

به نام خدا

شبکه‌های عصبی

دانشکده‌ی علوم و مهندسی کامپیوتر

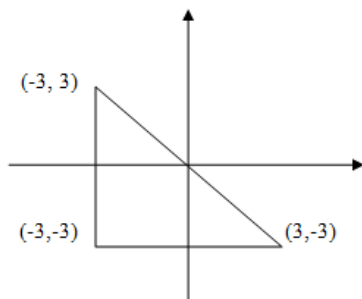


تاریخ تحویل:

تکلیف نخست - زمستان ۱۳۹۳

- قبل از اینکه شروع به حل تمرین‌ها کنید، حتماً یک بار فایل **Policies.pdf** (موجود در وبگاه درس) که حاوی نکات مهم در تحویل تکالیف هست را مطالعه فرمایید.
- طراح هریک از سوالات در زیر مشخص شده است. در صورتی که سوالی دارید می‌توانید سوالتان را در گروه مطرح کنید یا از طراح سوال پرسید.
-طراح سوال ۱ و ۳: خانم صنیعی
-طراح سوال ۲: آقای زندی
-طراح سوال ۴: آقای کاهانی

۱- الف) شکل روبرو را در نظر بگیرید. آیا یک پرسپترون تک لایه، قادر خواهد بود، تمامی نقاط داخل مثلث را از نقاط خارج مثلث جدا کند؟ چرا؟



ب) شبکه‌ای با سه نرون در لایه‌ی مخفی در نظر بگیرید که مقادیر وزن‌ها و بایاس آن به صورت زیر مشخص شده است. تابع انگیزش را **hardlim** در نظر بگیرید.

$$W = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -1 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 3 \\ 0 \end{bmatrix}$$

بررسی کنید این شبکه به ازای نقاط داخل مثلث $(-2, -1)$, $(-2, 1)$ و $(2, -5/2)$ چه مقادیری تولید می‌کند.

ج) بررسی کنید این شبکه به ازای نقاط خارج از مثلث $(-4, 1)$, $(-5, 1)$, $(-5, -1)$ و $(4, -2)$ چه مقادیری تولید می‌کند. آیا برای همه‌ی نقاط خارج از مثلث، مقادیر یکسانی تولید می‌شود؟ اگر جواب شما به این سوال منفی است، چه راهی برای حل این مشکل پیشنهاد می‌کنید؟

۲- یک بازی کامپیوتری تخیلی را در نظر بگیرید (یکی از آن اعتیادآورها!). فرض کنید که در هر لحظه از زمان گول بی‌شاخ و دم بازی یا از جادو استفاده می‌کند (عمل A) یا مشت محکمی بر شما می‌کوبد (عمل B).

الف) اگر احتمال وقوع A و B در هر لحظه از زمان با هم برابر باشد، می‌توان پیش‌بینی کرد که کدام عمل رخ خواهد داد؟

ب) اگر احتمال وقوع A و B وابسته به زمان و با توزیعی غیر از یکنواخت (مثلا گاوسی) باشد یا وابسته به عمل قبلی باشد چطور؟

پ) آنتروپی (یا به عبارتی میزان پیش‌بینی ناپذیری یا عدم قطعیت)، در حالت الف بیشتر است یا ب؟

ت) به طور مشابه (در دنیایی خارج از بازی‌های کامپیوتری)، فرض کنید کل تعداد کلمات یک زبان N است و احتمال وقوع آنها در هر مکانی از جمله مستقل از کلمات قبلی بوده و به علاوه برای همه کلمات نیز یکسان است. (در واقعیت اینطور نیست و احتمال هر کلمه به مکان آن و همچنین کلمات قبلی وابسته است) توضیح دهید که آیا یادگیری گرامر زبان برای انسان در این حالت سخت‌تر است یا آسان‌تر.

ث) آیا هر مسئله‌ای را می‌توان با یادگیری حل نمود؟ به عبارت دیگر آیا مرزی در یادگیری وجود دارد یا خیر؟ به طور دقیق‌تر آیا در برخی مسائل خطا را می‌توان از حد خاصی کمتر کرد؟ توضیح دهید.

راهنمایی: دقت نمائید که این سوال به طور کلی درباره‌ی مبحث یادگیری است و به ابزار خاصی مانند شبکه‌ی عصبی وابسته نیست. به علاوه نیاز به تحلیل دقیق ریاضی نیست ولی تحلیل با فرمول‌های ریاضی فمده‌ی امتیازی خواهد داشت.

۳- فایل Tehran2008.xls حاوی بخشی از اطلاعات آب و هوایی شهر تهران در هر یک از روزهای ۱۲ ماه سال ۲۰۰۸ میلادی است. برای هر روز مشخصات آب و هوایی در چندین ساعت مختلف ثبت شده است.

الف) یکی از کاربردهای شبکه‌ی عصبی، پیش‌بینی مقادیر مختلف در یک سری زمانی است. در این تمرین قصد داریم با نمونه‌ی ساده‌ای از این کاربرد آشنا شویم. ستونی که مربوط به دمای هوا (T: Temperature) است را در نظر بگیرید. قصد داریم با استفاده از داده‌های این ستون، شبکه‌ی عصبی را طراحی کنیم که قادر باشد دما را با استفاده از ۵ دمای ثبت شده‌ی قبلی، پیش‌بینی کند.

