به نام خدا

دانشکده سامانه‌های هوشمند

مبانی محاسبات (رایانش) نرم (۰۱-۰۳۹-۰۵-۸۳)
نیم‌سال اول ۱۴۰۵-۱۴۰۴

تمرین شماره ۳

تاریخ تحویل: ۱۴۰۴/۰۹/۲۱

---

۱. (۲۵٪) [پژوهش و محاسبه] برای هر کدام از سوالات زیر پاسخ را بدست آورید.

۱-۱ [۵٪] دو مفهوم Vanishing problem و Exploding problem را توضیح داده و راه‌حل‌هایی را برای برطرف کردن این مشکل‌ها پیشنهاد دهید. کدام یک از توابع فعال سازی می‌توانند باعث این دو پدیده شوند؟

۱-۲ [۵٪] در مورد L1 Regularization و L2 Regularization تحقیق کنید و تفاوت میان آن‌ها را با جزئیات شرح دهید و مشخص کنید هر کدام از آنها چگونه بر به‌روزرسانی وزن‌های مدل اثر می‌گذارند. کدام مورد را ترجیح می‌دهید؟ دلیل انتخاب خود را توضیح دهید.

۱-۳ [۵٪] پدیده شکست تقارن (Symmetry breaking) چیست؟ آن را توضیح دهید. برای جلوگیری از آن چه کار(های)ی می‌توان انجام داد؟

۱-۴ [۵٪] مفهوم Quantization مدل در شبکه‌های عصبی را شرح دهید، چند روش را نام برده و مزیت‌ها آن ها را بیان کنید.

۱-۵[۵٪] بهینه‌ساز یا Optimizer در شبکه های عصبی چه کاربردی دارد؟ روش کار، مزیت و معایب موارد زیر را بیان کنید.

- Stochastic Gradient Descent (SGD)
- Mini-Batch SGD
- SGD with Momentum
- Adagrad
- RMSProp
- Adam

---

۲. (۴۵٪) [پیاده‌سازی: دسته‌بندی تصاویر با شبکه عصبی پیچشی] در این بخش قرار است که با استفاده از یک چارچوب برنامه‌نویسی یادگیری عمیق، یک مدل برای دسته‌بندی تصاویر مجموعه داده‌های CIFAR-10 پیاده‌سازی کنیم. این مجموعه داده از ۶۰ هزار تصویر رنگی با ابعاد ۳۲×۳۲ تشکیل شده است (شامل ۵۰ هزار تصویر آموزشی و ۱۰ هزار تصویر آزمایشی). تصاویر این مجموعه داده در ده دسته قرار گرفته‌اند که نمونه‌ای از هریک از این دسته‌ها در شکل ۱ ملاحظه می‌شود.

h.veisi@ut.ac.ir

به نام خدا

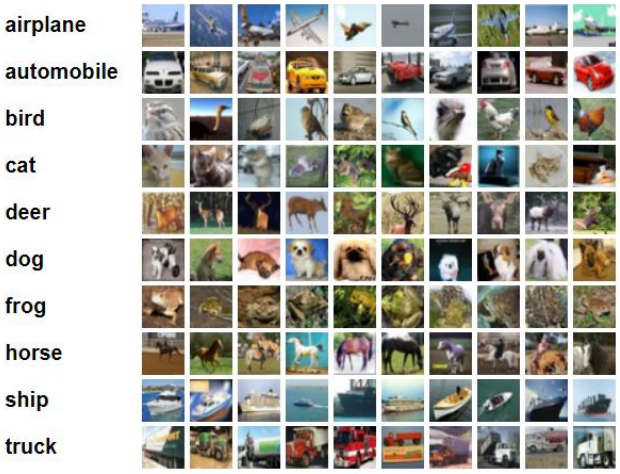
دانشکده سامانه‌های هوشمند

مبانی محاسبات (رایانش) نرم (۰۱-۰۳۹-۰۵-۸۳)

نیم‌سال اول ۱۴۰۵-۱۴۰۴

تمرین شماره ۳

تاریخ تحویل: ۱۴۰۴/۰۹/۲۱



شکل ۱ نمونه تصاویری از هریک از برچسب‌های مجموعه داده‌ی CIFAR-10

به صفحه‌ی رسمی این مجموعه داده[1] مراجعه کنید، آن را دریافت کرده و مجموعه‌های آزمایشی و آموزشی آن را در این بخش از تمرین مورد استفاده قرار دهید. داده‌های آموزشی این پایگاه داده در پنج دسته‌ی جداگانه (هریک با اندازه‌ی ۱۰ هزار تصویر) ارائه شده‌اند. در صورتی که منابع شما اجازه‌ی اجرای الگوریتم بر روی تمامی مجموعه داده را نمی‌دهد، می‌توانید به عنوان داده‌ی آموزشی، تنها دسته اول آن (data_batch_1) را مورد استفاده قرار دهید. توجه کنید که در این صورت لازم است این موضوع را در گزارش خود ذکر کنید.

