



هدف این تمرین آشنایی با پیاده‌سازی شبکه‌های عصبی تک لایه و چند لایه با توابع فعال‌سازی مختلف، نحوه به کارگیری شبکه‌ها برای مسائل دسته‌بندی، پرسپترون و قانون یادگیری در شبکه‌های پرسپترون تک لایه و نیز پیاده‌سازی توابع خطای مختلف است. در بخش‌های 1 تا 4 پیاده‌سازی شبکه عصبی تنها با استفاده از کتابخانه Numpy خواهد بود و **تنها در قسمت آخر امکان استفاده از کتابخانه torch وجود دارد.**

مهلت تحویل این تمرین پایان روز سه شنبه 30 آبان 1402 خواهد بود.

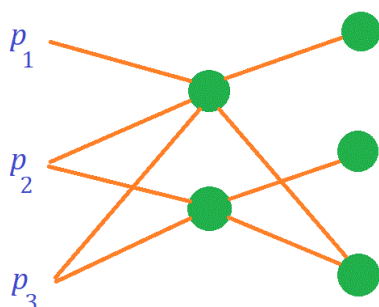
## 1 – معماری شبکه و توابع فعال‌سازی (Activation Functions)

**1-1:** یک شبکه عصبی تک لایه با 4 ورودی و 6 خروجی و تابع فعال‌سازی ReLU تشکیل دهید. همه پارامترهای این شبکه را به صورت تصادفی بین -1 تا 1 مقداردهی کنید. سپس داده زیر را به عنوان ورودی به آن بدهید و خروجی را به همراه همه پارامترها گزارش کنید.

$$\mathbf{p} = [1 \quad -2 \quad 5 \quad 1]^T$$

**1-2:** شبکه عصبی زیر را که دارای 2 لایه و 3 ورودی است، تشکیل دهید (دقت کنید که این شبکه fully connected نیست یعنی بعضی از وزن‌های آن حذف شده است). وزن‌ها و بایاس‌ها را به صورت تصادفی با توزیع یکنواخت بین -1 تا 1

مقداردهی کنید و خروجی را برای دو داده  $\begin{bmatrix} 4 \\ -1 \\ -2 \end{bmatrix}$  و  $\begin{bmatrix} -3 \\ 5 \\ 1 \end{bmatrix}$  به دست آورید.



**1-3:** توابع فعال‌سازی GeLU و Leaky ReLU را که روابطی به صورت زیر دارند، پیاده‌سازی کنید. سپس خروجی این توابع را در نقاط -0/5 و 0/5 محاسبه کنید (در Leaky ReLU مقدار  $a$  را برابر 0/01 در نظر بگیرید). با جستجو در اینترنت شکل این توابع را در گزارش خود بیاورید و بررسی کنید در چه مواقعی از این دو تابع به جای ReLU استفاده می‌شود.

GeLU	Leaky ReLU
$0.5x \left( 1 + \tanh \left[ \sqrt{2/\pi} (x + 0.044715x^3) \right] \right)$	$f(x) = \begin{cases} x & x \geq 0 \\ ax & x < 0 \end{cases}$

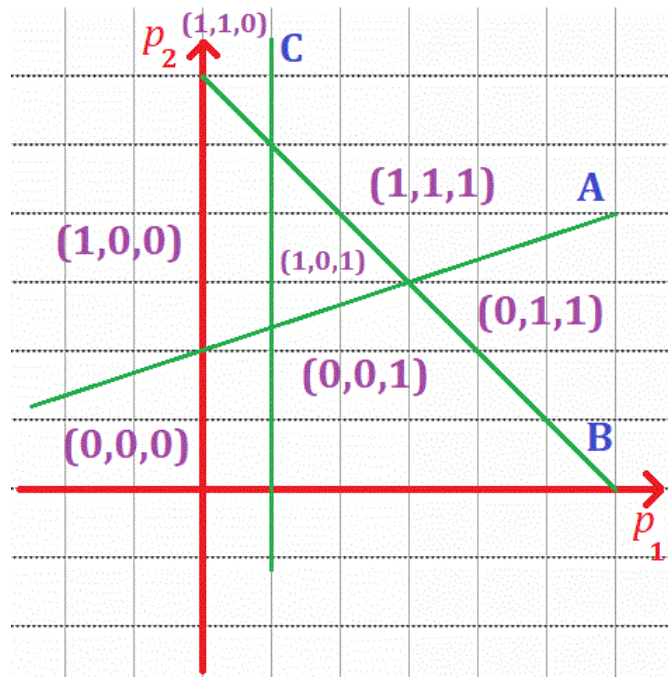
**1-4:** یک شبکه عصبی دو لایه با 3 ورودی و 4 خروجی طراحی کنید که در لایه مخفی آن از تابع فعال‌سازی Linear و در لایه خروجی آن از تابع فعال‌سازی tanh استفاده شده است. همه وزن‌ها را به صورت تصادفی و با توزیع نرمال با میانگین صفر و انحراف معیار 1 و همه بایاس‌ها را برابر صفر مقداردهی کنید. تعداد نورون‌ها را در لایه مخفی 12 عدد در نظر بگیرید. به نظر شما اضافه کردن یک لایه مخفی دیگر به این شبکه با تابع فعال‌سازی Linear چه تاثیری بر عملکرد آن دارد؟ اضافه کردن یک لایه مخفی با تابع ReLU چگونه؟

