

هدف این تمرین آشنایی با کتابخانه TensorFlow و آموزش دسته‌بند (۱۰ دسته‌ای) با استفاده از شبکه‌های چند لایه تمام متصل بر روی دادگان **MNIST** است. در تمامی قسمت‌ها، تابع هزینه را Cross Entropy بین خروجی لایه Softmax و بردار Label که به صورت one-hot کد شده است، در نظر بگیرید. ورودی شبکه نیز یک بردار ۷۸۴ تایی از تصاویر (ابعاد تصویر  $28 \times 28$ ) است.

۱. یک شبکه عصبی تمام متصل با یک لایه مخفی به تعداد ۵۰ نورون با تابع فعال‌ساز sigmoid آموزش دهید. برای بهینه‌سازی از Gradient Descent<sup>۱</sup> با Learning Rate برابر ۰,۵ استفاده کنید. همچنین تعداد داده‌ها (سایز mini-batch) در هر بار به‌روزرسانی وزن‌ها را برابر ۶۴ بگیرید. نمودار تابع هزینه و همچنین دقت دسته‌بندی را برای داده‌های یادگیری و اعتبارسنجی را در گزارش خود ارائه کنید. (برای رسم نمودارها از TensorBoard استفاده کنید.) همچنین دقت دسته‌بند را برای داده‌های تست به دست آورید.

۲. سوال ۱ را با تابع فعال‌ساز tanh (برای لایه مخفی) تکرار کرده و نتایج را در گزارش خود ارائه کنید. در کدام حالت، دقت دسته‌بند بیشتر است؟ نتیجه خود را تحلیل کنید.

۳. سوال ۲ را با ۵۰۰ نورون در لایه مخفی تکرار کرده و نتایج را در گزارش خود ارائه کنید.

۴. شبکه سوال ۲ را با یک شبکه عصبی تمام متصل با دو لایه مخفی که هر لایه مخفی ۵۰ نورون (با تابع فعال‌ساز tanh) دارد جایگزین کرده و یادگیری را مجدداً انجام دهید. نتایج را (مشابه سوال ۲) در گزارش خود بیاورید.

۵. تعداد پارامترهای سوال‌های ۲-۴ را به دست آورده و با هم مقایسه کنید. همچنین نتایج و نمودارهای به دست آمده در این سوال‌ها را با هم مقایسه و تحلیل کنید.

۶. سوال ۴ را با Learning Rate‌های مختلف (۰,۰۰۱، ۰,۰۱، ۰,۵ و ۱) پیاده‌سازی کنید و نتایج آن را در گزارش خود ارائه کنید. با توجه به نمودارها توضیح دهید Learning Rate چگونه در آموزش شبکه تاثیر دارد؟

۷. سوال ۴ را برای مقادیر متفاوت (۱۶، ۶۴، ۱۲۸ و ۲۵۶) برای سایز mini-batch پیاده‌سازی کنید و نتایج آن را در گزارش خود ارائه کنید. با توجه به نمودارها توضیح دهید سایز mini-batch چه تاثیری در آموزش شبکه دارد؟

۸. سوال ۴ را با بهینه‌سازهای RMSProp (با Learning Rate برابر ۰,۰۰۱) و Adam (با Learning Rate برابر ۰,۰۰۱) پیاده‌سازی کنید و نتایج آن را در گزارش خود ارائه و مقایسه کنید. (برای سایر hyper-parameterهای بهینه‌سازها از مقادیر پیش فرض آن‌ها در توابع tensorflow استفاده کنید.)

۹. گراف شبکه سوال ۴ را با استفاده از TensorBoard رسم کرده و در گزارش خود ارائه کنید. (قسمت‌های مختلف شبکه را با استفاده از تعیین scope در کد تفکیک کنید)

توجه: گزارش خود را به صورت pdf همراه کدهای قسمت‌های مختلف (در فایل‌های مشخص و مجزا) آپلود کنید.

<sup>۱</sup>tf.train.GradientDescentOptimizer