شبکه‌ای پیچشی با معماری زیر در نظر بگیرید:

۱) لایه‌ی پیچشی: اندازه‌ی کرنل: ۴×۴ ، (تعداد کانال‌های خروجی: ۶۴ و تابع فعال‌سازی: ReLU)
۲) لایه‌ی ادغام بیشینه: اندازه‌ی کرنل: ۲×۲
۳) لایه‌ی پیچشی: اندازه‌ی کرنل: ۴×۴ ، (تعداد کانال‌های خروجی: ۳۲ و تابع فعال‌سازی: ReLU)
۴) لایه‌ی ادغام بیشینه: اندازه‌ی کرنل: ۲×۲
۵) لایه خطی با تعداد ۲۵۶ (تعداد دسته‌ها) نرون خروجی

[1] https://www.cs.toronto.edu/~kriz/cifar.html

h.veisi@ut.ac.ir

۶) لایه خطی با تعداد ۱۰ (تعداد دسته‌ها) نرون خروجی

این شبکه را با بهینه‌ساز Adam، تابع خطای categorical cross entropy و نرخ یادگیری یک هزارم آموزش دهید (با اندازه دسته ۳۲). این معماری و نحوه‌ی آموزش را به عنوان حالت پایه در نظر گرفته. در هریک از بخش‌های **الف** تا **ذ** تغییرات گفته شده را بر روی آخرین معماری اعمال کنید.

نکته ۱: قسمتی از داده‌های آموزشی را به عنوان داده‌های validation جدا کرده (می‌توان به عنوان یک پارامتر هنگام فراخوانی تابع آموزش نیز تنظیم شود) و میزان خطای آن را معیاری برای زمان توقف الگوریتم قرار دهید. همچنین از دقت داده‌های آزمایشی به منظور ارزیابی مدل استفاده کنید.

نکته ۲: برای هریک از بخش‌های **الف** تا **ذ**، موارد زیر را گزارش کنید:

- نموداری از روند تغییرات خطای شبکه در حین آموزش، بر روی داده‌های train و validation (این دو نمودار را روی یکدیگر قرار دهید)
- دقت مدل بر روی داده‌های آزمایشی
- میانگین مدت زمان اجرای تکرار[2]ها
- تحلیل تغییرات «سرعت همگرایی بر اساس تعداد تکرار» و «توان شبکه در یادگیری الگو» در اجراهای مختلف
- بررسی تمام موارد بالا و تحلیل علت تغییرات آن‌ها

الف) نتایج مندرج شده از مشاهدات خود را از مدل پایه بنویسید. (آیا مدل دچار بیش برازش (Overfitting) شده است؟ اگر پاسخ شما بلی است، راه‌حلی پیشنهاد دهید.)

ب) تعداد کانال‌های خروجی لایه اول را به ۳۲ تغیر دهید و لایه drop-out با احتمال ۳۰ درصد را به مدل اضافه کنید. تغییرات این مدل را با مدل قبلی مقایسه کنید.

پ) تعداد لایه‌های پیچشی را از دو لایه به سه (تعداد فیلترها در هر سه لایه برابر با ۳۲ فیلتر) و چهار لایه (تعداد فیلترها در هر چهار لایه برابر با ۳۲ فیلتر) و اندازه‌ی کرنل لایه‌های پیچشی را به سایز ۳×۳ تغییر دهید.

ت) لایه‌ی Batch Normalization را بعد از هر لایه پیچشی به مدل اضافه کنید.

[2] epoch

h.veisi@ut.ac.ir

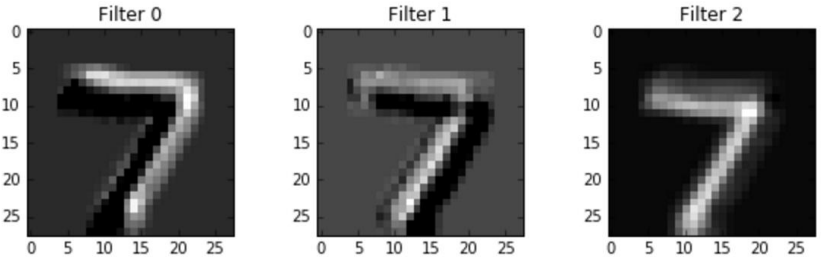
ث) ضریب یادگیری بهینه‌ساز را به 10^{-4}، 10^{-3} و 10^{-1} تغییر دهید.