## 2 – دسته‌بندی (Classification)

**2-1:** در شکل زیر یک شبکه پرسپترونی تک لایه با تابع فعال‌سازی Unit Step را مشاهده می‌کنید که توسط سه نورون A و B و C فضای ورودی را به 7 بخش تقسیم کرده است. پранتزه‌های بنفش رنگ، برچسب هر بخش را نشان می‌دهند. بیت اول (پر ارزش) هر بخش مربوط به نورون A، بیت دوم مربوط به نورون B و بیت سوم (کم ارزش) مربوط به نورون C می‌باشد. الف) معادله مرز تصمیم هر یک از نورون‌های A و B و C را به دست آورید. ب) وزن‌ها و بایاس‌های هر نورون را گزارش کنید.

پ) اگر یک داده به صورت  $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$  وارد شود، از روی نمودار واضح است که خروجی شبکه برای آن به صورت (0,0,1) خواهد بود، این موضوع را به صورت عملی آزمایش کنید (با اعمال آن به شبکه).

ت) با توجه به این که هر خط صفحه را به 2 قسمت تقسیم می‌کند، بنابراین سه خط باید صفحه را به  $2 \times 2 \times 2 = 8$  قسمت تقسیم کند. چرا در شکل زیر پس از افزایش صفحه فقط 7 بخش ایجاد شده است؟ در واقع چرا خروجی شبکه به ازای هیچ ورودی (0,1,0) نشده است؟



**2-2:** یک شبکه عصبی fully connected با یک لایه مخفی و سه ورودی را در نظر بگیرید. می‌خواهیم از این شبکه برای دسته‌بندی دو دسته‌ای استفاده کنیم. در لایه خروجی این شبکه از یک نورون با تابع فعال‌سازی Sigmoid بهره می‌بریم که خروجی احتمالاتی برای ما ایجاد می‌کند. در هر یک از حالت‌های زیر یک batch متشکل از چهار داده  $\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$ ،  $\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$ ،  $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$  را به شبکه اعمال کنید و خروجی شبکه و نیز دسته منتسب به هر داده را تعیین کنید (تمام پارامترهای شبکه به صورت تصادفی با توزیع نرمال با میانگین 0 و انحراف معیار 1 باید مقداردهی شوند و در تمام حالت‌های زیر از همان پارامترها استفاده شود).

(الف) تابع فعال‌سازی در لایه مخفی برای همه نورون‌ها ReLU باشد.

(ب) تابع فعال‌سازی در لایه مخفی برای همه نورون‌ها tanh باشد.

(پ) تابع فعال‌سازی در لایه مخفی برای همه نورون‌ها Sigmoid باشد.

(ث) تابع فعال‌سازی در لایه مخفی برای همه نورون‌ها Linear باشد.

(ج) اگر برچسب دو داده اول 0 و دو داده دوم 1 باشد، Binary Cross-Entropy را محاسبه کنید.

