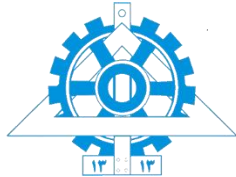


به نام خدا



سامانه‌های یادگیری ماشین توزیع شده (پاییز ۱۴۰۳)

تمرین کامپیوتری ۴

موعد تحویل: ۱۴۰۳/۱۰/۱۹

- هدف از انجام این تمرین آموزش و ارزیابی شبکه‌های عصبی به صورت توزیع شده و آشنایی با ابزار پروفایلر PyTorch می‌باشد. لطفاً قبل از انجام تمرین ویدیوهای مربوطه را مشاهده کنید و لینک‌های داده شده را مطالعه بفرمایید.
- تمرین اول و دوم و سوم باید بر روی کلاستر درس (بردهای رزبری پای) انجام شود. برای تمرین چهارم می‌توانید از کولب استفاده نمایید.
- آدرس بردهای رزبری پای به ترتیب 172.18.32.200 و 172.18.32.201 و 172.18.32.202 و 172.18.32.203 می‌باشد.
- دقت کنید که منابع کلاستر بین همه‌ی دانشجویان مشترک است. به همین دلیل سعی کنید بیشتر از مقدار مورد نیاز از آن استفاده ننمایید.
- سوالات خود را در گروه تلگرام درس مطرح نمایید یا فقط از طریق آدرس ایمیل fatemeh.javadi.y@gmail.com از من بپرسید. به هیچ وجه کد یا پاسخ سوالات را در گروه به اشتراک نگذارید.
- برای تمامی سوالات این تمرین از فایل دیتاست داده شده استفاده نمایید. دقت کنید که پوشه‌ی داده شده شامل چهار فایل است که شامل ویژگی‌ها و برجسب‌های داده‌های آموزش و تست است. داده‌ها را می‌توانید به کمک کتابخانه‌ی numpy بخوانید. فایل‌های داده‌های ویژگی برای کاهش سایز به صورت float16 ذخیره شده‌اند برای اینکه بتوانید از آن‌ها برای آموزش استفاده کنید لازم است تا آن‌ها را به float32 تبدیل کنید.
- در گزارش خود، کد تمام بخش‌ها را توضیح دهید (در مورد سولاتی که کدها یکسان هستند، یک بار توضیح دادن کافیست). هم‌چنین تحلیل و نتیجه‌گیری خود را بنویسید.
- به نحوه‌ی تحویل تمرین که در انتهای این فایل آورده شده است، دقت کنید.

سوال یک: (۳۵ نمره)

- هدف از این تمرین آشنایی با نحوه‌ی استفاده از torchrun برای آموزش شبکه‌ی عصبی به صورت توزیع شده به کمک slurm است.
- الف)** مدلی شامل لایه‌های FeedForward, ReLU و BatchNorm طراحی کرده و بسازید که با آن بتوان به دقت بالای ۸۰ درصد روی داده‌های تست رسید. (سایر پارامترهای مدل را به صورت دلخواه قرار دهید).
- ب)** چکپوینت مدل را در انتهای آموزش، تنها به وسیله‌ی یک process ذخیره نمایید.
- ج)** به کمک torchrun و slurm مدل را در حالت‌های زیر آموزش دهید.

- یک ماشین و یک هسته
- یک ماشین و دو هسته
- دو ماشین و یک هسته
- دو ماشین و دو هسته

(د) زمان آموزش و دقت مدل را در هر چهار حالت محاسبه کرده و گزارش کنید. نتایج را با هم مقایسه کرده و تحلیل خود را بنویسید.

سوال دو: (۳۵ نمره)

هدف از این تمرین آشنایی با نحوه استفاده از کتابخانه‌ی `huggingface accelerate` برای آموزش شبکه عصبی به صورت توزیع شده به کمک `slurm` است. این کتابخانه برای ساده سازی آموزش و استقرار شبکه‌های عصبی بر روی سخت افزارهای مختلف و با کانفیگ‌های مختلف مورد استفاده قرار می‌گیرد. به همین جهت می‌توان برای آموزش شبکه‌ی عصبی روی چند CPU از این کتابخانه استفاده کرد.

(الف) با استفاده از شبکه‌ای که در بخش الف سوال یک طراحی کردید، با کمک `slurm` و `accelerate` باید مدل را در حالت‌های مختلف آموزش دهید. برای اینکار لازم است تا تغییراتی را در کد خود اعمال کنید. این تغییرات را اعمال کرده و در گزارش خود بیاورید و توضیح دهید.

(ب) با کمک `slurm` و `accelerate` مدل را در حالت‌های زیر آموزش دهید. مدل برای استفاده از کتابخانه‌ی `accelerate` باید از طریق دستور زیر محیطی که این کتابخانه در آن نصب شده است را فعال کنید. از شما انتظار می‌رود که این دستور را در فایل `bash` خود بیاورید. عدم انجام اینکار موجب کسر نمره می‌شود.

