دانشگاه صنعتی امیرکبیر
دانشکده مهندسی کامپیوتر

پروژه درس رایانش عصبی و یادگیری عمیق

پروژه دوم

هدف: آشنایی با شبکه های عصبی چندلایه‌ی پرسپترونی.

کد: پیاده سازی این پروژه را به زبان پایتون انجام دهید؛ در این فعالیت مجاز به استفاده از tensorflow یا pytourch می‌باشید.

گزارش: ملاک اصلی انجام فعالیت، گزارش آن است و ارسال کد بدون گزارش فاقد ارزش است. برای این فعالیت یک فایل گزارش در قالب pdf تهیه کنید و در آن برای هر سوال، تصاویر ورودی، تصاویر خروجی و توضیحات کامل و جامعی تهیه کنید.

تذکر: مطابق قوانین دانشگاه هر نوع کپی برداری و اشتراک کار دانشجویان غیر مجاز بوده و شدیدا برخورد خواهد شد. استفاده از کدها و توضیحات اینترنت به منظور یادگیری بلامانع است، اما کپی کردن غیرمجاز است.

راهنمایی: در صورت نیاز میتوانید سوالات خود را در خصوص پروژه از تدریسیارهای درس، از طریق ایمیل زیر یا در گروه تلگرامی بپرسید.(لینک گروه تلگرامی در سایت کورسز در دسترس بوده و قبلا به همه‌ی دانشجویان ایمیل شده است)

Email: ann.ceit.aut@gmail.com

توجه: برای آموزش شبکه های عمیق می‌توانید از منابع و بسترهای سخت افزاری برخط رایگان نظیر Google Colab یا Kaggle استفاده نمایید.

تاخیر مجاز: در طول ترم، مجموعا مجاز به حداکثر ده روز تاخیر برای ارسال تمرینات هستید(بدون کسر نمره). این تاخیر را می‌توانید بر حسب نیاز بین تمرینات مختلف تقسیم کنید؛ اما مجموع تاخیرات تمام تمرینات شما نباید بیشتر از ده روز شود. پس از استفاده از این تاخیر مجاز، هر روز تاخیر باعث کسر ۱۰٪ نمره‌ی آن تمرین خواهد شد.

ارسال: فایل های کد و گزارش خود را در قالب یک فایل فشرده با فرمت StudentID_HW02.zip تا تاریخ ۱۴۰۲/۰۱/۱۴ ارسال نمایید.

در این تمرین با عملکرد شبکه های عصبی چند لایه پرسپترونی، آموزش، ارزیابی و فشرده‌سازی آن در دسته‌بندی آشنا خواهید شد. مجموعه داده‌ی مد نظر، تصاویر دست‌نویس ارقام انگلیسی هستند که با نام MNIST شناخته می‌شود و مشتمل بر ۷۰ هزار نمونه می‌باشد که در حالت اصلی ۶۰ هزار نمونه مربوط به آموزش و ۱۰ هزار نمونه مربوط به تست می باشد. شما برای اعتبارسنجی بایستی ۱۰ درصد از مجموعه داده آموزشی را در ابتدای کار تفکیک نمایید. دقت کنید که اگر با محدودیت سخت افزاری در زمان آموزش مواجه شدید، میتوانید تعداد نمونه های مورد استفاده را حداکثر تا ۲۰ درصد کاهش دهید. (مجموعه داده آموزشی برابر با ۱۲ هزار نمونه(۱۰ درصد از آن برای اعتبار سنجی) و ۲ هزار نمونه برای آزمون) البته توصیه می‌شود حد المقدور تعداد نمونه‌ها را کاهش ندهید.

لطفا به به هر یک از سوالات زیر حداکثر در دو بند پاسخ دهید؛ میتوانید پاسخ خود را با ارائه استدلال، ارجاع به مقالات یا مثال‌های عینی و عددی تقویت نمایید.

۱) آیا داده‌های اعتبارسنجی[1] در زمان آموزش شبکه‌های عمیق در بروزرسانی وزن ها تاثیر می گذارند و در آموزش شبکه دیده می‌شوند؟ اگر پاسخ بله است بیان کنید تفاوت آن با مجموعه داده آموزشی چیست و چرا با آن ترکیب نمی‌شود؛ اگر پاسخ منفی است بیان کنید چرا با مجموعه داده‌ی آزمون ترکیب نمی‌شود. (۵ نمره)

۲) دلیل استفاده از توابع فعال‌ساز[2] و اعمال توابع غیرخطی در شبکه‌های عصبی چیست؟ فرض کنید یک شبکه عصبی سه لایه‌ی پرسپترونی دارید و برای لایه‌های اول و دوم از tanh به عنوان تابع فعال‌ساز استفاده کرده و با آموزش کامل شبکه توانسته‌اید به دقت ٪۹۸ در امر دسته بندی دو کلاسه روی داده های آزمون دست یابید. با توضیحات و استدلال مناسب بیان کنید که تغییرات دقت شبکه به ازای تغییر تابع فعال‌ساز به Sigmoid و Linear چگونه خواهد بود. این فرض را داشته باشید که در شرایط مشابه و به ازای شبکه‌ی دو لایه در بهترین حالت به دقت ٪۷۲ دست می‌یابید. (۷ نمره)

۳) مقصود از بهینه‌سازها[3] چیست و چه کاری انجام می‌دهند؟ برای مثال می‌توانید بهینه‌ساز Adam را مختصرا معرفی و توضیح دهید. (۸ نمره)