**3-2:** فرض کنید برای یک مسئله دسته‌بندی سه دسته‌ای از یک شبکه دو لایه با 4 نورون و تابع فعال‌سازی ReLU در لایه مخفی و تابع Softmax در لایه خروجی استفاده شده است. پارامترهای شبکه را به صورت تصادفی با توزیع یکنواخت بین -4 تا +4 مقداردهی و گزارش کنید. خروجی شبکه و نیز دقت دسته‌بندی شبکه بر روی داده‌های زیر را مشخص کنید همچنین خطای خطای Categorical Cross-Entropy را گزارش کنید.

$$\text{Category I: } \left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\}, \text{Category II: } \left\{ \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\} \quad \text{Category III: } \left\{ \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$$

### 3 – قانون یادگیری پرسپترون (Perceptron Learning Rule)

**3-1:** مجموعه داده آموزشی زیر را برای یک مسئله دسته‌بندی دو دسته‌ای در نظر بگیرید:

$$\left\{ \mathbf{p}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, t_1 = 0 \right\} \left\{ \mathbf{p}_2 = \begin{bmatrix} 1 \\ 5 \end{bmatrix}, t_2 = 0 \right\} \left\{ \mathbf{p}_3 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, t_3 = 0 \right\} \left\{ \mathbf{p}_4 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}, t_4 = 0 \right\}$$

$$\left\{ \mathbf{p}_5 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, t_5 = 1 \right\} \left\{ \mathbf{p}_6 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, t_6 = 1 \right\} \left\{ \mathbf{p}_7 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, t_7 = 1 \right\} \left\{ \mathbf{p}_8 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, t_8 = 1 \right\}$$

یک شبکه پرسپترون تک لایه مناسب طراحی کنید که توانایی دسته‌بندی داده‌های بالا را داشته باشد. تمام وزن‌ها و بایاس‌های این شبکه را با عدد 10 مقداردهی اولیه کنید. این شبکه را به تعداد 5 epoch بر روی داده‌های آموزشی با استفاده از قانون یادگیری پرسپترون آموزش دهید و در انتهای هر epoch، دقت آموزش و تمام پارامترها را گزارش کنید. چند epoch برای رسیدن به دقت 100 درصد بر روی مجموعه آموزش نیاز است؟

**3-2:** داده‌های آموزشی زیر را در نظر بگیرید:

$$\text{Category I: } \left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\}, \text{Category II: } \left\{ \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$$

$$\text{Category III: } \left\{ \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}, \text{Category IV: } \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$$

یک شبکه پرسپترون تک لایه مناسب طراحی کنید که توانایی دسته‌بندی داده‌های بالا را داشته باشد. آن را به حدی آموزش دهید که دقت آموزش به 100 درصد برسد. تعداد epoch های مورد نیاز و پارامترهای نهایی مدل را گزارش کنید. سپس داده‌های آزمایشی زیر را به شبکه آموزش داده شده اعمال کنید و خروجی‌ها را گزارش کنید.

$$\mathbf{p}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{p}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{p}_3 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \mathbf{p}_4 = \begin{bmatrix} -2 \\ -5 \end{bmatrix}$$

**3-3:** فرض کنید که در یک مسئله دسته‌بندی سه دسته‌ای توزیع برچسب داده‌های آموزشی به صورت زیر است:

$$y = \begin{cases} 0 & p_1^2 < p_2 - 2 \\ 1 & -p_1^2 > p_2 + 2 \\ 2 & \text{Otherwise} \end{cases}$$

که  $\mathcal{Y}$  برچسب واقعی داده‌ها و  $p_1$  و  $p_2$  به ترتیب ویژگی (feature) اول و دوم هر داده را نشان می‌دهد.  
الف) آیا این داده‌ها جدایی‌پذیر خطی هستند؟ چرا؟

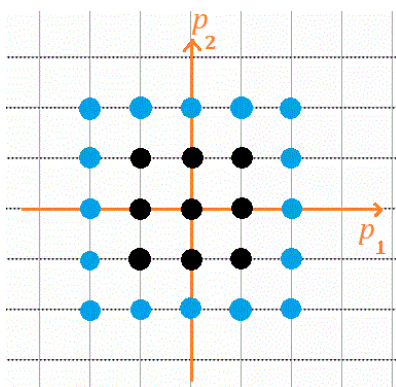
ب) به صورت تصادفی از هر دسته 100 داده تولید کنید و مجموعه 300 داده حاصل را مجموعه آموزشی بنامید.

پ) یک شبکه پرسپترونیک تک لایه با تعداد نورون مناسب را طراحی کنید و آن را به منظور دسته‌بندی مجموعه آموزشی تولید شده آموزش دهید. فرآیند آموزش را به تعداد 10 epoch تکرار کنید.

ت) وزن‌ها و بایاس‌های نهایی و نیز دقت شبکه بر روی مجموعه آموزشی را گزارش کنید.

ث) آیا ممکن است که داده‌ای توسط شبکه در دسته 3 (دسته‌ای که اصلاً وجود ندارد) دسته‌بندی شود؟ در صورت وجود، سعی کنید که با سعی و خطا حداقل یک داده با خروجی 3 را پیدا کنید و علت را تحلیل کنید.

3-4: داده‌های آموزشی زیر را در نظر بگیرید (مقیاس نمودار واحد است):



نقاط آبی رنگ: برچسب 0

نقاط سیاه رنگ: برچسب 1

الف) یک شبکه پرسپترونیک تک لایه مناسب را به منظور دسته‌بندی این داده‌گان تشکیل دهید و آن را به تعداد 200 epoch با استفاده از قانون یادگیری پرسپترون آموزش دهید. مقدار اولیه همه وزن‌ها و بایاس‌ها را به ترتیب 0/1 و 0 در نظر بگیرید.

ب) پس از پایان فرآیند یادگیری، ماتریس وزن و بردار بایاس نهایی و نیز دقت آموزش (train) را برای این شبکه به دست آورید.

پ) اگر در مرحله آزمایش (test) دو نمونه به صورت زیر وارد شود، شبکه چه برچسبی را برای هر کدام پیش‌بینی می‌کند؟

$$\mathbf{p}_1 = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \quad t_1 = 0, \quad \mathbf{p}_2 = \begin{bmatrix} -1 \\ 0.5 \end{bmatrix} \quad t_2 = 1$$

## 4- رگرسیون (Regression)

1-4: برای یک شبکه دو لایه با 3 ورودی و یک خروجی خطای MSE را بر روی مجموعه داده زیر محاسبه کنید (همه

پارامترهای شبکه را در لایه اول 1 و در لایه دوم -1 در نظر بگیرید):

$$\mathbf{p}_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \quad t_1 = 0.4, \quad \mathbf{p}_2 = \begin{bmatrix} -4 \\ 3 \\ -3 \end{bmatrix} \quad t_2 = -1.6, \quad \mathbf{p}_3 = \begin{bmatrix} 1.7 \\ -1 \\ 0 \end{bmatrix} \quad t_3 = 3.7, \quad \mathbf{p}_4 = \begin{bmatrix} -2 \\ -3 \\ 1.5 \end{bmatrix} \quad t_4 = -0.9$$

## 5 – پایتورچ (Pytorch)

**1-5:** سوال 2-2 را با استفاده از کتابخانه پایتورچ انجام دهید. وزن‌هایی که به لایه‌های شبکه داده شده را استخراج کرده و به عنوان وزن‌های اولیه به لایه‌های تعریف شده در کتابخانه پایتورچ دهید و خطای Cross-Entropy را محاسبه کنید و حاصل را با قسمت 2-2 مقایسه کنید.

**2-5:** سوال 2-3 را با استفاده از کتابخانه پایتورچ انجام دهید. همانند سوال قبل وزن‌های لایه‌های شبکه را استخراج کنید و به عنوان وزن‌های اولیه به لایه‌های تعریف شده در کتابخانه پایتورچ دهید. انتظار می‌رود که این دو خروجی یکسان باشند. در صورت عدم تطابق، علت را شرح دهید.

**2-5:** سوال 1-4 را با استفاده از کتابخانه پایتورچ انجام دهید. همانند سوال قبل وزن‌های لایه‌های شبکه را استخراج کنید و به عنوان وزن‌های اولیه به لایه‌های تعریف شده در کتابخانه پایتورچ دهید. انتظار می‌رود که این دو خروجی یکسان باشند. در صورت عدم تطابق، علت را شرح دهید.

### ❖ مواردی که باید تحویل داده شود:

- فایل (یا فایل‌های) کد برنامه مورد استفاده برای پیاده‌سازی تمرین با فرمت ipynb. در یک پوشه به نام Code
- به استثنای بخش 5، استفاده از هر گونه کتابخانه آماده (به جز Numpy و Matplotlib) ممنوع است و در صورت استفاده هیچ نمره‌ای به تمرین تعلق نمی‌گیرد.
- فایل گزارش با نام Doc.pdf

فایل‌های کد و گزارش را به صورت یک فایل فشرده در قالب zip و با نام NN\_Name\_Family\_HW1 در سایت کوئرا بارگذاری کنید (به جای Name نام و به جای Family نام خانوادگی خود را قرار دهید).

**مهلت تحویل این تمرین تا پایان روز سه‌شنبه 30 آبان خواهد بود.**

موفق باشید