

بسم الله الرحمن الرحيم

دانشگاه علم و صنعت ایران

بهار ۱۴۰۰

تحويل: شنبه ۸ خرداد

تمرین سری دهم

مبانی یادگیری عمیق

۱. فرض کنید خروجی یک لایه به ازای N داده به صورت زیر باشد (ستون‌ها D و ردیف‌ها N می‌باشند). نتیجه batch normalization و layer normalization را محاسبه کنید (برای محاسبات مجاز هستید از توابع پایه numpy استفاده کنید).

29	39	75	38	35
60	31	49	27	96
48	26	13	29	90
65	97	62	34	11
46	71	76	40	87

۲. در کدامیک از لایه‌های batch normalization, instance normalization, layer normalization و group normalization نیاز است تا میانگین و واریانس داده‌های آموزشی برای زمان تست ذخیره شوند؟

۳. در این سوال هدف آشنایی بیشتر با مدل Word2Vec است. در این کد ابتدا می‌توانید مدل word2vec از مدل از پیش آموزش داده شده load کنید. توجه کنید که مدل آموزش دیده در زبان انگلیسی است و در قسمت های بعد باید کلمات انگلیسی مورد نظر را انتخاب کنید.

الف) در این قسمت یک کلمه انتخاب کنید. پس از اجرای سلول قسمت A کلمات مشابه با کلمه مورد نظران را ارزیابی کنید که مدل تا چه میزان خوب عمل می‌کند و چه اشتباهاتی وجود دارد؟

ب) در قسمت B ابتدا یک کلمه انتخاب کنید. دو کلمه دیگر نیز انتخاب کنید به طوری که فاصله یکی از آن‌ها تا کلمه اول کم و دیگری زیاد باشد. آیا انتظاری که برای دور بودن و نزدیک بودن کلمات داشتید برآورده شد؟ اگر نه فکر می‌کنید مشکل چیست؟

پ) در قسمت C سه کلمه انتخاب کنید. هدف این است دنبال کلماتی بگردید که به کلمات اول و دوم شبیه و متفاوت از کلمه سوم باشد. همچنین کلماتی که به کلمات اول و سوم شبیه و متفاوت از کلمه دوم باشد. نتایج را ارزیابی کنید.

ت) نتایج حاصل از قسمت D را تحلیل کنید (توجه داشته باشید که Word2Vec یکی از روش‌های موجود برای یافتن word embedding است و قرار نیست کاملاً مطابق با پیش‌بینی ما عمل کند).

۴. در این سوال می‌خواهیم یک مثال ساده از یادگیری ویژگی‌های بصری با استفاده از رویکرد یادگیری خودنظارتی را پیاده‌سازی کنیم. مراحل زیر را بر روی مجموعه داده CIFAR10 انجام دهید. برای حل این تمرین یک شبکه با قابلیت یادگیری بالا با استفاده از لایه‌های کانولوشنی و دیگر لایه‌های خوانده شده طراحی کنید و تمام مراحل زیر را با استفاده از آن انجام دهید. در این آزمایش، از داده‌های آموزشی هر کلاس تنها ۲۰ داده را دارای برچسب نگه می‌داریم و باقی داده‌ها را بدون برچسب استفاده خواهیم کرد. به عبارت دیگر، در مجموع ۲۰۰ داده آموزشی دارای برچسب و ۴۹۸۰۰ داده آموزشی بدون برچسب برای آموزش مدل خواهیم داشت و ۱۰۰۰۰ داده تست دارای برچسب برای ارزیابی مدل خواهیم داشت (در این کد، نحوه آماده‌سازی داده‌ها مشخص شده است).

الف) مدل خود را تنها با استفاده از داده‌های آموزشی دارای برچسب آموزش دهید و بر روی داده‌های تست ارزیابی کنید.

ب) با استفاده از داده‌های آموزشی بدون برچسب، مسئله تشخیص زاویه تصویر را حل کنید. سپس، لایه انتهایی شبکه را حذف کرده و بجای آن یک لایه دارای ۱۰ نورون برای دسته‌بندی قرار دهید و مدل خود را با این وزن‌های اولیه و با استفاده از داده‌های آموزشی دارای برچسب آموزش دهید (با نرخ آموزش کوچکتر) و ارزیابی کنید.

پ) مدل خود را به گونه‌ای تغییر دهید که دارای دو خروجی باشد (یک خروجی برای دسته‌بندی زاویه و یک خروجی برای دسته‌بندی ۱۰ کلاس). سپس، مدل خود را با تمام ۵۰۰۰۰ داده آموزشی آموزش دهید (۴۹۸۰۰ نمونه از داده‌ها دارای برچسب نیستند و بنابراین برای این داده‌ها خروجی مطلوب دسته‌بندی ۱۰ کلاس را برابر با بردار صفر قرار دهید تا اثری روی تابع ضرر آن نداشته باشند). مدل آموزش دیده را بر روی داده‌های تست ارزیابی و با نتایج قبل مقایسه کنید. در این حالت، میزان اثر هر تابع ضرر باید به درستی تنظیم شود (با توجه به کم بودن داده‌های دارای برچسب، اثر آنها در مجموع کم خواهد بود). چند ضریب مختلف برای تابع ضرر تخمین زاویه را امتحان کنید و نتایج خود را با دقت تحلیل کنید.

* برای تعریف یک مدل با چند خروجی می‌توانید از مدل `functional` در `keras` استفاده کنید. همچنین، برای تعیین وزن هر کدام از توابع ضرر می‌توانید از `loss_weights` در هنگام `compile` مدل استفاده کنید. برای راهنمایی بیشتر می‌توانید از این [لینک](#) کمک بگیرید (البته توجه داشته باشید که در مسئله ما، فقط ورودی دو مسئله مشترک نیست بلکه بخش عمده شبکه CNN برای دو مسئله مشترک است).

نکات تکمیلی

- ۱) لطفاً پاسخ سوالات (تئوری و توضیحات پیاده‌سازی) را به طور گویا و به زبان فارسی و در صورت امکان تایپ همراه با سورس کدهای نوشته شده، در یک فایل فشرده شده به شکل `HW10_YourStudentID.zip` قرار داده و بارگذاری نمایید.
- ۲) منابع استفاده شده را به طور دقیق ذکر کنید.
- ۳) برای سهولت در پیاده‌سازی‌ها و منابع بیشتر، زبان پایتون پیشنهاد می‌شود. لطفاً کدهای مربوطه را به طور جداگانه در فرمت `ipynb` ارسال نمایید.
- ۴) ارزیابی تمرین‌ها براساس صحیح بودن راه حل‌ها، گزارش مناسب، بهینه بودن کدها و کپی نبودن می‌باشد.
- ۵) در مجموع تمام تمرین‌ها، تنها ۷۲ ساعت تاخیر در ارسال پاسخ‌ها مجاز است اما پس از آن به صورت خطی از نمره شما کسر خواهد شد (معادل با روزی ۵۰ درصد).
- ۶) تمرین‌ها باید به صورت انفرادی انجام شوند و حل گروهی تمرین مجاز نیست.
- ۷) در رابطه با پرسش و پاسخ در رابطه با تمرین‌ها می‌توانید در گروه مربوطه مطرح کنید.

موفق باشید.