۴) بروزرسانی وزن‌های شبکه‌های عمیق بر اساس خطای[4] محاسبه شده به ازای خروجی مطلوب و خروجی شبکه انجام می‌پذیرد. برای محاسبه‌ی خطا، توابع خطای متعددی نظیر توابع سنجش فاصله[5] مورد استفاده قرار می‌گیرد. یکی از توابع مورد استفاده در دسته‌بندی دو کلاسه آنتروپی متقابل[6] می‌باشد. نحوه عملکرد، هدف و رابطه آن را بررسی کنید. (۸ نمره)

۵) چرا در آموزش شبکه‌های عصبی مستقیما دقت شبکه را بیشینه نکرده و کمینه کردن خطا را مد نظر قرار می دهیم؟ (۵ نمره)

۶) مجموعه داده مورد نظر را بارگذاری کرده و پس از درهم ریختن[7]، از هر کلاس ۱۰ نمونه را نمایش دهید. (یک تصویر ۱۰*۱۰ که هر ستون یا هر سطر آن مربوط به یک کلاس می‌باشد) شما در مجموعه داده هر تصویر را بصورت یک ماتریس ۲۸*۲۸ در اختیار دارید. مجموعه داده را طوری تغییر شکل دهید که بردار ویژگی مربوط به هر تصویر بصورت یک بردار ۷۸۴ تایی در آید. (۷ نمره)

۷) تعیین تعداد لایه و تعداد نورون در طراحی معماری شبکه‌های عصبی چندلایه‌ی پرسپترونی یکی از چالش‌ها و جعبه‌های سیاه[8] حوزه یادگیری عمیق می‌باشد؛ هر چند اخیرا روش های تئوری بسیار قوی نظیر ReduNet [۱] در خصوص چگونگی تعیین این ابرپارامتر [9] ها ارائه شده است، اما همچنان روش های جست و جو نظیر grid-seach مورد استفاده قرار می گیرد. با استفاده از توابع آماده برای هدف grid-search، یک معماری مناسب برای دسته بندی تصاویر یافته و آن را بصورت کامل و حداقل در ۲۵-۳۰ تکرار آموزش دهید. برای آموزش نهایی خود ماتریس درهمریختگی[10]، نمودار خطا و دقت را رسم نمایید. معماری انتخابی و آموزش طوری باشد که دقت شبکه بروی مجموعه داده‌ی آزمون حداقل ۸۵-۹۰٪ باشد. (۳۰ نمره)

---

[1] Validation Set
[2] Activation Function
[3] Optimizers
[4] Loss
[5] Distance Function
[6] Cross Entropy
[7] Shuffle
[8] Black box
[9] Hyper-parameter
[10] Confusion Matrix

فشرده‌سازی تکنیکی است برای کاهش اندازه‌ی داده‌ها مانند تصویر، ویدئو، صدا یا مدل شبکه‌های عصبی. این در حالی است که اطلاعات ضروری آن را حفظ می‌کند. هدف فشرده‌سازی، کاهش هزینه‌های ذخیره‌سازی، انتقال و سرعت بخشیدن به پردازش آنها است. در شبکه‌های عصبی، فشرده‌سازی میتواند برای کاهش اندازه و هزینه محاسباتی مدل‌های بزرگ، در عین حفظ دقت یا بهبود آن استفاده شود و این عمل قرارگرفتن مدل در دستگاه‌های دارای محدودیت مانند گوشی‌های همراه را ممکن میسازد و باعث کارآمدتر شدن مدل و سریعتر شدن آن در دستگاه‌های متعدد نظیر ساعت های هوشمند و دستگاه های اینترنت اشیاء(IoT) می‌شود و میزان مصرف حافظه را کاهش می‌دهد. فشرده‌سازی همچنین میتواند شبکه‌های عصبی را قابل تفسیرتر کند و باعث بهبود قدرت تعمیم دهی آن شود.

همانطور که در بخش فوق گفته شد فشرده‌سازی یک شبکه عصبی شامل کاهش اندازه شبکه (تعداد پارامترها، محاسبات و حافظه مورد نیاز) و در عین حال حفظ دقت یا حتی بهبود آن است. این کار را میتوان از طریق تکنیک‌های مختلفی انجام داد. چند تکنیک برای فشرده‌سازی شبکه‌های عصبی عبارت است از:

الف) هرس: شناسایی و حذف نورون یا برخی از وزن‌ها و اتصالات بی‌اهمیت یا کم ارزش برای کاهش تعداد پارامترهای شبکه.(روشی راحت با پیاده‌سازی آسان و ملموس)

ب) کوانتیزه کردن: نمایش و ذخیره وزن‌ها با بیت‌های کمتر برای کاهش نیازهای حافظه و محاسبات.

ج) فاکتورسازی: نمایش ماتریس های وزنی به عنوان حاصل ضرب دو ماتریس.

د) تقطیر دانش[11]: آموزش یک شبکه کوچک‌تر برای تقلید از پیش‌بینی‌های یک شبکه بزرگتر و از قبل آموزش دیده.

ه) کدگذاری هافمن: فشرده‌سازی فعالسازها در فاز انتشار به جلو[12].

۸) یکی از روش های فشرده‌سازی فوق را انتخاب و روی شبکه‌ی چند لایه‌ی پرسپترونی آموزش دیده شده در سوال هفتم اعمال نمایید. فرآیند اعمال روش را توضیح و در ادامه به ازای مقادیر مختلف نرخ فشرده سازی (۰ تا ۹۰ درصد) نمودار دقت بر حسب آن را رسم کنید. چه نتیجه‌ای میتوان از نمودار حاصل گرفت؟ (۳۰ نمره)

[1] Chan, K., et al. "ReduNet: A white-box deep network from the principle of maximizing rate reduction." Journal of machine learning research 23.114 (2022).

موفق باشید

[11] Knowledge Distillation
[12] Forward Propagation