ج) تابع فعال‌ساز را به leaky ReLU و Swish تغییر دهید.

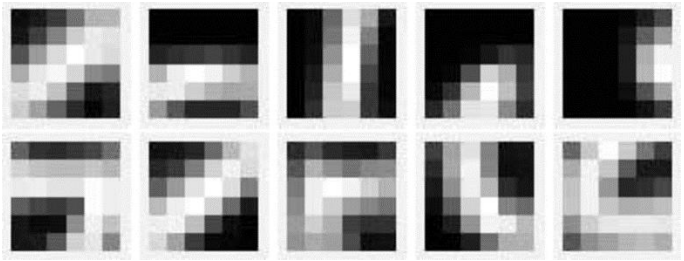
چ) بهینه‌ساز را به SGD و AdamW تغییر دهید.

ح) اندازه‌ی batch-size را به ۴ و ۱۲۸ تغییر دهید.

خ) پس از آموزش مدل، یکی از تصاویر آزمایشی را در نظر بگیرید، آن را نمایش دهید و آن را به عنوان ورودی به شبکه بدهید. پس از عبور تصویر از لایه‌ی اول پیچشی، هریک از کانال‌ها تصویری خاکستری[3] با ویژگی خاصی از تصویر ورودی را نمایش می‌دهد (شبیه به شکل ۲). همچنین وزن‌های مربوط به یکی از کانال‌های کرنل را نیز می‌توان به عنوان یک تصویر خاکستری در نظر گرفت و آن را نمایش داد (شبیه به شکل ۳). حال شما تصویر کانال‌های مختلف لایه اول پیچشی را در کنار تصویر وزن‌های کرنل‌های همان کانال نمایش دهید و ارتباط آن‌ها را توضیح دهید.



شکل ۲ نمونه سه تصویر که توسط سه فیلتر مختلف در لایه‌های میانی شبکه‌ی عصبی تولید شده است



شکل ۳ نمونه ۱۰ تصویر از وزن‌های کرنل (فیلتر) با اندازه‌ی ۶×۶

[3] gray-scale

h.veisi@ut.ac.ir

د) تعدادی از داده‌های تست را انتخاب کنید و نتایج پیش‌بینی شده توسط مدل را با نتایج واقعی مدل نمایش دهید.

ذ) ماتریس درهم‌ریختگی (confusion matrix) را نمایش دهید.

ر) از بین معماری های ResNet50, VGG19 و EfficientNetB0 به دلخواه دو مورد را در نظر بگیرید. در این بخش، با به‌کار گیری نمونه‌ی پیش‌آموزش دیده‌شده‌ی این معماری‌ها بر روی مجموعه‌داده‌ی ImageNet[4] (که به صورت آماده در چارچوب‌های برنامه‌نویسی tensorflow و pytorch قرار داده‌شده است)، می‌خواهیم مدلی برای دسته‌بندی مجموعه‌ی داده‌ی CIFAR-10 آموزش دهیم. برای این منظور لازم است به انتهای بخش استخراج ویژگی[5] این شبکه، با وزن‌های آموزش دیده‌شده، لایه‌های خطی (با تعداد خروجی‌های ۲۵۶ و در نهایت ۱۰ دقیقا مانند شبکه CNN ۴ لایه پیاده سازی شده در قسمت‌های پیشین) اضافه شود. سرعت همگرایی این مدل بر حسب تکرار و دقت آن بر روی داده‌های آزمایشی را با حالت پایه مقایسه، علت تفاوت‌ها را بیان کرده و نکات مثبت و منفی استفاده از مدل پیش‌آموزش دیده‌شده را بررسی کنید. ایده و نوع اوری هر شبکه را بیان کنید.

---

۳. (۳۰٪) [**پیاده‌سازی: تولید تصاویر اعداد دست‌نویس با شبکه مولد متخاصمی**] در این پرسش به پیاده‌سازی یک شبکه مولد تخاصمی GAN می‌پردازیم. هدف ما در این تمرین بررسی بخش‌های مختلف شبکه‌های مولد و یادگیری مفهوم و درک توانایی این شبکه می‌باشد. مجموعه دادگان این تمرین، تصاویر MNIST هستند.

