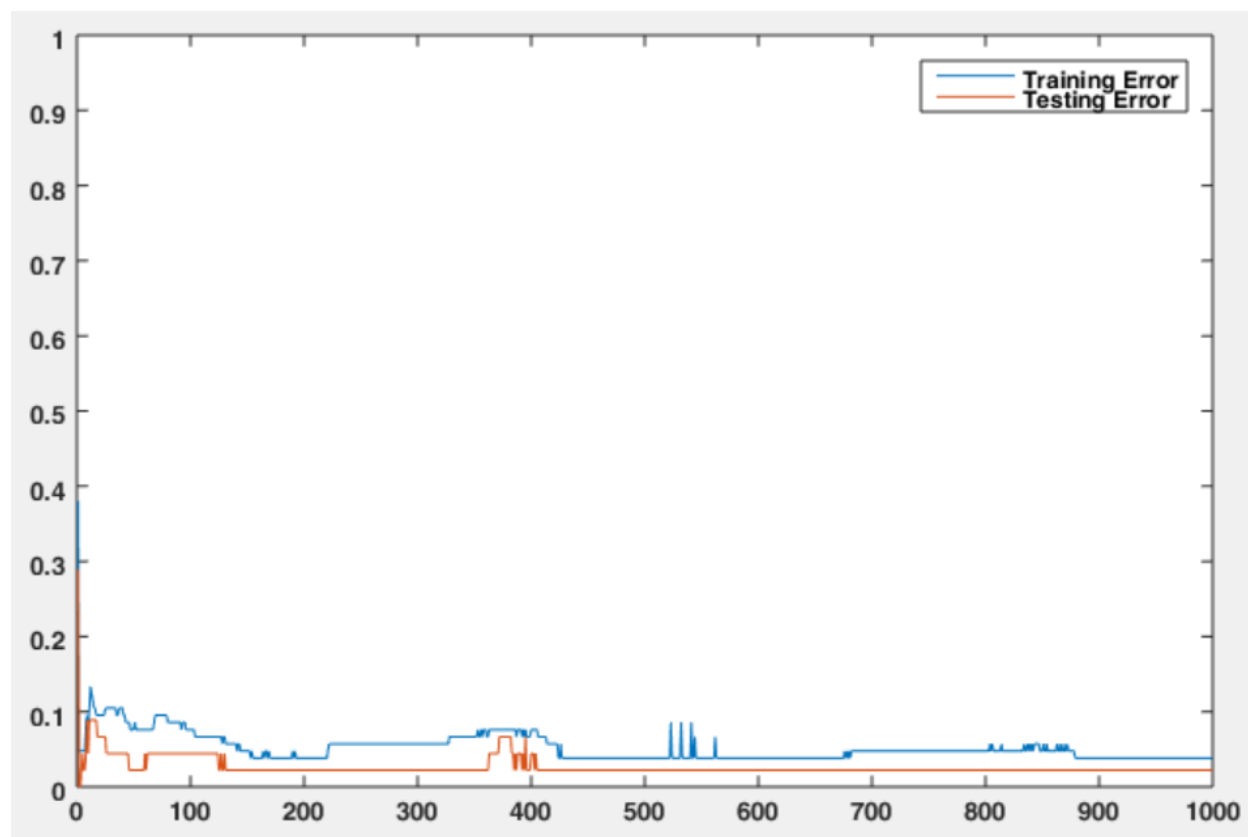
الگوریتم را با نرخ‌های صفر تا یک با گام‌هایی به اندازه ۰,۱ آزمودیم که نتایج آن به شکل زیر هستند.



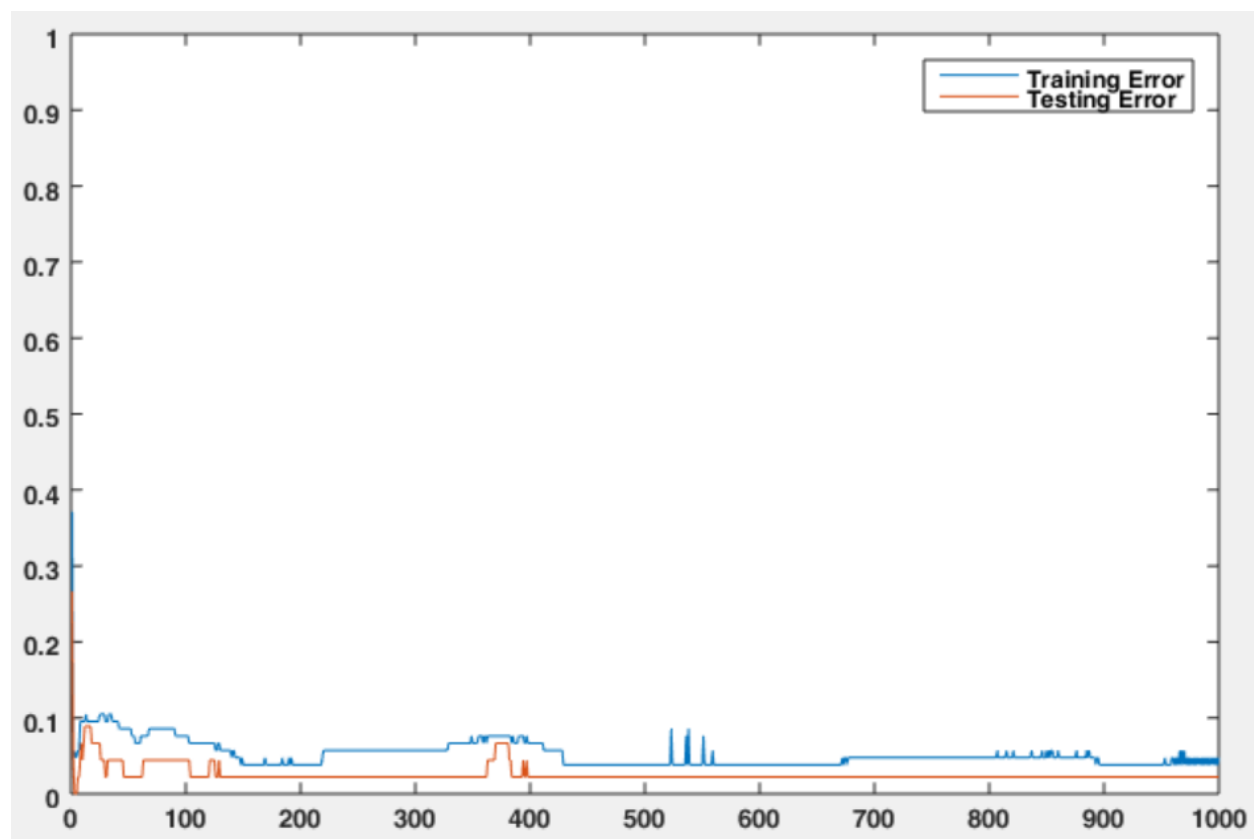
شکل ۶- نرخ یادگیری صفر



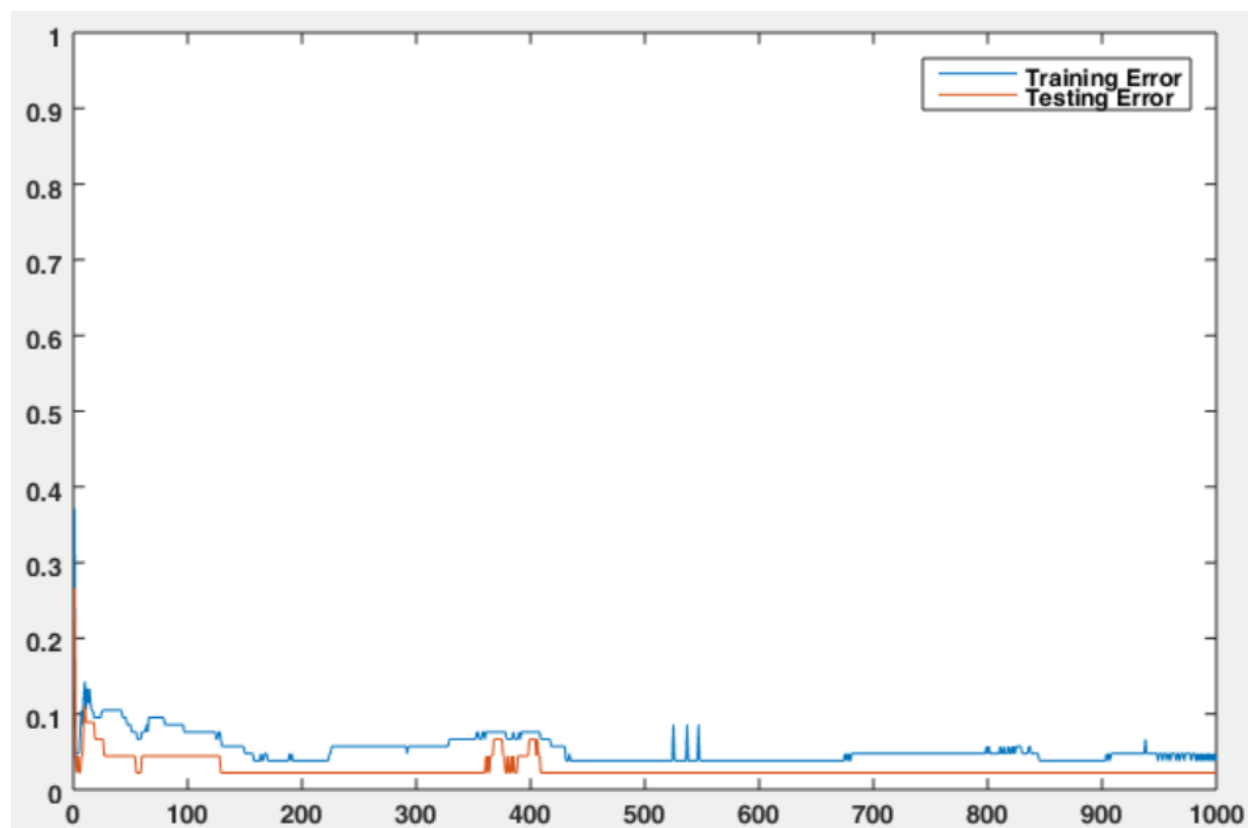
شکل ۷- نرخ یادگیری ۰,۱



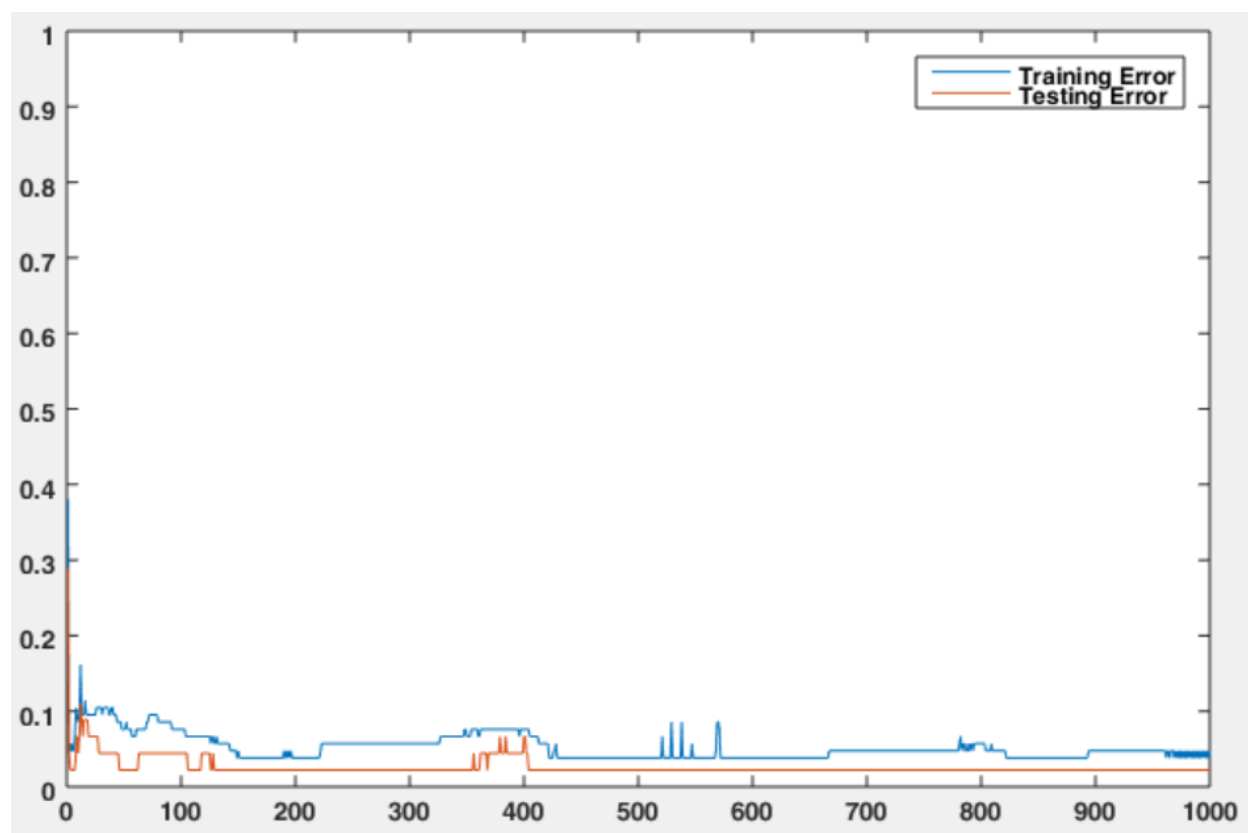
شکل ۱- نرخ یادگیری ۰,۲



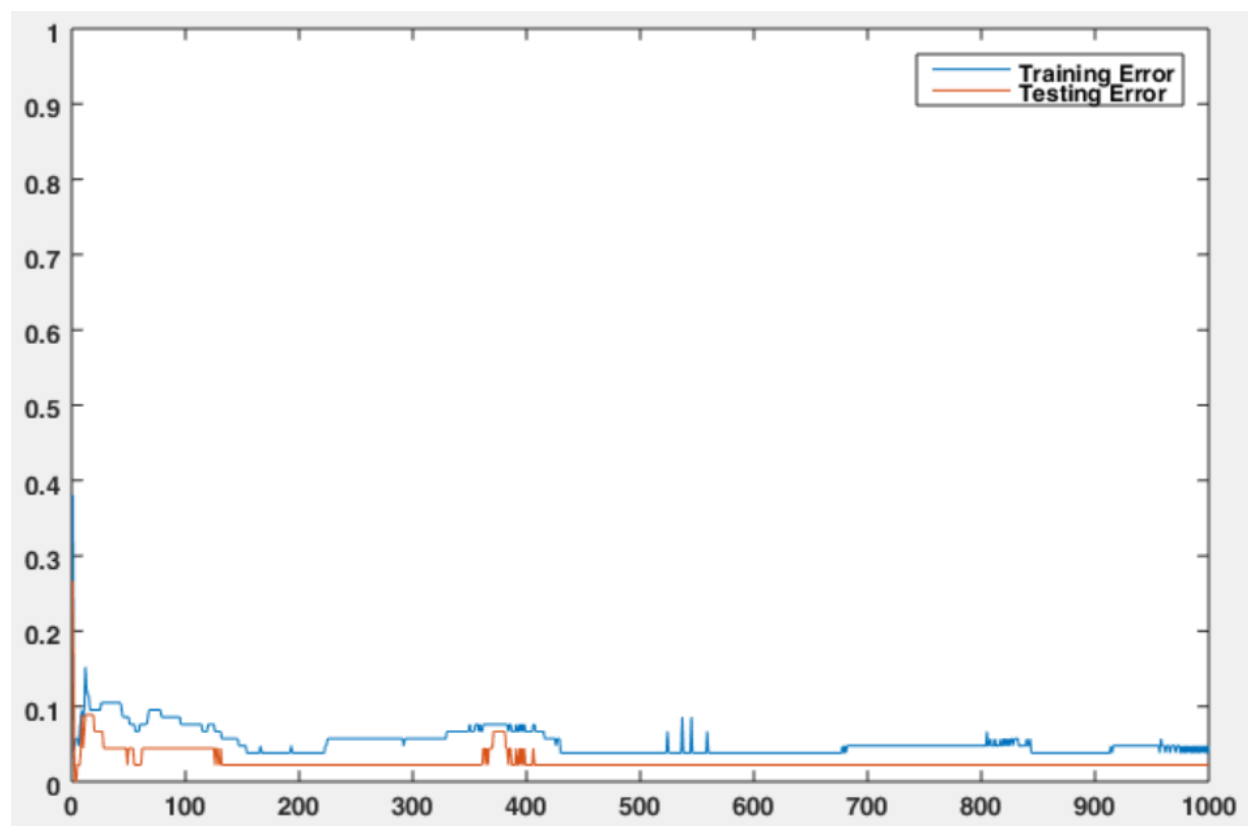
شکل ۹- نرخ یادگیری ۰,۳



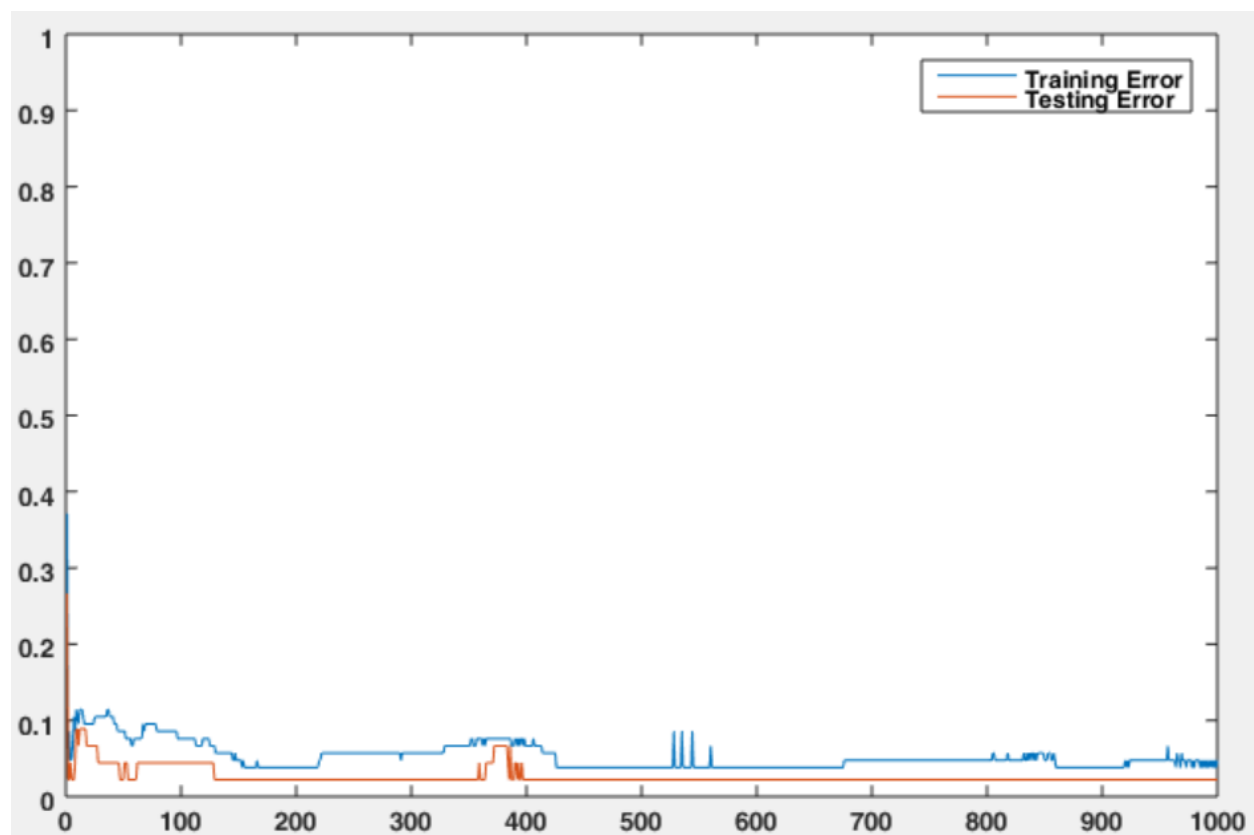
شکل ۱۰ - نرخ یادگیری ۰,۴



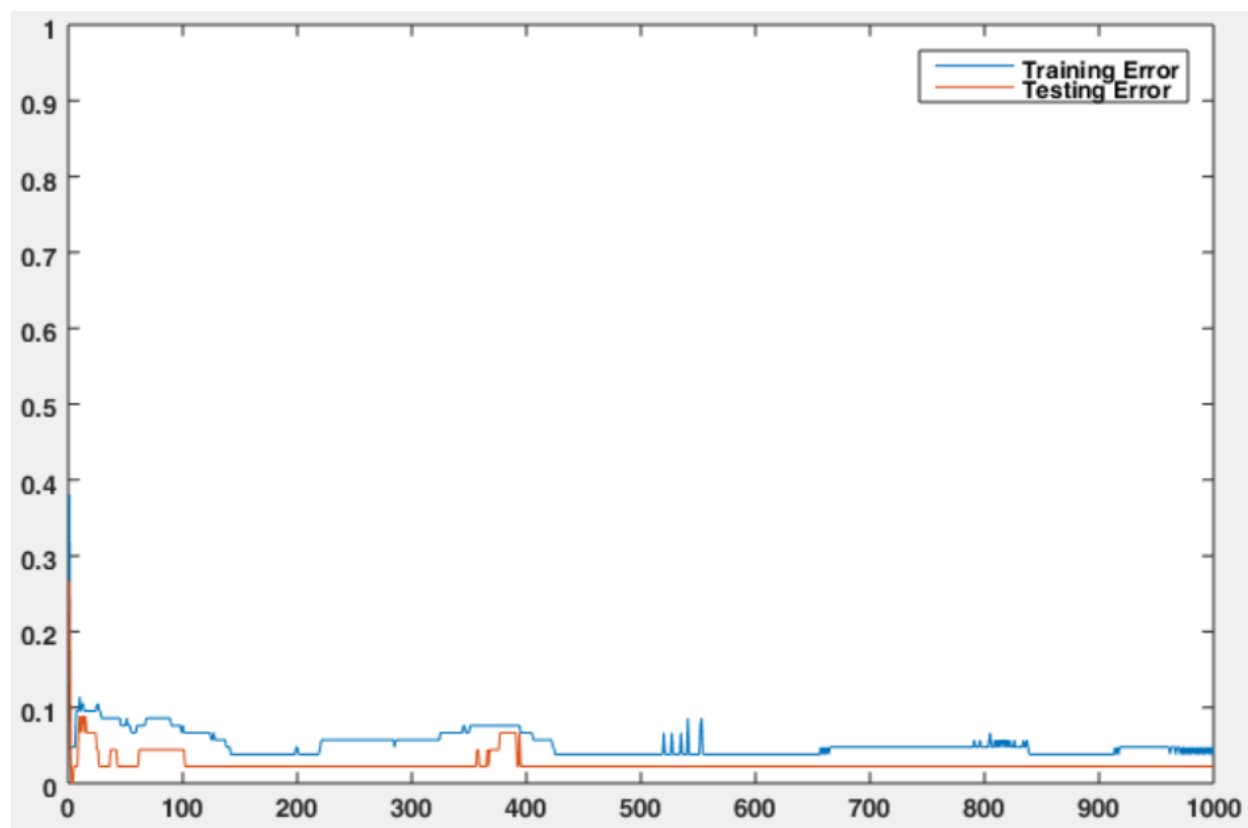
شکل ۱۱- نرخ یادگیری ۰,۰۰۱



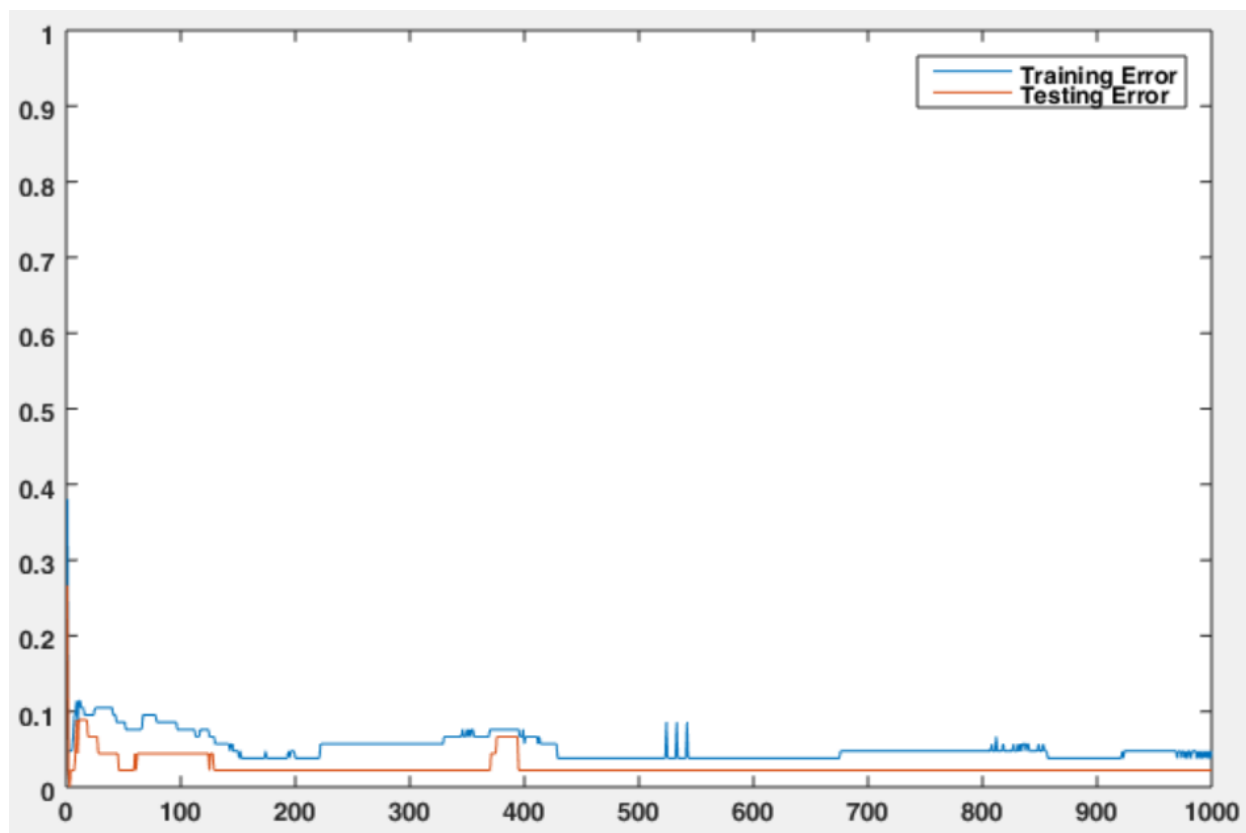
شکل ۱۲- نرخ یادگیری ۰,۷



شکل ۱۳ - نرخ یادگیری ۰,۱



شکل ۱۴ - نرخ یادگیری ۰,۹



شکل ۱۵- نرخ یادگیری یک (۱)

همان طور که مشاهده می شود با افزایش نرخ یادگیری، نمودار با شیب بیش تری نزول میکند اما نوسانات آن هم بیش تر است.

نتیجه گیری:

نتیجه ای که از مشاهدات بالا حاصل می شود این است که داده های کلاس ۱ از سایر داده ها به صورت خطی جداپذیر است، چرا که می توان با آموزش یک پرسپترون، آن ها را از هم جدا کرد و خطا به حدود صفر می رسد.

