



تمرین سری دوم
درس یادگیری عمیق

نام مدرس: دکتر محمدی
دستیاران آموزشی مرتبط: آقای احمدی – آقای براتی
مهلت تحویل: ۱۴۰۱/۱۲/۲۷

۱- در این سوال قصد داریم که مفاهیم مطرح شده درباره MLP را مرور کنیم. لطفا سوالاتی که در ادامه آورده می شود را به طور کامل توضیح دهید.

الف) دلیل استفاده از توابع فعال سازی در شبکه های MLP چیست؟

ب) فرض کنید مقادیر یکی از لایه های یک شبکه عصبی به صورت زیر باشد. در صورتی که برای منظم سازی از dropout استفاده شده باشد و ماسک dropout به صورت زیر باشد مقادیر نهایی این لایه بعد از اعمال dropout را در مرحله آموزش و آزمون مشخص کنید:

- Dropout rate = 0.5
- Output : [1.5, 6, 3, -0.75, -4.5]
- Mask : [1,0,1,0,1]

ج) فرض کنید یک شبکه ساده با یک لایه مخفی که دارای دو نورون است را برای طبقه بندی تصاویر سگ از گربه استفاده می کنیم. تاثیر اضافه کردن تعداد نورون در هر یک از حالات گفته شده چیست؟ اضافه کردن تعداد نورون های لایه مخفی به تعداد ۲ نورون، ۲۵۶ نورون و ۱ میلیون نورون

د) فرض کنید می خواهیم یک شبکه MLP را برای یک مسئله رگرسیون بکار بگیریم. درحالی که شبکه MLP در تلاش است که همگرا شود اما مقدار training loss نوسان دارد و به آرامی کاهش نمی یابد.

شما احتمال می دهید که شبکه MLP دچار مشکل ناپدید شدن گرادیان است. توضیح دهید که چرا این مشکل در MLP ها ممکن است رخ دهد و راه حل های ممکن برای رفع این مشکل را پیشنهاد کنید. برای مطالعه می توانید به این [لینک](#) مراجعه کنید.

۲- درست یا غلط بودن گزاره های زیر را مشخص کنید و دلیل پاسخ خود را نیز بیان کنید.

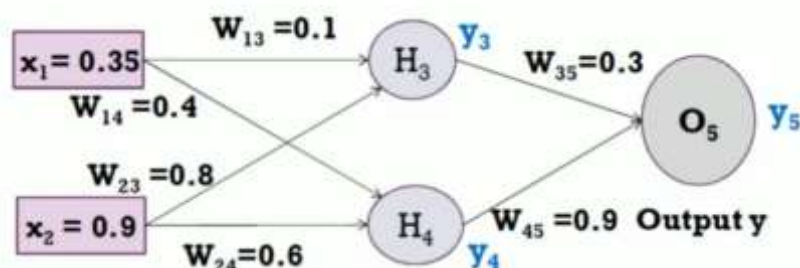
الف) استفاده از منظم سازی (regularization)، ممکن است باعث تضعیف عملکرد مدل شود.

ب) اضافه کردن تعداد زیاد ویژگی های جدید، باعث جلوگیری از بیش برآزش (overfitting) می شود.

ج) با زیاد کردن ضریب منظم سازی (regularization)، احتمال بیش برآزش (overfitting) بیشتر می شود.



۳- فرض کنید در هنگام آموزش یک شبکه عصبی وزن ها به صورت زیر درآمده است:



حال می خواهیم داده ی آموزشی زیر را برای آموزش این شبکه بکار ببریم؛

x1	x2	y
0.35	0.9	0.5

که x_1, x_2 ورودی های شبکه و y خروجی مدنظر با این ورودی ها (خروجی واقعی که انتظار داریم، شبکه با این ورودی ها آن را تولید کند) می باشد.

الف) خروجی تولید شده توسط این شبکه برای ورودی گفته شده را محاسبه نمایید. (محاسبه y_5)

ب) با استفاده از الگوریتم گرادیان کاهشی یک بار (یک دوره) وزن های این شبکه را بروزرسانی کنید.

توابع تمام نورون ها را سیگموئید و نرخ آموزش را نیز برابر یک و بایاس تمام نورون ها را برابر صفر در نظر بگیرید.

۴- در این بخش می خواهیم با استفاده از کتابخانه پایتورچ، آموزش یک شبکه پرسپترون چند لایه بر روی مجموعه داده MNIST پیاده سازی نماییم؛ برای این کار کافی است تا نوتبوک موجود در کنار این فایل را تکمیل نمایید (بخش هایی که باید تکمیل شود با TODO مشخص شده است و برای راحتی کار مابقی بخش ها نوشته شده است).

این کار را با بهینه سازها، توابع ضرر و توابع فعال سازی مختلف انجام دهید و نتایج را با یکدیگر مقایسه و تحلیل نمایید. (پیشنهاد می شود تابع اصلی را به صورت ماژولار بنویسید به طوری که پارامترهای مختلف شبکه را به صورت ورودی بگیرد و شبکه را با استفاده از این ساختار تشکیل و آموزش دهد و نتایج را بازگرداند و سپس این تابع را با ورودی های مختلف (بهینه سازها، توابع ضرر و توابع فعال سازی مختلف) صدا بزنید و نتایج را گزارش کنید).

تعداد لایه ها و تعداد نورون های لایه مخفی به عهده شما می باشد. سعی کنید از overfit شدن شبکه جلوگیری نمایید و این مورد (overfit بودن شبکه) را نیز چک نمایید و در صورت استفاده از راهکارهایی برای جلوگیری از آن، این راهکار (ها) را شرح دهید.

لطفا سند قوانین انجام و تحویل تمرین های درس را مطالعه و موارد خواسته شده را رعایت فرمایید

موفق و سلامت باشید