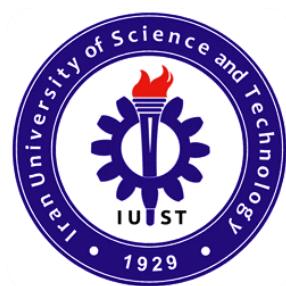


به نام خدا



دانشکده مهندسی کامپیوتر

درس هوش محاسباتی

دکتر ناصر مزینی

تمرین سری دوم

دستیار آموزشی:

ملیکا محمدی فخار

تاریخ تحویل:

۱۴۰۳/۰۸/۱۸

## نکات تکمیلی:

۱. پاسخ سؤالات را به صورت کامل در یک فایل PDF و کدهای سؤالات در فرمت ipynb. (همراه با خروجی سلول های نوت بوک) در یک فایل فشرده به شکل HW2\_StudentID.zip قرار داده و تا زمان تعیین شده بارگذاری نمایید.
۲. در مجموع تمام پروژه ها، ۷۲ ساعت تاخیر در ارسال پاسخ ها مجاز است و پس از آن به ازای هر روز تاخیر ۱۰ درصد از نمره کسر می گردد.
۳. چنانچه دانشجویی پروژه را زودتر از موعد ارسال کند و ۷۵ درصد از نمره را کسب کند، تا سقف ۲۴ ساعت به ساعات مجاز تاخیر دانشجو اضافه می گردد.
۴. حداکثر تأخیر برای ارسال هر تمرین ۴ روز می باشد و بعد از ۴ روز امکان ارسال وجود ندارد.
۵. لازم به ذکر است توضیح کدها و نتایج حاصله باید در فایل PDF آورده شوند و از کد بدون گزارش نمره کسر می شود.
۶. ارزیابی تمرین بر اساس صحیح بودن راه حل، گزارش کامل و دقیق است.
۷. تمرین ها باید به صورت انفرادی انجام شوند و حل گروهی تمرین مجاز نیست.
۸. لطفاً برای انجام تمرین زمان مناسب اختصاص دهید و انجام آن را به روزهای پایانی موکول نکنید.
۹. سؤالات خود را می توانید در گروه مربوطه مطرح نمایید.

## سؤال ۱:

تحلیل خود را در مورد سؤالات زیر ارائه دهید: (۱۵ نمره)

- (a) توابع فعال‌سازی در شبکه‌های MLP چه نقشی دارند؟
- (b) آیا هر تابع غیرخطی می‌تواند به‌عنوان تابع فعال‌سازی در شبکه استفاده شود؟ توضیح دهید.
- (c) افزودن لایه‌های بیشتر به شبکه MLP چه مزایا و معایبی دارد؟
- (d) توابع فعال‌سازی زیر را به‌اختصار مقایسه کرده و مزایا و معایب هر یک را توضیح دهید:

- تابع Sigmoid
- تابع Softmax
- تابع ReLU
- تابع Tanh

## سؤال ۲:

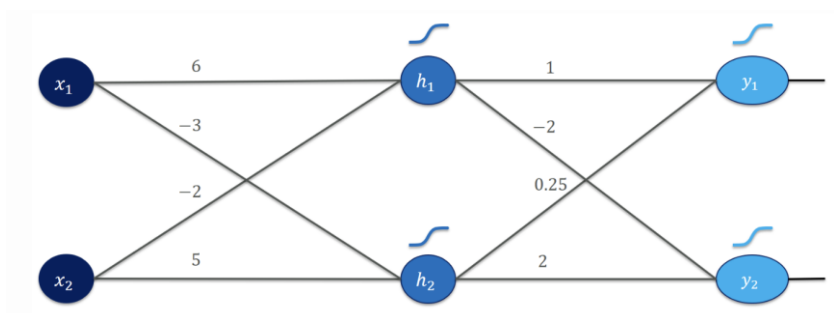
برای حل یک مسئله دسته‌بندی دودویی از یک شبکه MLP استفاده کرده‌ایم. این شبکه از یک آستانه (threshold) ۰.۵ استفاده می‌کند، به این صورت که خروجی‌های بزرگ‌تر یا مساوی ۰.۵ به‌عنوان کلاس ۱ و خروجی‌های کمتر از ۰.۵ به‌عنوان کلاس ۰ دسته‌بندی می‌شوند. خروجی نهای Z است و خروجی شبکه MLP به شکل زیر محاسبه می‌شود:

$$y = \text{Sigmoid}(\text{ReLU}(z))$$

با در نظر گرفتن این ساختار (توابع فعال‌سازی ReLU و Sigmoid و آستانه ۰.۵) چه مشکلاتی را در عملکرد شبکه پیش‌بینی می‌کنید؟ (۱۵ نمره)

### سؤال ۳:

شبکه عصبی نمایش داده شده در شکل زیر را در نظر بگیرید که وزن های هر نورون مشخص شده اند و تابع فعال سازی آن sigmoid است.



اگر داده های آموزشی به شکل زیر داده شود، مراحل محاسبات مربوط به فرایند آموزش مدل (شامل مرحله feed forward و backpropagation) را به صورت گام به گام توضیح دهید. در این سوال فرض شده است که بایاس نادیده گرفته شده است. (۲۰ نمره)

Training data	Inputs		Targets	
	$x_1$	$x_2$	$t_1$	$t_2$
	3	1	1	0
	-1	4	0	1

### سؤال ۴:

با مراجعه به نوت بوک پیوست شده و مطالعه توضیحات هر بخش، توابع مشخص شده را کامل کنید و سپس به سؤالات زیر پاسخ دهید: (۲۵ نمره)

(a) اگر در ابتدای کار شبکه عصبی، به جای مقداردهی تصادفی به وزن ها، همه وزن ها را برابر با صفر قرار دهیم، چه اتفاقی می افتد؟

(b) آیا در الگوریتم های گرادیان کاهشی، گرادیان همواره کاهش می یابد؟

(c) چه زمانی می گوئیم شبکه دچار حالت های underfit یا overfit شده است؟ چه روش هایی را برای رفع مشکلات overfit و underfit پیشنهاد می دهید؟ هر یک را توضیح دهید.

(d) باتوجه به نمودار دقت و تابع ضرر، مدل را از نظر underfit، overfit بررسی کنید

## سؤال ۵:

دیتاست Fashion-MNIST شامل تصاویر لباس‌ها در ۱۰ دسته مختلف است. با استفاده از Keras ، یک مدل MLP طراحی کنید که بتواند بر روی مسئله دسته‌بندی Fashion-MNIST به دقتی حداقل ۸۵ درصد دست یابد. (۲۵ نمره)

- مدل شما باید حداقل ۲ لایه‌ی مخفی داشته باشد. تعداد نوروها در هر لایه‌ی مخفی را به دلخواه انتخاب کنید. توابع فعال‌سازی و سایر پارامترهای مدل را متناسب با مسئله تعیین کنید.
- از روش‌های آموخته‌شده در درس برای جلوگیری از overfitting استفاده کنید.
- پارامترهای شبکه، از جمله تعداد لایه‌ها، تعداد نوروها، توابع فعال‌سازی، تابع ضرر و دلایل انتخاب هر کدام را توضیح دهید.
- در پایان، نمودارهای Loss و Accuracy را برای داده‌های آموزشی و تست رسم کرده و نتایج را تحلیل کنید.
- همچنین بررسی کنید که با استفاده از توابع فعال‌سازی متفاوت (مانند tanh یا sigmoid) چه تغییراتی رخ می‌دهد. نتایج را با تابع فعال‌سازی اصلی (مانند ReLU) مقایسه کرده و دلیل تفاوت‌ها را توضیح دهید.

موفق باشید.