- از ۸۰۰ داده‌ی اول به عنوان داده‌ی آموزشی و از ۳۹۸ داده‌ی باقی‌مانده به عنوان داده‌ی تست استفاده کنید (برای سادگی، اطلاعات مربوط به ساعات روزها را نادیده بگیرید. هم‌چنین می‌توانید سطرهایی از جدول که دمای هوا در آن‌ها مشخص نشده است را حذف کنید).

- این مساله را با پرسپترون‌های چند لایه با تعداد لایه‌ها و تعداد نرون‌های مختلف حل کنید. بررسی کنید کدام معماری از نظر میانگین مربعات خطا، بر روی داده‌های تستی نتیجه‌ی بهتری می‌دهد.

- در این تمرین می‌توانید از جعبه ابزار شبکه عصبی متلب استفاده کنید.

۴- فرض کنید مجموعه داده‌ای ۲ بعدی در اختیار داریم. هدف جداسازی کلاس‌های این مجموعه با پرسپترون تک لایه است. در این تمرین برای پیاده سازی، از جعبه ابزار شبکه عصبی متلب استفاده نکنید.

الف) یک شبکه‌ی پرسپترون با یک نرون را پیاده سازی کنید. وزن های شبکه را برای ماتریس نمونه‌ها (X1) و خروجی مطلوب آن (D1)، بدست آورید. تابع انگیزش را hardlim در نظر بگیرید.
- برای هر نرون بردار وزن به صورت یک بردار ۱ در ۳ خواهد بود که مقدار اول آن همان بایاس است.

ب) -اختیاری- کد بخش الف را به پرسپترون تک لایه با دو نرون گسترش دهید. این بار شبکه را دو بار آموزش دهید:

- بار اول با استفاده از ماتریس نمونه‌ها (X2) و خروجی مطلوب (D2).

- بار دوم با ماتریس نمونه‌ها (X2) و خروجی مطلوب (D3).

همگرایی شبکه را در دو حالت بالا بررسی کرده و در صورت همگرا نشدن هریک، علت آن را شرح دهید.

ج) -اختیاری- چرا با این شیوه نمی‌توان مسئله‌ی XOR را حل کرد؟

محتویات فایل Data1.mat:

- ماتریس X1: یک ماتریس ۳ در ۱۲۰ است که ۶۰ نمونه‌ی اول متعلق به کلاس A، ۶۰ نمونه‌ی دوم متعلق به کلاس B است. هر ستون یک نمونه است. و سطر اول هر نمونه مقدار ۱ دارد. (ورودی مربوط به وزن اول، بایاس است).

- ماتریس X2: یک ماتریس ۳ در ۱۲۰ است که ۳۰ نمونه‌ی اول متعلق به کلاس A، ۳۰ نمونه‌ی دوم متعلق به کلاس B، ۳۰ نمونه‌ی سوم و چهارم به ترتیب متعلق به کلاس C و D هستند. هر ستون یک نمونه است. و سطر اول هر نمونه مقدار ۱ دارد. (ورودی مربوط به وزن اول، بایاس است).

- ماتریس D1: یک ماتریس ۱ در ۱۲۰ است که هر ستون، خروجی مطلوب نرون را برای نمونه‌ی متناظر با آن در X1، نشان می‌دهد. خروجی مطلوب برای هر کلاس به صورت زیر در نظر گرفته شده است:

Class A = 0

Class B = 1

- ماتریس D2: یک ماتریس ۲ در ۱۲۰ است که هر ستون، خروجی مطلوب نرون‌ها را برای نمونه‌ی متناظر با آن در X2، نشان می‌دهد. خروجی مطلوب برای هر کلاس به صورت زیر در نظر گرفته شده است:

Class A = [1 0]

Class B = [0 1]

Class C = [1 1]

Class D = [0 0]

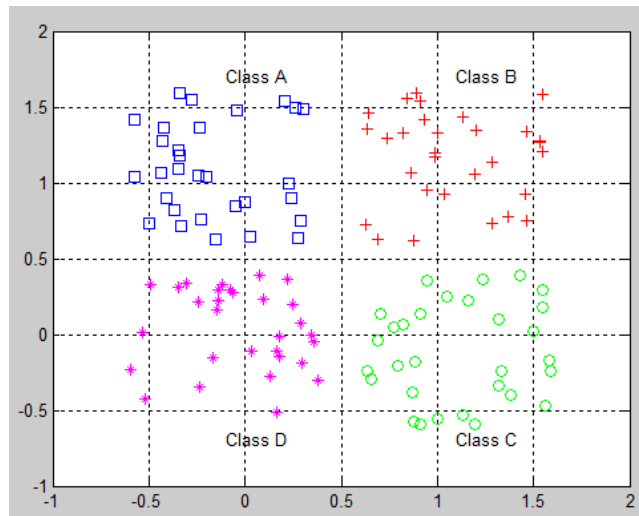
- ماتریس D3: مشابه ماتریس D1، با تغییر خروجی مطلوب برای هر کلاس:

Class A = [0 1]

Class B = [1 1]

Class C = [1 0]

Class D = [0 0]



شکل ۱- نمایش داده‌های مرتبط با سوال ۴

موفق باشید: