به نام خدا

یادگیری ماشین

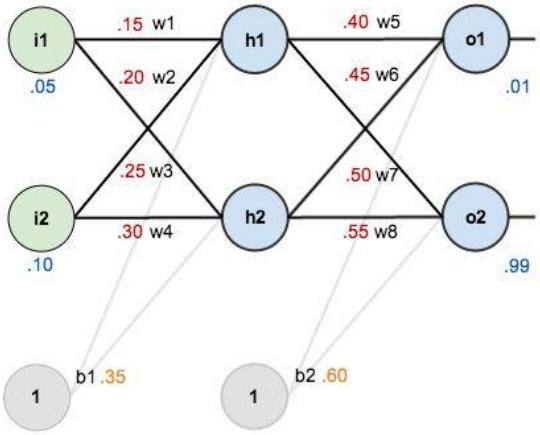
تکلیف سوم– پاییز ۱۳۹۶ مهلت ارسال: ۸ دی

قبل از اینکه شروع به حل تمرین‌ها کنید، حتما یک بار فایل Policies.pdf (موجود در وبگاه) که حاوی نکات مهم در تحویل تکالیف هست را مطالعه فرمایید.

حل سوالاتی که امتیازی هستند، اجباری نبوده و نمره ی اضافی خواهد داشت.

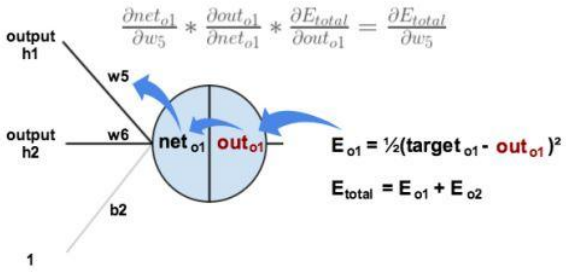
طراح هریک از سوالات در زیر مشخص شده است .در صورتی که سوالی دارید می توانید سوالتان را در گروه مطرح کنید یا از دستیار درس بپرسید.

۱- در این مساله می‌خواهیم یک‌بار مراحل back-propagation جهت به دست آوردن اوزان بهینه را طی نماییم. همان‌طور که در شکل زیر مشخص است. اوزان اولیه، مقادیر bias داده‌های train ورودی و خروجی داده شده است.



شکل ۱

الف) با انجام یک مرحله back-propagation مقادیر update شده w5، w6، w7 و w8 را به دست آورید.(راهنمایی: می‌توانید از شکل زیر استفاده نمایید.)



شکل ۲

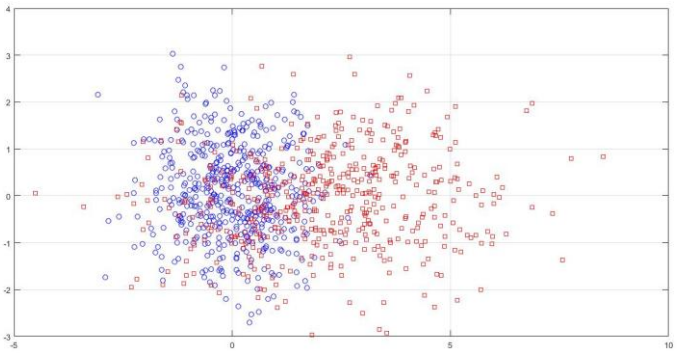
ب) (امتیازی) با استفاده از نتایج بخش اول مقادیر update شده w1، w2، w3 و w4 را بدست آورید.

۲- می‌دانیم که برای یک شبکه عصبی می‌توانیم توابع activation متفاوتی داشته باشیم. برای نمونه تابع خطی، threshold activation و تابع سیگموئید نمونه‌هایی از این توابع هستند.

الف) ادعا می‌شود که threshold activation برای training مناسب نیستند. آیا این ادعا درست است. توضیح دهید.

