



دانشکده فنی دانشگاه تهران

دانشکده برق و کامپیوتر

تمرین ۳ سیستم‌های هوشمند

رایانامه

sharifi.sina1377@gmail.com

طراحان:

سینا شریفی

نیم سال اول ۱۳۹۹-۱۴۰۰

دانشجویان عزیز، قبل از پاسخ‌گویی به سوالات به نکات زیر توجه کنید:

۱. کدها و گزارش باید با الگو `IS_HW3_StudentNumber.zip` در محل تعیین شده آپلود شوند.
۲. گزارش کار شما نیز از معیارهای ارزیابی خواهد بود، در نتیجه زمان کافی برای تکمیل آن اختصاص دهید.
۳. شما می‌توانید سوالات خود را از طریق ایمیل sharifi.sina1377@gmail.com پرسید.

از مسائل مهم در زمینه‌ی یادگیری ماشین، **Classification** و **Regression** می‌باشند. تا به حال روش‌های مختلفی برای این کار پیشنهاد شده است، اما یکی از روش‌هایی که در چندسال اخیر توانسته به دقت و سرعت خیلی خوبی دست یابد، شبکه عصبی می‌باشد.

۱. به سوالات تشریحی زیر پاسخ دهید.

(آ) در مورد نرمال کردن و استاندارد کردن داده‌ها قبل از آموزش مدل تحقیق کنید و علت این کار و تفاوت بین این دو روش را توضیح دهید.

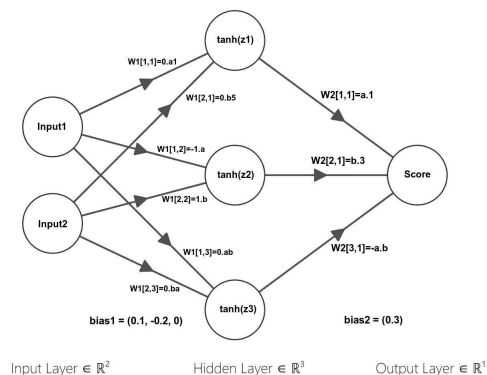
(ب) در مورد Dropout تحقیق کنید و علت استفاده از آن را توضیح دهید.

(ج) در صورتی که در هنگام آموزش شبکه، مقدار Loss برای داده‌های **train** و **validation** به هم خیلی نزدیک شود، به چه معناست، توضیح دهید.

اگر این موضوع مطلوب ما نباشد، راه جلوگیری از آن چیست؟ و اگر مطلوب ما باشد به چه شکل می‌توانیم به آن دست یابیم؟
(د) شبکه عصبی شکل ۱ را با استفاده از روش **gradient descent** تا ۲ مرحله **update** کنید، همچنین تابع هزینه را **MSE** فرض کنید (در لایه آخر تابع فعال ساز وجود ندارد). ($\text{learning rate} = 0.1$)

که منظور از **a** و **b** در شکل ۱ به ترتیب رقم دهگان و یکان شماره دانشجویی شما می‌باشد. (برای مثال $a=1, b=2$), $-a.b$ برابر -1.2 میشود) مسئله ما از نوع **Regression** می‌باشد و فرض کنید که ورودی شبکه و خروجی مطلوب ما: $input =$

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}, output = 2$$



شکل ۱: شبکه عصبی تک لایه

۲. (آ) هدف از این سوال، پیاده سازی پایه‌ای (from scratch) و آشنایی با مفاهیم ابتدایی شبکه عصبی می‌باشد.
می‌خواهیم با استفاده از یک شبکه عصبی با یک لایه پنهان یا همان **hidden layer** داده‌های موجود را دسته‌بندی کنیم. ابتدا به صورت دستی، رابطه گرادینان را برای هر دو ماتریس وزن بدست آورید و آن را در گزارش کار خود بنویسید. (از تابع هزینه **MSE** و تابع فعال ساز **tanh** استفاده کنید).

(ب) برای شبکه خود تابعی به اسم **InitWeight** تعریف کنید که در آن وزن‌های اولیه شبکه تعیین می‌شود. (تعداد **node** برای **hidden layer** را ۱۵۰ فرض کنید).

(ج) در دنیای فعلی شبکه‌های خیلی بزرگی ساخته شده‌اند طبیعتاً برای آموزش آن‌ها زمان زیاد صرف شده است، بدیهیست که آموزش مجدد این شبکه‌ها کار پرهزینه‌ای می‌باشد، به همین منظور دو تابع با نام‌های **SaveWeight(path)** و **LoadWeight(path)**

تعریف کنید که در آن‌ها وزن‌های شبکه را ذخیره و بارگذاری می‌کند. (در توابع گفته شده **path** همان موقعیت مطلوب ما برای ذخیره و یا محل یافتن وزن‌ها می‌باشد)، در گزارش کار خود ذکر کنید از چه **Method** برای وزن دهی اولیه استفاده کردید.

(د) در این قسمت شما باید یک تابع **forward** برای شبکه خود تعریف کنید که به ازای ورودی مطلوب به شما خروجی تحویل دهد.

(ه) حال به اصلی‌ترین بخش شبکه رسیدیم، شما می‌بایست با استفاده از الگوریتم پس انتشار یا همان **back propagation** وزن‌های شبکه را به روز رسانی کنید.

(و) نتایج، شامل نمودار دقت و **loss** روی داده‌های **Train** و **Validation** نمایش دهید، و همچنین دقت و **loss** روی داده‌های **Test** را بعد از پایان آموزش شبکه نشان دهید.

(ز) سوال قبل را برای حالتی که تعداد **node** های **hidden layer** برابر ۱۰۰ و ۳۰۰ هست تکرار کنید و نتیجه را مقایسه کنید. همچنین تابع فعال ساز را هم **ReLU** را نیز امتحان کنید و آن را با **tanh** مقایسه کنید.

توجه کنید در این سوال مجاز به استفاده از کتابخانه‌هایی شامل **Pytorch**، **Keras** و **scikit learn** نیستید و باید مراحل طراحی و آموزش شبکه را خودتان پیاده‌سازی کنید.

۳. در این بخش می‌خواهیم با یکی از **framework** های معروف و ساده برای طراحی شبکه عصبی آشنا شویم. در این بخش قصد داریم از کتابخانه **keras** برای پیاده‌سازی این سوال استفاده کنید. می‌خواهیم دیتاست **fashion mnist** را دسته بندی کنیم. یک شبکه دلخواه طراحی کنید و موارد زیر را انجام دهید و نتایج را در گزارش کار خود بیاورید.

(ا) از تابع فعالساز **sigmoid** و **relu** استفاده کنید و نتیجه را مقایسه کنید.

(ب) یک بار با استفاده از لایه **dropout** و یک بار بدون استفاده از آن شبکه را آموزش دهید. چه تفاوتی ایجاد می‌شود. توضیح دهید.

(ج) در تمامی موارد فوق دقت روی داده‌های تست و نمودار دقت روی داده‌های **train** و **Validation** را در گزارش کار بیاورید.

موفق باشید