

به نام خدا



درس: یادگیری عمیق

استاد: دکتر فاطمی زاده

گزارش تمرین عملی شماره ۳

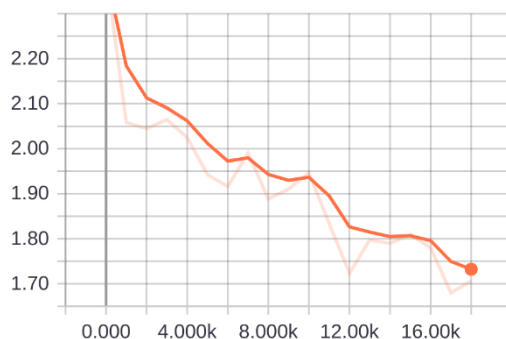
سید محمد امین منصوری طهرانی

۹۴۱۰۵۱۷۴

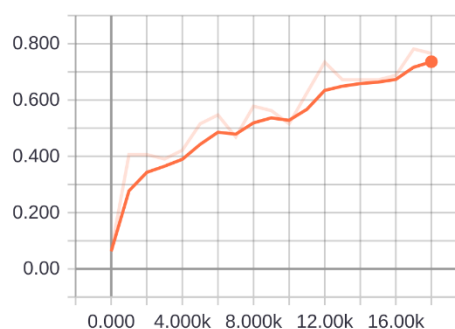
توجه: کدها به صورت کامل کامنت گذاری شده اند و لذا از آوردن توضیح قطعه های کد در گزارش خودداری می کنیم.

۱. نتایج در زیر آورده شده اند. طریقه بدست آوردن نمودارها: مطابق کدی که در پیوست قرار گرفته است، برای حالتی که یک لایه نورون مخفی داریم، هر بار با رسیدن mini-batch جدید مدل را train می کنیم و در هر ۱۰۰۰ بار، یک بار خطا را به همراه دقت با برداشتن یک mini-batch جدید (که احتمال استفاده شدن آن در مدل به خاطر تعداد زیاد داده ها کم است) محاسبه می کنیم. در آخر نیز یک mini-batch از داده تست استفاده می کنیم و خطای مدلی که کامل آموزش دیده است را محاسبه می کنیم. نتیجه آن در زیر دو نمودار زیر آورده شده است.

Cross_Entropy_validation



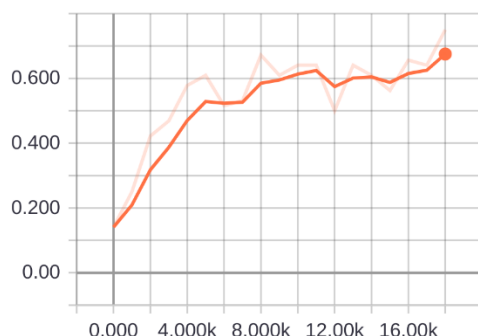
accuracy_train



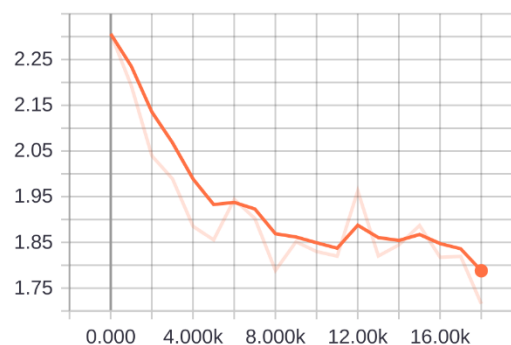
The accuracy of test = 65.4865

۲. در کد مربوط به خروجی لایه مخفی، از tanh استفاده کردم و نتیجه مطابق شکل های زیر شد. دقت ها تقریباً برابر است و با تکرار محاسبه برای هر دو مشاهده می شود که tanh اندکی بهتر است و در بعضی موارد نیز تفاوت چندانی مشاهده نمی شود و نتیجه گیری قطعی امکان پذیر نیست.

accuracy_validation



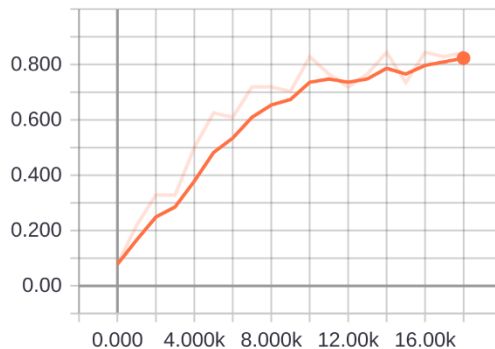
Cross_Entropy_validation



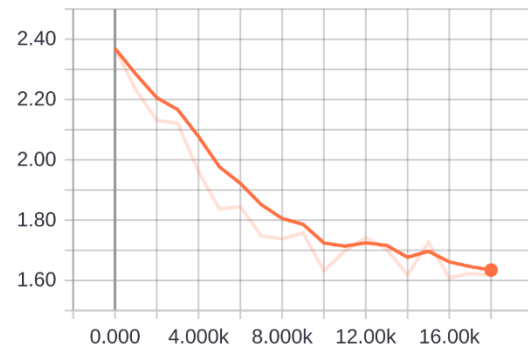
The accuracy of test:0.703125

۳. پارامتر تعداد نورون‌های لایه مخفی را به ۵۰۰ تغییر دادم. نتایج در زیر آورده شده‌اند.

accuracy_validation



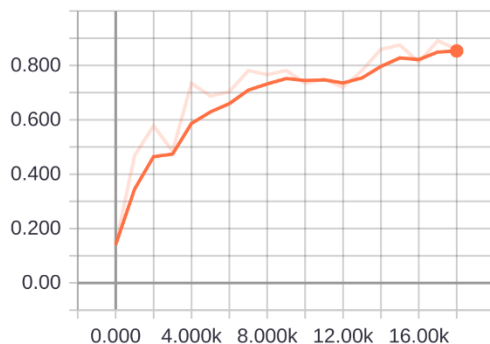
Cross_Entropy_validation



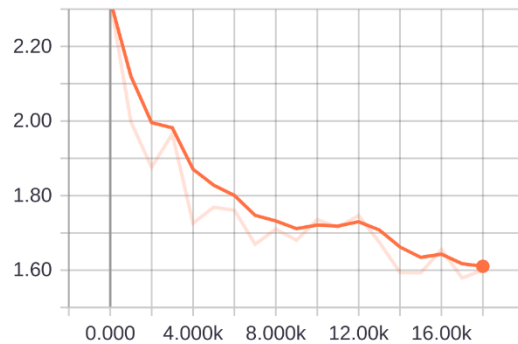
The accuracy of test:0.8125

۴. کد این قسمت به بعد در فایل دیگری قرار داده شده و خروجی لایه مخفی اول پس از گذشتن از tanh، در بردار وزنی جدید ضرب شده و با بایاس جمع می‌شود تا ورودی لایه مخفی دوم را تشکیل دهد و ادامه آن مثل قبل است. پارامتر تعداد نورون لایه مخفی نیز ۵۰ تنظیم می‌شود. نتایج در زیر قابل ملاحظه‌اند.

accuracy_validation



Cross_Entropy_validation



The accuracy of test:0.859375

۵. دقت شبکه دو لایه‌ای با ۵۰ نورون در هر لایه از دقت شبکه با ۵۰۰ نورون در لایه مخفی بیشتر بوده و دقت شبکه اخیر از دقت شبکه تک لایه‌ای با ۵۰ نورون بیشتر است. در نمودارها نیز مشاهده می‌شود که تابع هزینه به ترتیبی که گفته شد سرعت افت خود را از دست می‌دهد. دقت ذکر شده در نمودارها نیز قابل مشاهده است. تعداد پارامترهای شبکه دو لایه بین دو شبکه دیگر قرار و دقت آن از هر دو بیشتر است.

تعداد پارامترها برای مدل‌ها به شرح زیر است:

$$\text{مدل سوال ۲: } 784 \times 50 = 39200$$

مدل سوال ۳: $784 \times 500 = 392000$

مدل سوال ۴: $784 \times 50 + 50 \times 50 = 41700$

۶. فایل‌های event این بخش برای نمایش مناسب در tensorboard، در folder های جداگانه قرار گرفته‌اند (تا هنگام فراخوانی tensorboard رنگ‌ها متفاوت باشد) و به همین علت برای نمایش همزمان نیاز دارید با برچسب آن‌ها را در تنسوربرد اجرا کنید و مشاهده کنید. تصاویر نتایج را خلاصه می‌کنند. در نرخ‌های کم سرعت یادگیری کم است و برای همین دقت خوب نشده و بهترین نرخ ۱ بوده است.

دقت بر روی داده‌های تست به ازای نرخ یادگیری ۰,۰۰۱:

The accuracy of test:0.046875

دقت بر روی داده‌های تست به ازای نرخ یادگیری ۰,۰۱:

The accuracy of test:0.203125

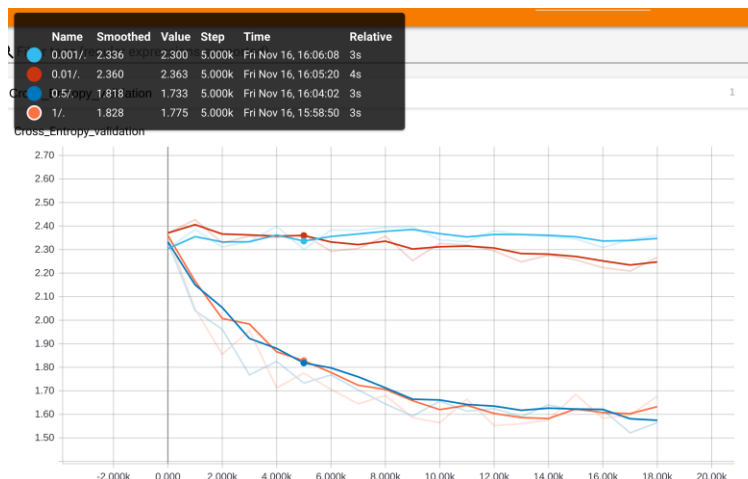
دقت بر روی داده‌های تست به ازای نرخ یادگیری ۰,۵:

The accuracy of test:0.78125

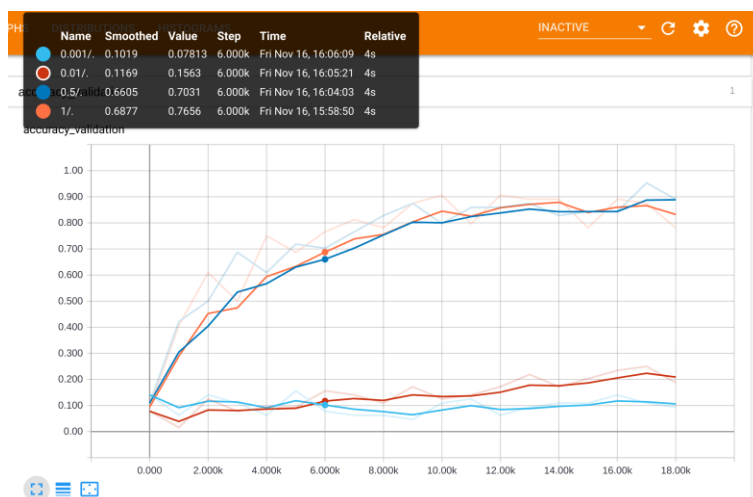
دقت بر روی داده‌های تست به ازای نرخ یادگیری ۱:

The accuracy of test:0.890625

مقایسه cross entropy برای هر ۴ حالت:



مقایسه accuracy برای هر ۴ حالت:



۷. اگر این مقدار کم شود، به حالت Online یا sample mode نزدیک می‌شویم که دقت بیشتر شده ولی زمان اجرا بیشتر می‌شود. با افزایش این اندازه، زمان هر آپدیت شدن زیاد می‌شود (تعداد محاسبات گرادیان به صورت خطی افزایش می‌یابد). اما گرادیان دسته بزرگتر تقریب بهتری از گرادیان است تا تقریب یک نمونه. (دومی اصطلاحاً نویزی تر است). البته نویز دار بودن می‌تواند ما را از دره‌هایی که در آن‌ها گیر می‌کنیم فراری دهد و احتمالاً به همین دلیل در این مسأله جواب دسته‌های کوچک‌تر بهتر شده است.

دقت بر روی داده‌های تست به ازای اندازه mini-batch برابر ۱۶:

The accuracy of test:0.9375

دقت بر روی داده‌های تست به ازای اندازه mini-batch برابر ۶۴:

The accuracy of test:0.875

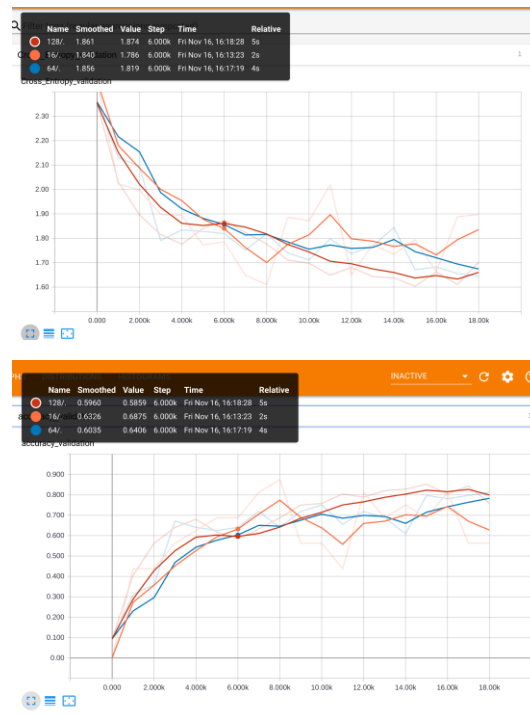
دقت بر روی داده‌های تست به ازای اندازه mini-batch برابر ۱۲۸:

The accuracy of test:0.859375

دقت بر روی داده‌های تست به ازای اندازه mini-batch برابر ۲۵۶:

The accuracy of test:0.78125

تابع هزینه و دقت اعتبارسنجی:



۸. مشاهده نشان می‌دهد که دقت دو مدل و هزینه آن‌ها بسیار نزدیک است و مشابه هم عمل می‌کنند.

تابع هزینه و دقت اعتبارسنجی:



