



دانشگاه صنعتی امیرکبیر  
(پلی تکنیک تهران)

دانشکده مهندسی کامپیوتر و فناوری اطلاعات

## تکلیف شماره ۱ درس رایانش عصبی

استاد  
دکتر صفابخش

امیرحسین کاشانی

### سوال الف )

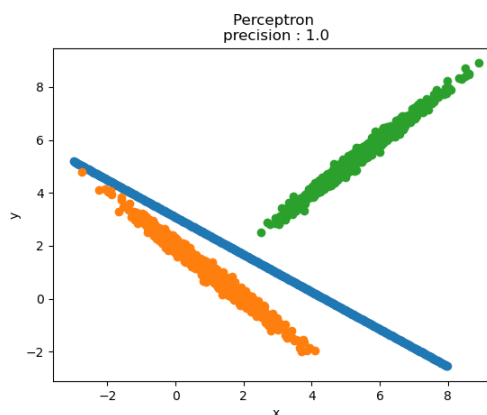
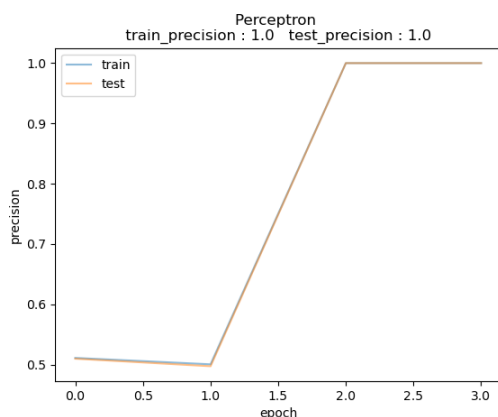
نرون آدالین در برخی موارد نمی تواند به دقت ۱۰۰ درصد برسد. برای این مورد می توان دو دلیل ذکر کرد، یک اینکه در خیلی از موارد در صورتی که تعداد گام ها و نرخ یادگیری به درستی انتخاب نشود ممکن است دچار پدیده انفجار گرادیان شده و به طور کلی همگرا نشود و حتی **stack overflow** نیز به دلیل بزرگ شدن ناگهانی خطا ها نیز بوجود آید یا از طرف دیگر به دلیل کافی نبودن **epoch** ها به جواب مورد نظر نرسد. دلیل دومی که می توان برای این موضوع بیان کرد این است که آدالین یک عبارت را دارد بهینه می کند و نسبت به **activation function** مورد استفاده دیدی ندارد و ممکن است در برخی از حالات مرزی صرفا به کاهش خطا توجه کند نه به اینکه مقدار گزارش داده شده از مرز تصمیم گیری ما هست عبور کند. برای مثال فرض کنید ۳ داده داریم که ارزش حقیق آن ها ۱- می باشد. از نظر آدالین پیش بینی 0.1 0.1 0.1 پیش بینی بهتری نسبت به 0.1 0.1 1- می باشد در صورتی که دقت پیش بینی اول 0 و پیش بینی دوم 0.33 می باشد.

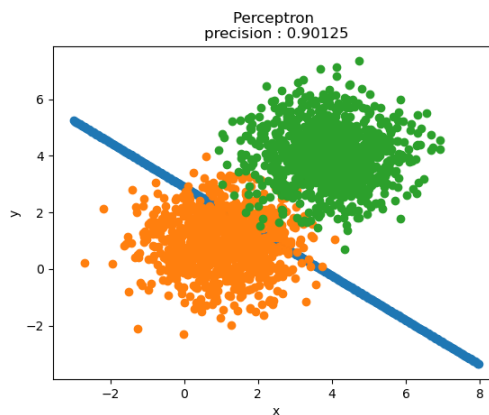
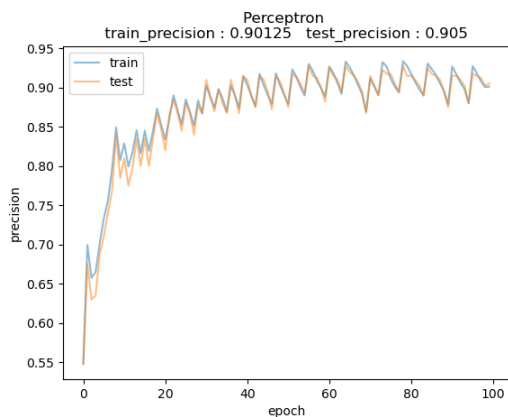
### سوال ب)

آدالین بهتر عمل می کند.

با توجه به اینکه آدالین یک مجموعه پیوسته از ضرایب و مقادیر را در این مساله تحلیل می کند اما پرسپترون فقط به خروجی نگاه می کند و همچنین از نرخ آموزش نیز استفاده نمی کند احتمال گیر کردن در مینیمم های محلی برای پرسپترون بیشتر است و در نتیجه آدالین نتیجه بهتری نسبت به پرسپترون خواهد داشت.