➤ `source /home/shared_files/pytorch_venv/bin/activate`

در انتهای این فایل لینک‌های مورد نیاز جهت انجام این تمرین برای شما آورده شده است. می‌توانید از فایل `config` خود `accelerator` استفاده کنید، اما توصیه می‌شود با استفاده از `command line` در فایل `bash` آرگومان‌های لازم را تعریف کنید.

- یک ماشین و یک هسته
- یک ماشین و دو هسته
- دو ماشین و یک هسته
- دو ماشین و دو هسته

(ج) زمان آموزش و دقت مدل را در هر چهار حالت محاسبه کرده و گزارش کنید. نتایج را با هم مقایسه کرده و تحلیل خود را بنویسید.

سوال سوم: (۱۵ نمره)

کتابخانه‌ی `accelerate` امکان استفاده از یک روش بهینه‌سازی برای آموزش به اسم `mixed precision` را فراهم می‌کند که در حالت دیفالت غیرفعال است، در رابطه با این روش آموزش و مقادیری که این آرگومان می‌تواند بگیرد تحقیق کرده و آن‌ها را در گزارش خود بیاورید. این آرگومان علاوه بر مقدار دیفالت می‌تواند ۳ مقدار دیگر را داشته باشد. دو مقدار آن را می‌توانید برای حالت‌های

زیر روی کلاستر درس آزمایش کنید. برای مقدار سوم در حالت یک ماشین و یک هسته کد را اجرا کرده و علت اینکه نمی‌توان از این مقدار استفاده کرد را بیان کنید.

- یک ماشین و یک هسته
- یک ماشین و دو هسته

زمان آموزش و دقت مدل را در هر دو حالت و برای مقادیر متفاوت این آرگومان محاسبه کرده و گزارش کنید. نتایج را با هم مقایسه کرده و تحلیل خود را بنویسید.

سوال چهارم: (۱۵ نمره)

هدف از این تمرین آشنایی با ابزار پروفایلر PyTorch می‌باشد. ابزار پروفایلر مدت زمان و حافظه‌ی مصرفی برای هرماژول مدل را محاسبه می‌کند.

الف) داده‌های تست سوال اول را مطابق با توضیحات قبلی لود نمایید و مدلی دقیقاً مشابه به مدل بخش الف قسمت یک لود نمایید. مدل را بر روی CPU لود کنید. در این بخش نیازی به آموزش مدل ندارید. مدل را بر روی صد batch از داده‌ی تست پروفایلر نمایید و زمان و حافظه‌ی مصرفی هر یک از ماژولهای Linear، BatchNorm و ReLU را گزارش کرده و با یکدیگر مقایسه نمایید.

ب) به جای تابع فعالساز ReLU، از توابع فعالساز Tanh، Sigmoid و GeLU استفاده نمایید و زمان و حافظه‌ی مصرفی هر چهار تابع فعالساز را با همدیگر مقایسه نمایید.

لینک‌های مورد استفاده:

1. <https://huggingface.co/docs/accelerate/en/index>
2. https://huggingface.co/docs/accelerate/en/package_reference/cli
3. <https://github.com/huggingface/accelerate>
4. <https://huggingface.co/docs/transformers/en/accelerate>
5. <https://www.digitalocean.com/community/tutorials/multi-gpu-on-raw-pytorch-with-hugging-faces-accelerate-library>
6. https://pytorch.org/tutorials/recipes/recipes/profiler_recipe.html

نحوه تحویل پروژه

- یک فایل گزارش به صورت pdf شامل تمامی موارد خواسته شده
- نام گذاری کدها حتما به فرم زیر انجام شود.

تمامی فایل‌های خواسته شده را در یک فایل zip در سامانه آپلود کنید.

نام فایل	بخش	سوال
Q1_SMSC.py, Q1_SMSC.sh	بخش ج (یک ماشین و یک هسته)	1
Q1_SMMC.py, Q1_SMMC.sh	بخش ج (یک ماشین و دو هسته)	
Q1_MMSC.py, Q1_MMSC.sh	بخش ج (دو ماشین و یک هسته)	
Q1_MMMC.py, Q1_MMMC.sh	بخش ج (دو ماشین و دو هسته)	
Q2_SMSC.py, Q2_SMSC.sh	بخش ب (یک ماشین و یک هسته)	2
Q2_SMMC.py, Q2_SMMC.sh	بخش ب (یک ماشین و دو هسته)	
Q2_MMSC.py, Q2_MMSC.sh	بخش ب (دو ماشین و یک هسته)	
Q2_MMMC.py, Q2_MMMC.sh	بخش ب (دو ماشین و دو هسته)	
Profiler.ipynb	شامل تمامی بخش‌ها	4