درمورد داده های کلاس ۲ می توان گفت که به نظر می رسد تقریباً به صورت خطی جداپذیرند اما نه به طور کامل؛ چرا که خطا به نزدیکی صفر میرسد اما در همان حدود متوقف می شود و به صفر نمی رسد.

در مورد داده های کلاس ۳ باید گفت که این داده ها مشخصاً به صورت خطی جداپذیر نیستند؛ چرا که با حدود ۱۰۰۰ دور آموزش الگوریتم، پرسپترون شماره به خطایی کم تر از حدود ۰,۶۵ نمی رسد.

نتیجه دیگر این است که با استفاده از الگوریتم پرسپترون و تکنیک دسته‌بندی یکی در برابر همه، برای این مسئله نمی‌توان نتیجه مناسبی گرفت.

از دلایل آن می‌توان به این اشاره کرد که از آن‌جا که بیشترین امتیاز بین سه دسته‌بندی کننده ملاک انتخاب ماست، نمی‌توان مطمئن بود که مقیاس امتیاز این سه دسته‌بندی کننده با هم برابر باشد.

مشکل دیگر این تکنیک این است که داده‌های آموزشی نامتعادل^۱ هستند. برای مثال اگر ده کلاس با تعداد داده‌های آموزشی برابر داشته باشیم، هر دسته‌بندی کننده خطی روی داده‌هایی آموزش می‌بندد که ۹۰٪ منفی و ۱۰٪ مثبت هستند و در نتیجه تقارن اولیه مسئله به هم می‌ریزد.^۲

منابع مورد استفاده:

<https://class.coursera.org/ml-005/lecture/38>