نکته: در این تمرین مجاز به استفاده از کتابخانه‌های اماده هستید.

نکته: در این تمرین تنها بخش‌هایی که ناقص هستند باید توسط دانشجویان تکمیل شود. این بخش‌ها در کد که در پیوست ارسال شده مشخص شده‌اند.

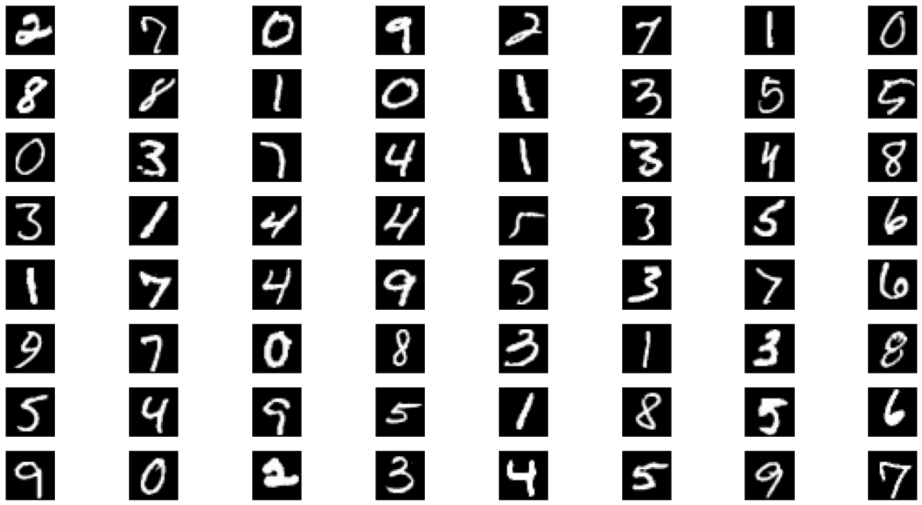
[4] http://image-net.org/
[5] Feature extraction

h.veisi@ut.ac.ir

به نام خدا

تاریخ تحویل: ۱۴۰۴/۰۹/۲۱

تمرین شماره ۳



نمونه ای از مجموعه دادگان تصاویر MNIST

۳-۱ [۵٪] [تولید توزیع گوسی] در بسیاری از مسائل واقعی، داده‌های ما توزیع‌های پیچیده‌ای دارند و GANها ابزاری هستند که می‌توانند این توزیع‌ها را مدل‌سازی کرده و داده‌های جدید تولید کنند. در این بخش قصد داریم یک توزیع گوسی (نرمال) یک‌بعدی را از یک توزیع یکنواخت تولید کنیم. هدف اصلی این بخش آشنا کردن شما با ساختار اولیه یک شبکه مولد متخاصم (GAN) و گام هایی است که برای پیاده‌سازی این مدل باید طی شود. این تمرین به شما کمک می‌کند تا مفاهیم اساسی شبکه‌های مولد را به صورتی بنیادین درک کنید. در این بخش شما باید قسمت های مختلف شبکه GAN شامل بخش‌های:

- متمایز کننده (Discriminator)
- مولد (Generator)
- تابع آموزش شبکه و توابع مکمل آن

را تکمیل کنید.

۳-۲ [۲۰٪] [تولید تصویر عدد دست نویس] در این بخش با برداشتن گامی فراتر، هدف ما تولید تصاویر اعداد دست نویس از بردار نویز می‌باشد. با کامل کردن بخش‌های مشخص شده که شامل توابع

h.veisi@ut.ac.ir

مولد، متمایزکننده، ضرر برای هر دو قسمت شبکه و نهایتا تابع آموزش، این شبکه خواهید ساخت.

پارامترهای شبکه به صورت زیر هستند:

- اندازه بردار نویز برابر با ۱۰۰
- نرخ یادگیری برابر با ۰.۰۰۲
- ابعاد تصاویر ۲۸*۲۸ در فرمت GREYSCALE
- اندازه batch برابر با ۶۴
- تعداد epoch حداقل برابر با ۱۰۰ دور

۳-۳ [۵٪] **[سوالات تکمیلی]** به سوالات زیر پاسخ دهید.

۳-۳-۱ با توجه به پیاده سازی انجام شده، به نظر شما مشکلات این شبکه چیست؟

۳-۳-۲ چه راه کاری برای بهبود پیشنهاد می‌دهید؟ آیا تغییر پارامترها به تنهایی کافیست؟ در صورت مخالفت، راهکار پیشنهادی خود را بیان کنید.