ب) آیا می‌توانیم با تغییراتی در یک شبکه عصبی با لایه مخفی، شبکه عصبی بدون لایه مخفی با همان کارکرد بسازیم.(توضیح دهید)

۳- در سوال ۷ ام سری نخست تمرینات با استفاده از Bayesian Decision Boundary مرز تصمیمی برای جدا نمودن دو کلاس C_1 و C_2 پیدا نمودید حال سعی کنید با استفاده از داده‌های مربوط به آن سوال به پرسش های زیر پاسخ دهید:



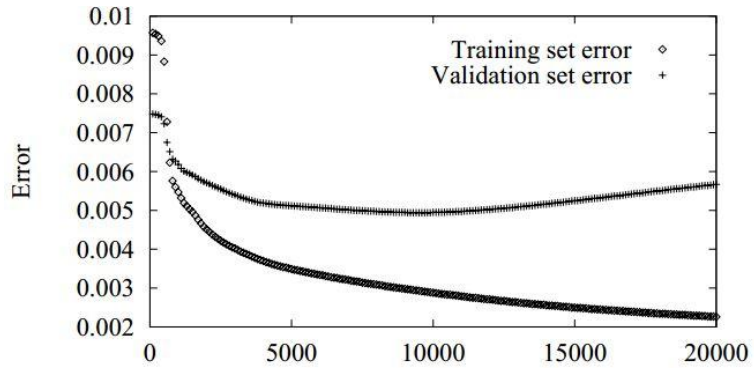
شکل ۳

الف) سعی کنید با استفاده از شبکه عصبی دو لایه این مرز را ترسیم نمایید.

ب) قسمت قبل را با استفاده از شبکه عصبی سه لایه با تعداد واحدهای مخفی مختلف (۲، ۷، ۱۵) بدست آورید.(برای نمونه شکل ۱، دارای ۲ واحد مخفی است.)

۴- در یک مساله تشخیص پزشکی، می‌خواهیم با استفاده از الگوریتم Back-propagation یک شبکه عصبی train نماییم. برای تخمین مقدار داده های لازم برای training از یک تکنیک validation ساده استفاده نموده‌ایم. برای این منظـور، ما هبه صورت تصادفی داده ها را به دو بخش با اندازه مساوی تقسیم کرده‌ایم. یک بخش به عنوان training و بخـش دیگر به عنوان validation. منحنی خطای این دو مجموعه داده در شکل زیر رسم شده است. همانطور کـه مشـخص

است training Error بصورت یکنواخت با زیاد شدن تعداد batch update کاهش می‌یابد؛ در حالی که این اتفاق بـرای validation Error رخ نمی‌دهد.



شکل 4-منحنی های خطا با استفاده از Batch-mode Gradian Descent

حال فرض کنید می‌خواهیم با استفاده از الگوریتم قبلی یک شبکه عصبی دیگر train نماییم با این تفاوت که به تعداد ۱۰ برابر داده های قسمت قبلی داده در اختیار داریم.

الف) پیشبینی نمایید در این حالت منحنی training error به چه صورت می‌شود. شکلی تقریبی آن را رسم کنید و درمورد تغییرات احتمالی که می‌دهید توضیح دهید.

ب) پیشبینی نمایید در این حالت منحنی validation error به چه صورت می‌شود. شکلی تقریبی آنرا رسم کنید و درمورد تغییرات احتمالی که می‌دهید توضیح دهید.

در ادامه فرض کنید ما از یک k-fold cross-validation بجای استفاده از validation set استفاده نماییم.(به هرکدام از بخش هایی که در هربار استفاده از cross-validation به عنوان test set استفاده می‌کنیم یک fold گفته می‌شود.)

ج) با توجه به اینکه با این تکنیک در هر fold از یک test set مستقل استفاده می‌کنیم، آیا منحنی validation Error برای هر fold نیز مستقل می‌شود.

د) (امتیازی) به نظر شما مقدار k مناسب جهت training را چگونه باید انتخاب نماییم.

۵- (امتیازی) در این سوال قصد داریم با استفاده از python برای داده های دو کلاس با استفاده از Logistic regression classifier و Perceptron جداسازی بیابیم.

برای این سوال داده با استفاده از داده های مساله و کد زیر، دو کلاس به صورت تصادفی می‌سازیم.

```
iris = np.genfromtxt("data/iris.txt",delimiter=None)
X, Y = iris[:,0:2], iris[:,-1]   # get first two features & target
X,Y  = ml.shuffleData(X,Y)       # reorder randomly (important later)
X,_  = rescale(X)                # works much better on rescaled data

XA, YA = X[Y<2,:], Y[Y<2]        # get class 0 vs 1
XB, YB = X[Y>0,:], Y[Y>0]        # get class 1 vs 2
```

الف) با استفاده از Logistic regression classifier جداسازی برای دو کلاس بیابید.

ب) با استفاده از الگوریتم Perceptron جداسازی برای دو کلاس بیابید.