Bishop, Christopher M. (2006). Pattern Recognition and Machine Learning. Springer.

Rifkin, Ryan. "Multiclass Classification." Lecture Slides. February (2008).
(<http://www.mit.edu/~9.520/spring09/Classes/multiclass.pdf>)

سوال ۲

دسته‌بندی داده‌های دو بعدی و دو کلاسه با استفاده از پرسپترون.

شرایط آزمایش:

| پارامتر | مقدار |
|------------|-------|
| تعداد اپیک | ۱۰۰۰ |

^۱ imbalanced

^۲ Bishop, Christopher M. (2006). Pattern Recognition and Machine Learning. Springer. Page 354.

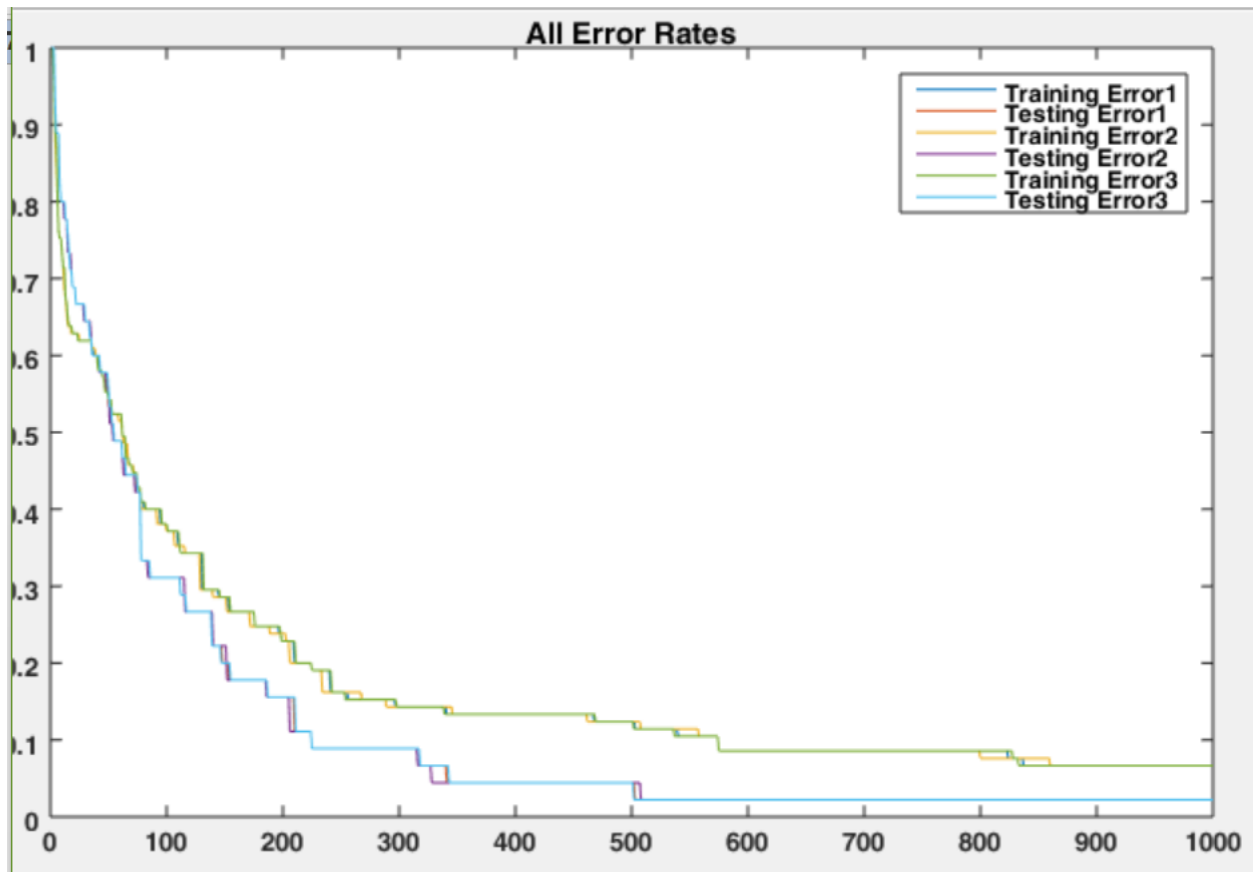
| | |
|------------------|----------|
| نرخ یادگیری | ۰.۵ |
| درصد داده آموزشی | ٪۷۰ |
| درصد داده آزمون | ٪۳۰ |
| تابع فعالیت | سیگموئید |

برای این آزمایش برای این که داده های با کلاس های مختلف را مخلوط کنیم و مخلوط برای تمام دفعات اجرای آزمایش ثابت باشد، ابتدا با استفاده از متلب، ترکیب داده ها را به صورت تصادفی به هم زدیم (برنامه DataShuffler.m به همین منظور نوشته شده است).

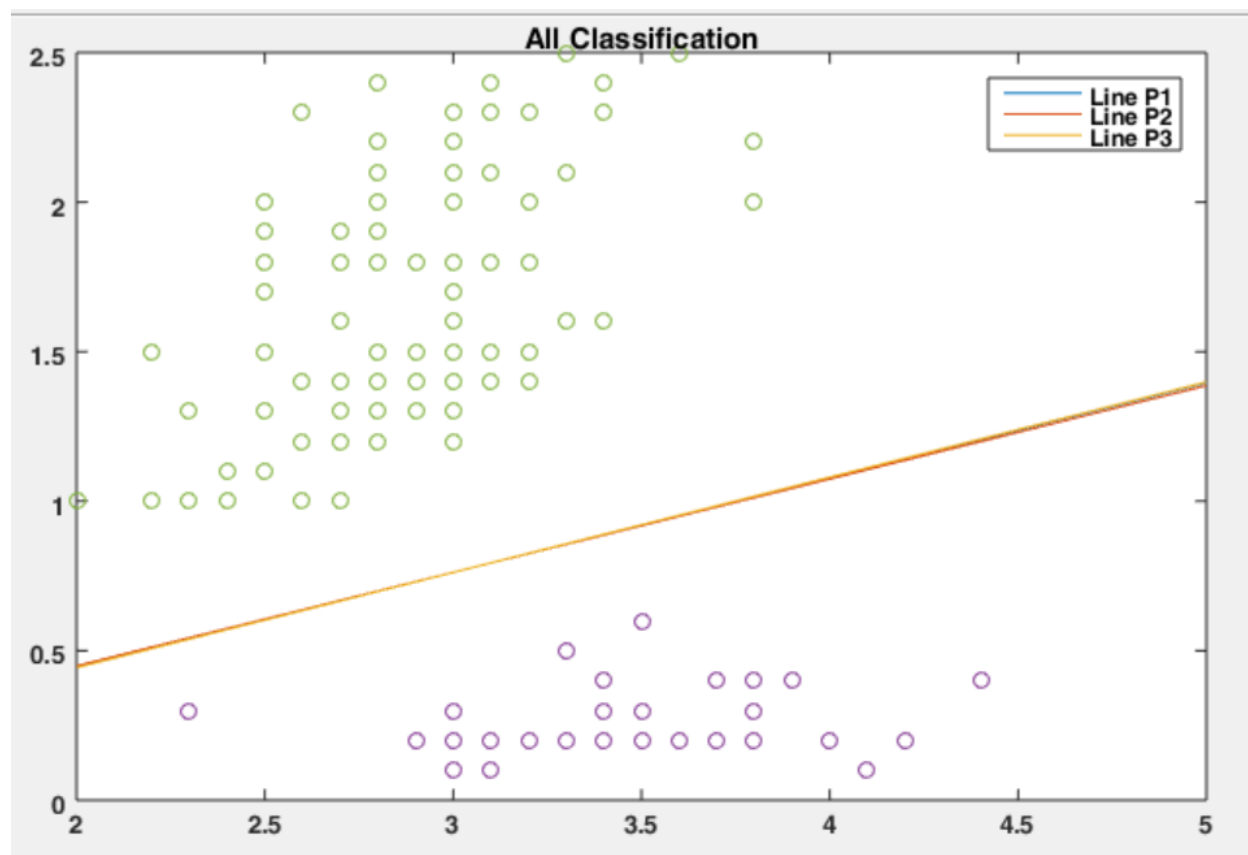
برای برنامه اصلی از فایل P_1_2_G.m استفاده کردیم.

نتیجه انجام آزمایش:

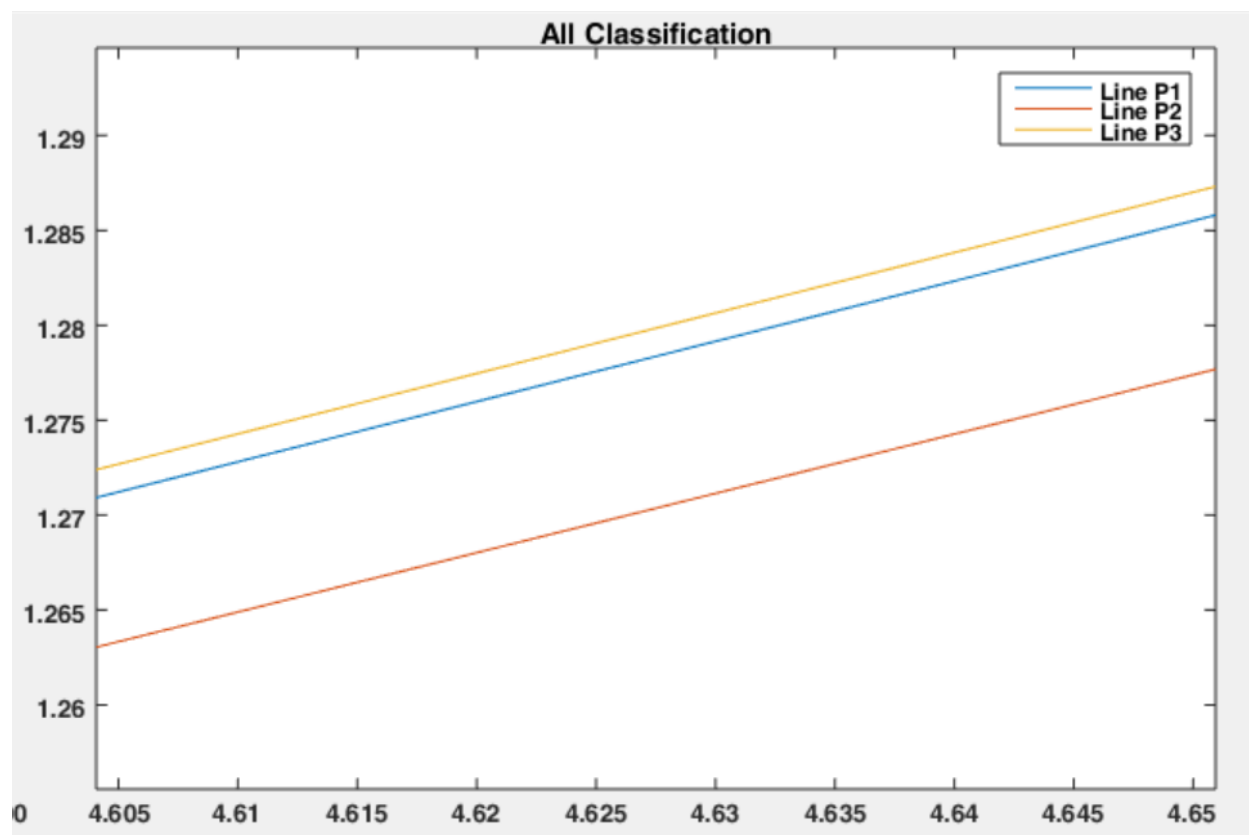
پس از اجرای این آزمایش نمودارهای به دست آمده نشان می دادند که با تعداد تکرار بالای الگوریتم، خطوط دسته بندی جداکننده تفاوت های بسیار کمی دارند و تقریباً روی هم می افتند. نمودارهای جداگانه و ترکیب آنها در یک نمودار در شکل های زیر مشاهده می شوند.



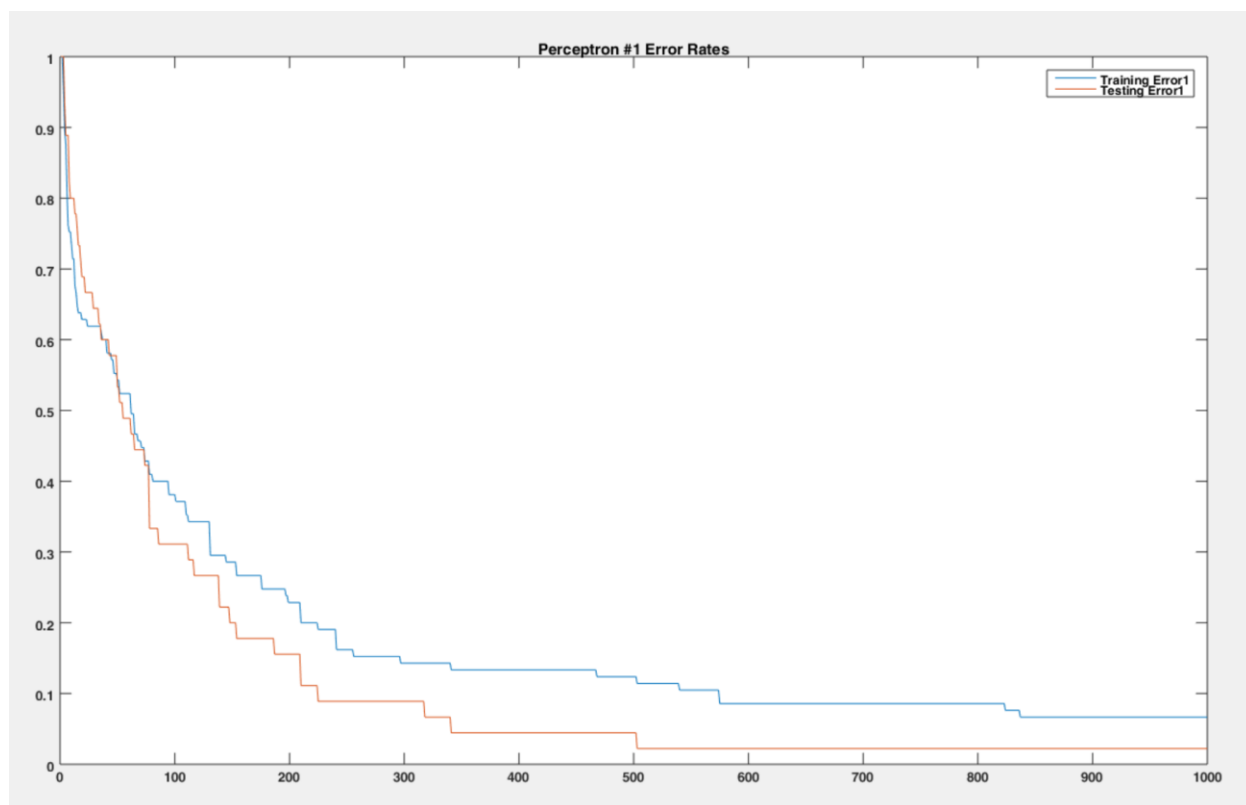
شکل ۱۶- خطای آموزش و آزمایش برای سه پرسپترون ۱ و ۲ و ۳



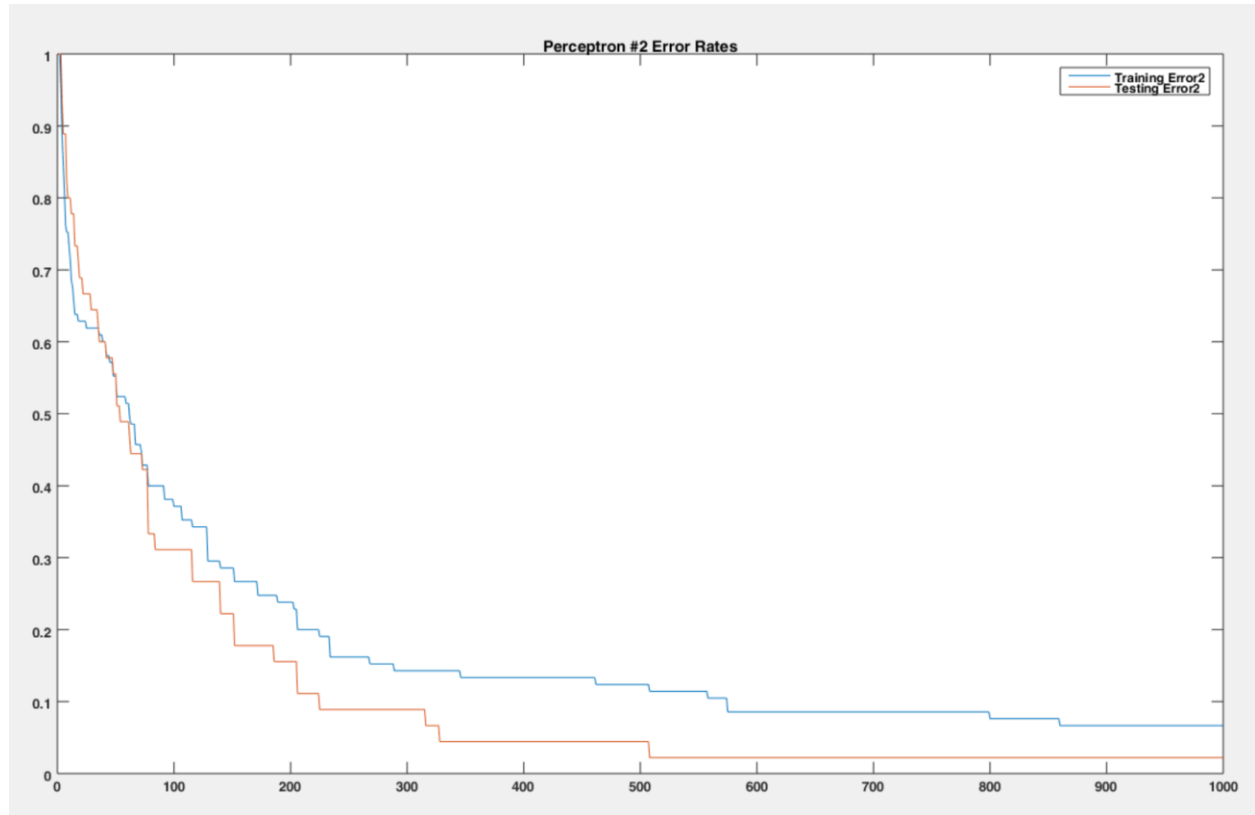
شکل ۱۷- داده‌های دو کلاس و سه خط جداکننده حاصل از سه پرسپترون



