



پردیس دانشکده های فنی

به نام خدا  
دانشکده مهندسی برق و کامپیوتر  
تمرین سری سوم یادگیری  
ماشین  
پاییز ۱۴۰۱



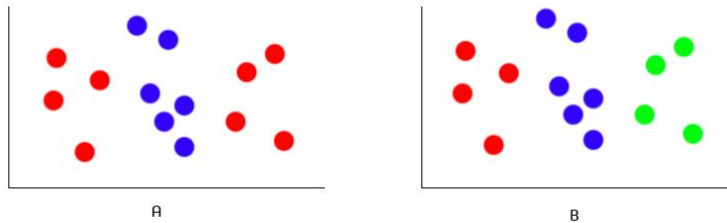
دانشگاه تهران

سلام بر دانشجویان عزیز، چند نکته مهم:

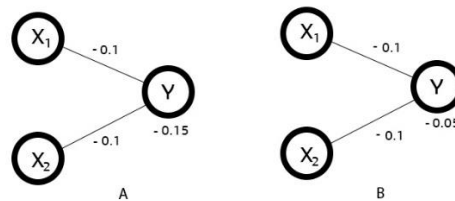
۱. حجم گزارش به هیچ عنوان معیار نمره دهی نیست، در حد نیاز توضیح دهید.
۲. نکته ی مهم در گزارش نویسی روشن بودن پاسخ ها می باشد، اگر فرضی برای حل سوال استفاده می کنید حتما آن را ذکر کنید، اگر جواب نهایی عددی است به صورت واضح آن را بیان کنید.
۳. کدهای ارسال شده بدون گزارش فاقد نمره می باشند.
۴. برای سوالات شبیه سازی، فقط از دیتاست داده شده استفاده کنید.
۵. فایل نهایی خود را در یک فایل زیپ شامل، pdf گزارش و فایل کدها آپلود کنید. نام فایل زیپ ارسالی الگوی [Student ID Number]\_ML\_HW3 داشته باشد.
۶. از بین سوالات تئوری حتماً حداقل به ۶۰ نمره، و از بین سوالات پیاده سازی حداقل به ۴۰ نمره پاسخ دهید.
۷. نمره تمرین ۱۰۰ نمره می باشد و حداکثر تا نمره ۱۱۰ ( ۱۰ نمره امتیازی) می توانید کسب کنید.
۸. هرگونه شباهت در گزارش و کد مربوط به شبیه سازی، به منزله تقلب می باشد و کل تمرین برای طرفین صفر خواهد شد.
۹. مهلت این تکلیف تا ۲۱ آذر ماه است و به هیچ عنوان قابلیت تمدید ندارد.
۱۰. در صورت داشتن سوال، از طریق ایمیل [sara.jahedazad@gmail.com](mailto:sara.jahedazad@gmail.com) سوال خود را مطرح کنید.

سوال ۱: (۱۵ نمره)

به سوالات زیر در ارتباط با مفاهیم اولیه‌ی شبکه‌های عصبی پاسخ دهید.  
الف) برای طبقه‌بندی هر یک از داده‌های زیر یک ساختار شبکه‌ی عصبی **mlp** ارائه‌دهید و علت انتخاب خود را بیان کنید. ( ساده ترین مدل ممکن را پیشنهاد دهید. نیاز و یا عدم نیاز به لایه‌ی مخفی و تعداد نورون‌های مورد نیاز در خروجی را ذکر کنید).



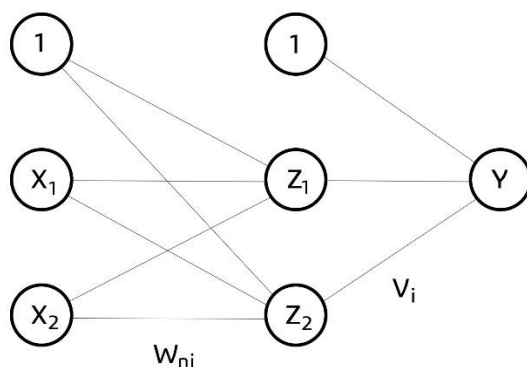
ب) هر کدام از شبکه‌های زیر نشانگر کدام گیت منطقی است؟ با محاسبه‌ی خروجی به ازای مقادیر باینری مختلف ورودی بیان کنید. آستانه‌ی تحریک در قسمت پایینی نورون درج شده است.



ج) یک شبکه‌ی عصبی ارائه دهید که گیت **XNOR** را نمایش دهد. درستی عملکرد شبکه‌ی خود را با حساب خروجی به ازای مقادیر ورودی مختلف نشان دهید. (مقدار وزن‌ها و آستانه‌ی تحریک را ذکر کنید. در این سوال مجاز به اضافه کردن بایاس نیز هستید.)