### سوال ج )

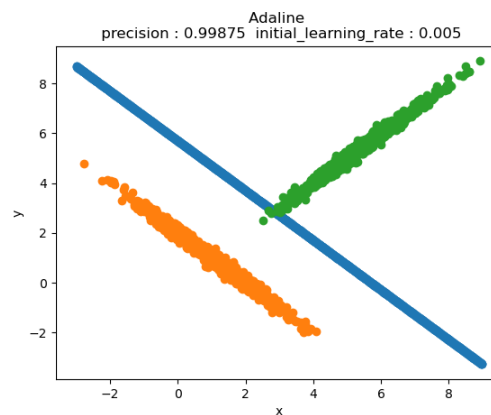
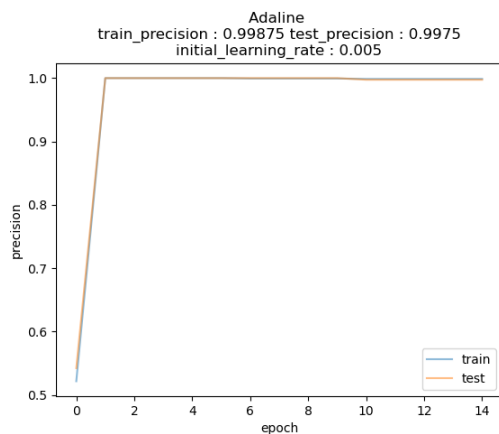
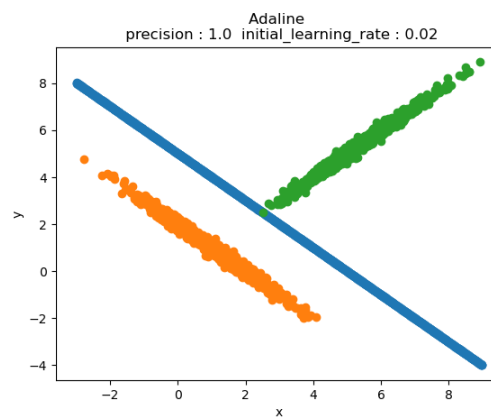
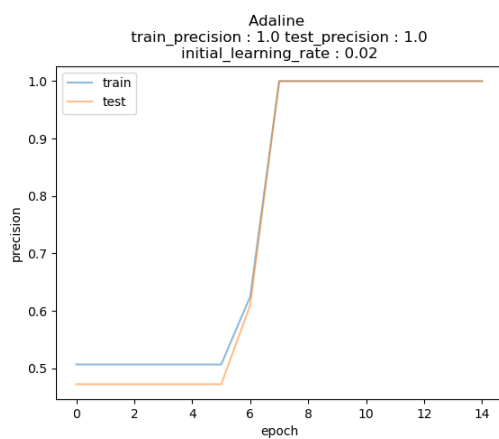


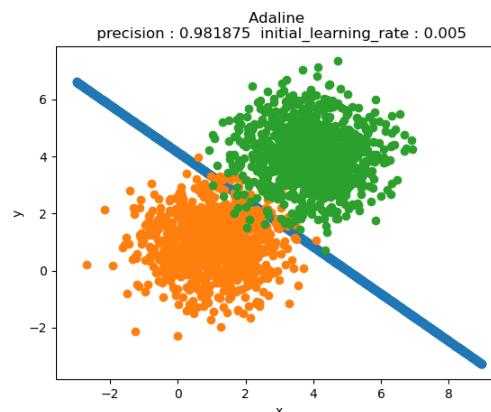
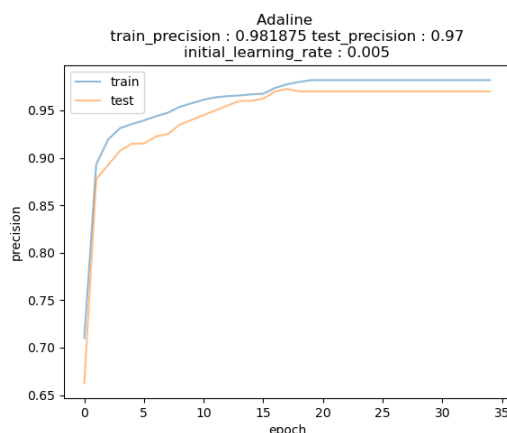


سوال د)

علت اینکه باید برچسب داده ها تغییر کند این است که خروجی آدالین یک تابع bipolar است که خروجی آن یا یک یا منفی یک می باشد. در پیاده سازی به جای تغییر در دیتا در موقعی که خطا محاسبه شده و زمان تصمیم گیری با افزودن یک شرط این مساله بر طرف گردیده است. ضریب یادگیری آدالین به مرور زمان در هر epoch با فرمول زیر کاهش می یابد.

$$\eta = \eta - \eta * \left( \frac{epoch\_iteration}{maximum\_epoch * 1.1} \right)$$





سوال ه)

با توجه به نتایج بدست آمده می بینیم که آدالین در حالتی که کاملاً جدایی پذیر خطی هستند به عملکرد ۱۰۰ درصدی نرسیده است از طرفی در حالتی که جداپذیر خطی نبوده اند حدود ۷ درصد عملکرد بهتری داشته است.

نکته حائز اهمیت در این پیاده سازی این است که در پرسپترون ۱ زحالات پایه استفاده شده که ضریب برای یادگیری فرض نمی شود مانند فرمول زیر

$$w_i = w_i + w_{i-1} * (t - y)$$

در نتیجه نرخ یادگیری در این پرسپترون برابر یک است و دلیل این اختلاف این می باشد. در صورتی که نرخ یادگیری به معادله پرسپترون اضافه گردد و این مقدار برابر مقداری کوچک در نظر گرفته شود می توان حتی امیدوار بود در برخی از تکرار ها نتایج بهتر از آدالین باشد.

سوال و)

روش به کار گرفته شده در این بخش به اینصورت می باشد که به جای استفاده از  $x, y$  از  $x^2, y^2, xy$  متغیر به عنوان ورودی استفاده شده است که متغیرهای به کار گرفته شده عبارت اند از  $x, x^2, y, y^2, xy$  در نهایت معادله خط درجه ۲ بدست می آید که با استفاده از `matplotlib` به نمایش در آمده است. در این سوال نیز از کاهش نرخ آموز که در صفحه قبل نشان داده شده استفاده نمودیم.