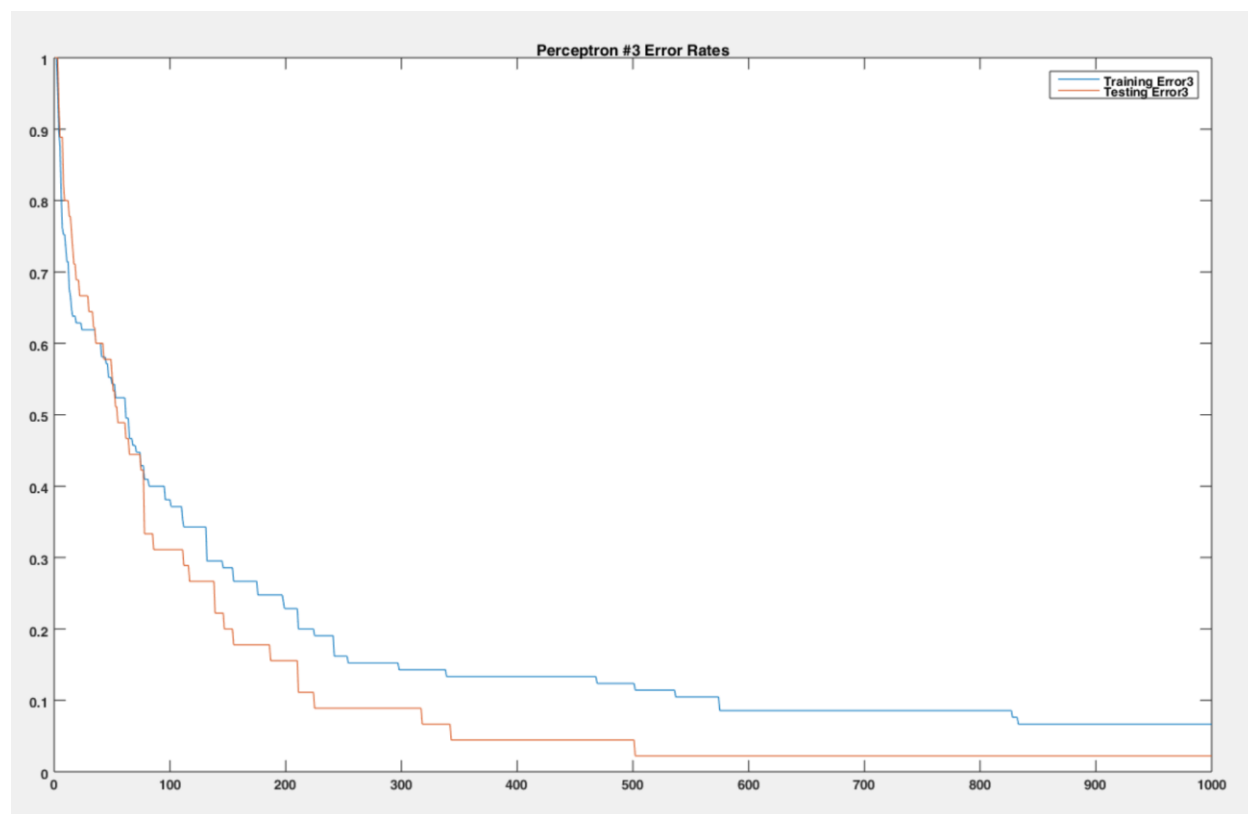
شکل ۱۸- بزرگ‌نمایی شکل ۱۷ برای نشان دادن تمایز ظریف سه خط جداکننده