سوال ۲: (۱۵ نمره)

برای شبکه‌ی عصبی پایین، یک مرحله، با کمک پس‌انتشار خطا، وزن‌ها را به روز رسانی نمایید. ابتدا تمام مراحل را به صورت پارامتری بنویسید، سپس عددی حل کنید.



ورودی:  $x_1 = 1, x_2 = 0$

خروجی مطلوب: صفر

تابع فعالساز لایه‌ی مخفی: سیگموئید

تابع فعالساز خروجی: همانی

نرخ یادگیری: دو دهم

تابع خطا:

$$0.5(predicted - target)^2$$

$$\begin{aligned} W_{01} &= 0.4, W_{02} = 0.6, W_{11} = 0.7, \\ W_{12} &= -0.4, W_{21} = -0.2, W_{22} = 0.3 \\ V_0 &= -0.3, V_1 = 0.5, V_2 = 0.1 \end{aligned}$$

سوال ۳: (۱۵ نمره)

شخصی قصد دارد لباسی خریداری کند. تعدادی لباس با ویژگی‌هایی مختلف در اختیار او قرار می‌گیرند و او تصمیم می‌گیرد که آیا تمایل به خرید آن‌ها دارد یا خیر. اطلاعات مربوط به انتخاب‌های او در جدول زیر قابل مشاهده است. با انجام محاسبات برای معیار آنتروپی، کدام ویژگی برای تعیین این که این شخص به خرید یک لباس تمایل دارد یا نه ملاک مناسب تری است؟ با توجه به این ویژگی یک درخت تصمیم به ارتفاع یک رسم کنید. دقت و صحت را گزارش نمایید.

قیمت	جنس	رنگ	تمایل به خرید
گران	پشمی	آبی	خیر
گران	پشمی	آبی	بله
گران	پشمی	سبز	خیر
گران	نخی	قرمز	بله
ارزان	نخی	قرمز	بله
گران	نخی	آبی	خیر
گران	پشمی	قرمز	خیر
ارزان	پشمی	آبی	خیر
ارزان	پشمی	سبز	خیر
ارزان	نخی	قرمز	بله

سوال ۴: (۱۵ نمره)

فرض کنید داریم  $p(x) \sim U(0, a)$  که  $U$  توزیع یونیفرم در بازه‌ی ۰ تا  $a$  است. (۱۵ نمره)

پنجره‌ی پارزن به صورت زیر تعریف می‌گردد:

$$\varphi(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

الف) نشان دهید برای میانگین تخمین پارزن داریم:

$$\bar{P}_n(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{a} \left( 1 - e^{-\left(\frac{x}{h_n}\right)} \right) & 0 \leq x \leq a \\ \frac{1}{a} \left( e^{\frac{a}{h_n}} - 1 \right) e^{-\frac{x}{h_n}} & a \leq x \end{cases}$$

ب)  $h_n$  چه قدر باید کوچک باشد تا اندازه‌ی بایاس تخمین پارزن، در نود و نه درصد از طول بازه‌ی 0 تا a، کمتر یک درصد باشد؟

سوال ۵: (۱۵ نمره)

فرض کنید  $D = \{x_1, x_2, x_3, \dots, x_n\}$  مجموعه‌ای از نقاط مستقل دارای برچسب باشد، و  $D_k(x) = \{x'_1, x'_2, \dots, x'_k\}$  مجموعه‌ی k نزدیک‌ترین همسایه به x باشد. در یک مسئله‌ی دو کلاسه فرض می‌کنیم  $p(\omega_1) = p(\omega_2) = \frac{1}{2}$  است. همچنین، احتمالات شرطی  $p(x|\omega_i)$  در کره‌هایی به شعاع واحد یکنواخت هستند. فاصله‌ی این دو کره از یک دیگر ده واحد در نظر گرفته می‌شود. با داشتن این فرض‌ها به سوالات زیر پاسخ دهید: (۱۵ نمره)

الف) نشان دهید که در صورت فرد بودن k، میانگین احتمال خطا به صورت زیر خواهد بود:

$$P_n(e) = \frac{1}{2^n} \sum_{j=0}^{\frac{k-1}{2}} \binom{n}{j}$$

ب) نشان دهید برای  $k > 1$  نرخ خطای تک همسایگی از نرخ خطای k همسایگی کمتر است.

ج) اگر k اجازه داشته باشد که با n افزایش یابد ولی دارای شرط  $k < a/\sqrt{n}$  باشد نشان دهید که با میل کردن n به بینهایت  $P_n(e)$  به صفر همگرا می‌گردد.

سوال ۶: (۳۰ نمره) (پیاده‌سازی)

در این سوال شما مجاز به استفاده از کتابخانه‌های آماده (از جمله Keras و یا هر کتابخانه‌ی دیگر) هستید. هدف این سوال به کارگیری یک CNN است. در این سوال تمامی پارامترها و هایپرپارامترها به اختیار شما انتخاب می‌شوند. نیاز است تمامی فرض‌ها را ذکر نمایید. پیشنهاد می‌گردد از محیط google colab استفاده کنید. دیتاست CIFAR-10 در این سوال به کار گرفته می‌شود. شما می‌توانید این دیتاست را به هر روشی لود کنید (از جمله استفاده از دیتاست کراس).

الف) این موارد را در گزارش خود ذکر کنید:

- دیتاست را به سه بخش آموزش، تست و سنجش تقسیم کنید. درصد اختصاص داده به هر یک از این موارد را ذکر نمایید.

- معماری شبکه‌ای که به کار گرفته‌اید را در یک جدول ارائه دهید. (تعداد لایه‌های convolutional و

dense، تعداد نورون‌ها در هر لایه، اندازه‌ی فیلترها، تابع فعالساز و ...)

- تابع هزینه، تعداد نورون و تابع فعالساز لایه‌ی آخر (چرا این‌ها را انتخاب کردید؟)

- نرخ یادگیری

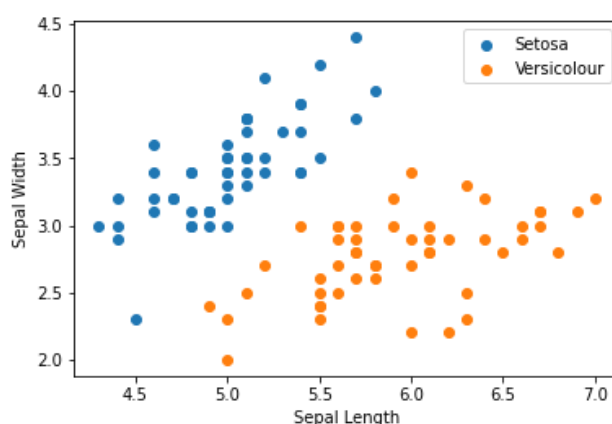
- بهینه‌ساز

ب) شبکه را آموزش دهید.

- نمودار loss و accuracy را بر حسب ایپاک برای داده‌ی آموزش و سنجش رسم کنید.

- داده‌ی تست را به شبکه دهید و ماتریس آشفتگی را برای پیش‌بینی شبکه رسم کنید.

سوال ۷: (۲۰ نمره) (پیاده‌سازی) (در این سوال مجاز به استفاده از دستورهای آماده برای پرسپترون نیستید).  
درباره‌ی دیتاست Iris (زنبق) مطالعه کنید. می‌توانید از [این مسیر](#) و یا هر روش دیگری دیتاست را لود کنید.  
در یک تصویر، برای داده‌های دو گونه‌ی setosa و versicolor، طول کاسبرگ در مقابل عرض کاسبرگ را رسم کنید. تصویر شما باید چیزی مانند شکل زیر باشد:



با کمک پیاده‌سازی الگوریتم پرسپترون و به کارگیری ویژگی‌های طول و عرض کاسبرگ، خطی پیدا کنید که دو گونه‌ی گل زنبق یادشده را از یک دیگر جدا کند. شیب و عرض از مبدأ را گزارش کرده و خط پیدا شده را با داده‌ها در یک تصویر رسم نمایید.

سوال ۸: (۲۰ نمره) (پیاده‌سازی) (در این سوال مجاز به استفاده از دستورهای آماده نیستید).

ابتدا با کمک دستور زیر داده تولید کنید:

```
import numpy as np
N = 1000
np.random.seed(1)
X = np.concatenate((np.random.normal(0, 1, int(0.3 * N)),
np.random.normal(5, 1, int(0.7 * N))))[:, np.newaxis]
```

توزیع داده‌ها را رسم نمایید. برای این کار شما نیاز به پیاده‌سازی پنجره‌ی پارزن دارید.  
برای کرنل گوسی تاثیر اندازه‌ی پنجره‌ی پارزن را روی توزیع تخمین زده شده بررسی کنید و توضیح دهید.  
(حداقل سه اندازه‌ی مختلف؛ مثلاً 0.1، 1 و 10)