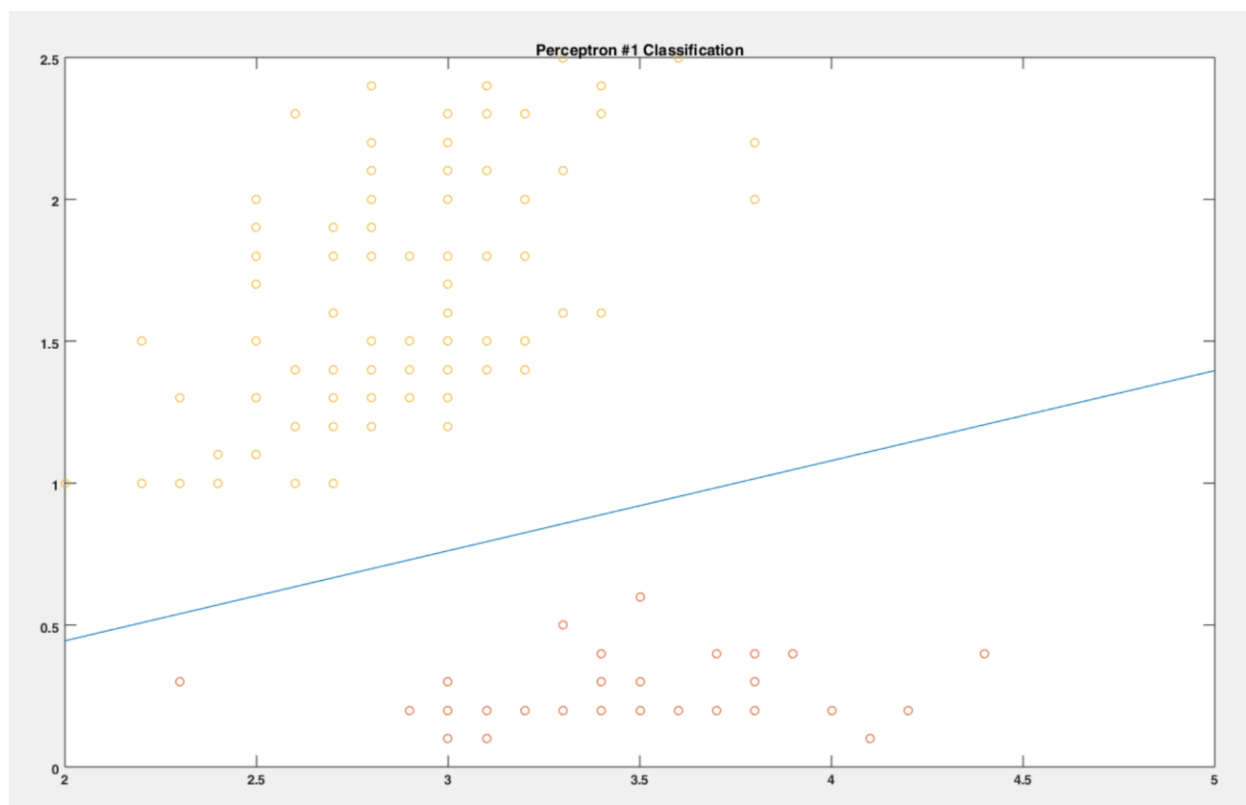
شکل ۱۹- نمودار خطای پرسپترون شماره ۱



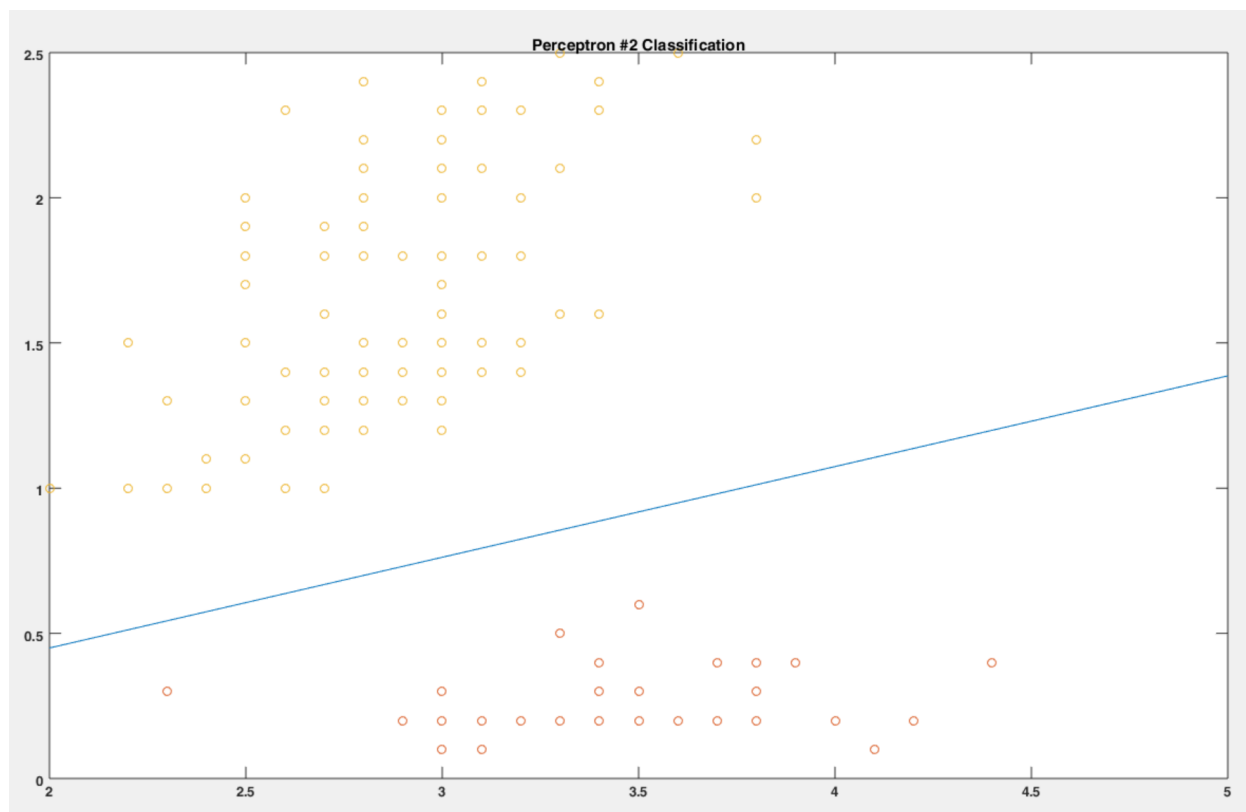
شکل ۲۰ - نمودار خطای پرسپترون شماره ۲



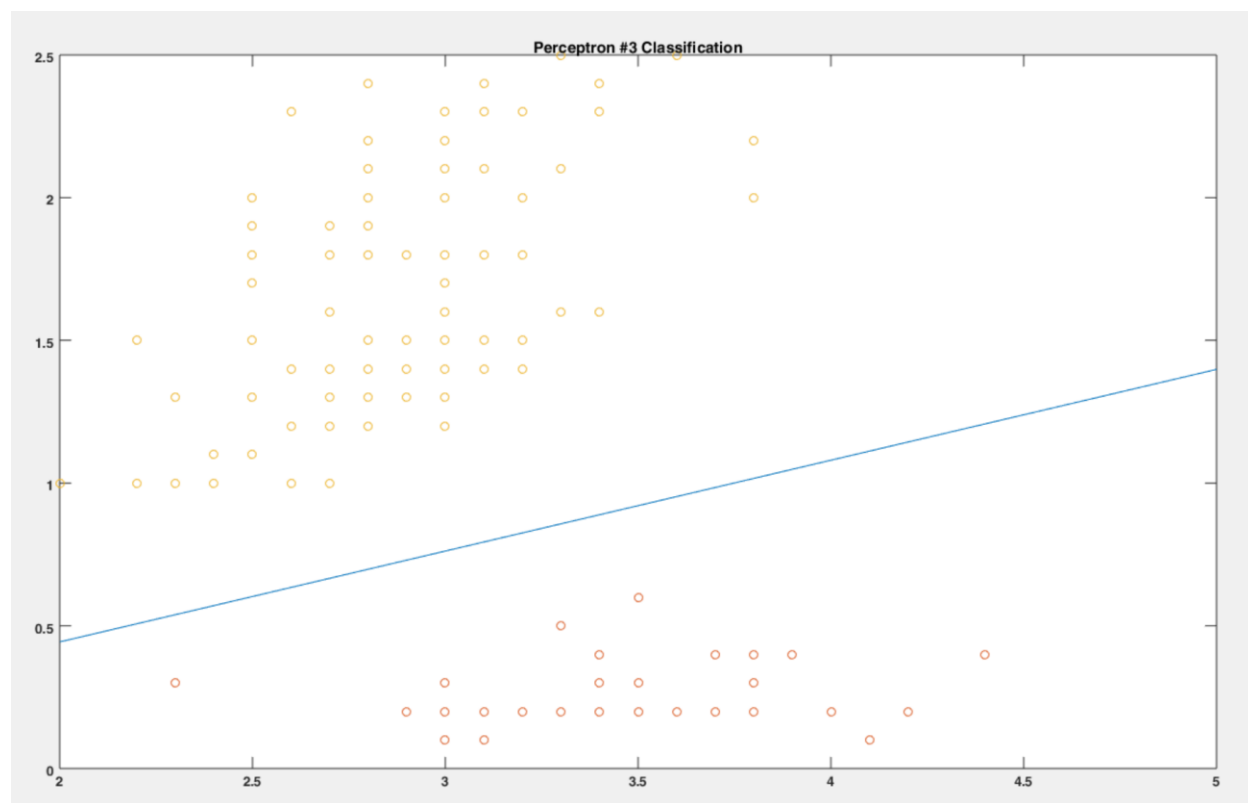
شکل ۲۱- نمودار خطای پرسپترون شماره ۳



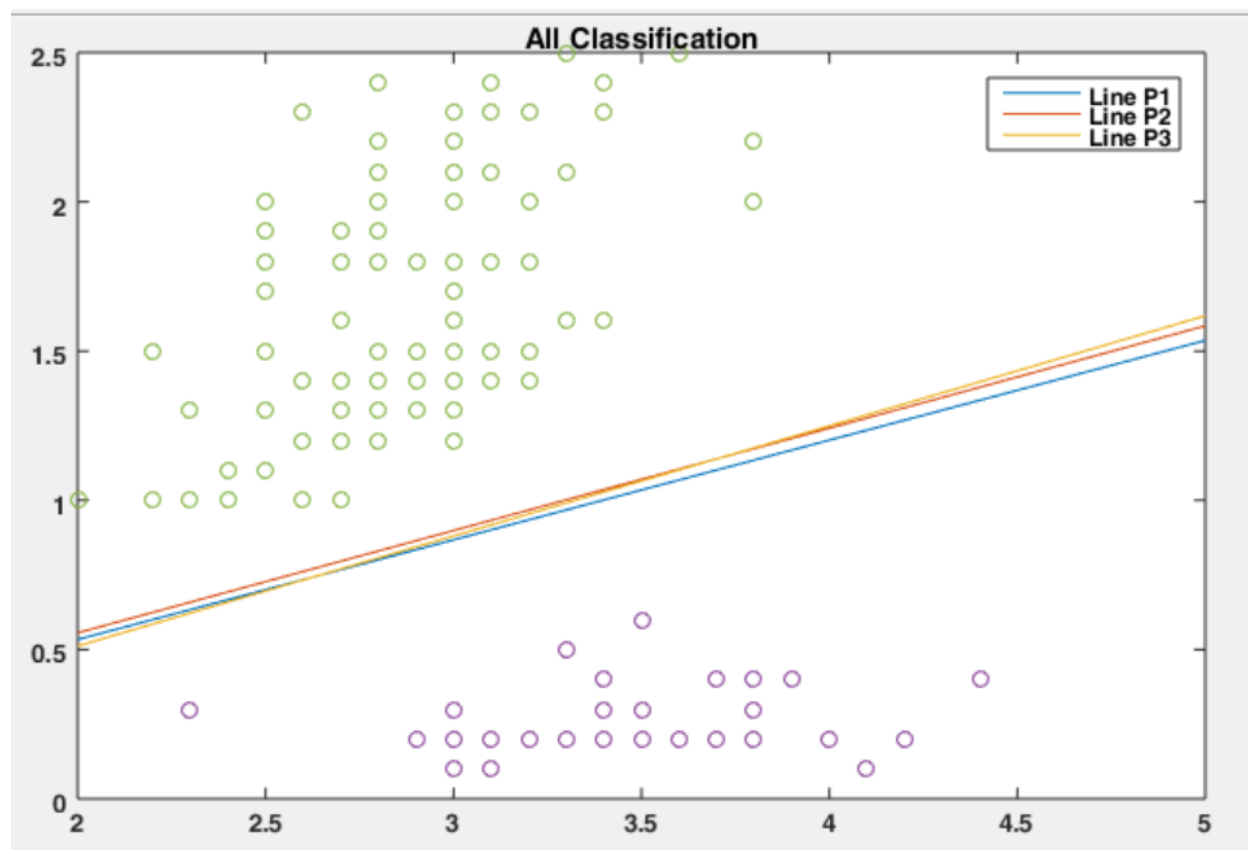
شکل ۲۲- نتیجه دسته‌بندی پرسپترون شماره ۱



شکل ۲۳ - نتیجه دسته‌بندی پرسپترون شماره ۲



شکل ۲۴ - نتیجه دسته‌بندی پرسپترون شماره ۳



شکل ۲۵ - نتیجه حاصل دسته‌بندی سه پرسپترون با تعداد ایک ۱

همان‌گونه که در شکل‌ها مشاهده می‌شود، در تعداد تکرار ۱ الگوریتم، سه خط جداکننده حاصل فاصله خیلی محسوس‌تری نسبت به حالت قبل دارند.

نتیجه‌گیری:

با توجه به نتایج به دست آمده می‌توان این نتیجه را گرفت که با شروع از هر نقطه‌ای، با تعداد تکرار مناسب، وزن‌های پرسپترون به نقطه بهینه هم‌گرا می‌شود.

نتیجه دیگر این که با توجه به زیاد بودن داده‌های آموزشی ما، با یک بار تکرار الگوریتم روی این مجموعه آموزشی، باز هم سه خط به نسبت خیلی خوبی به هم نزدیک می‌شوند.

نتیجه دیگر این که مطابق صورت سوال، داده‌ها به صورت خطی از هم جدا شدند و این نکته مشاهده شد.