



به نام خدا  
درس یادگیری عمیق  
تمرین سری هفتم  
استاد درس : دکتر محمدرضا محمدی  
دستیاران : فاطمه ستوده، رضا علیدوست،  
علی سبحانی و مرتضی حاجی آبادی  
دانشگاه علم و صنعت ایران، دانشکده مهندسی کامپیوتر  
نیمسال دوم تحصیلی ۱۴۰۳ - ۱۴۰۴

## مهلت تحویل : ۱۴۰۴/۰۵/۰۵

لطفا به نکات موجود در سند قوانین انجام و تحویل تمرین ها دقت فرمایید.

### سوالات تئوری



۱. روش‌های پیشرفته‌ی یادگیری ماشین خودکار مانند جستجوی معماری عصبی (NAS) و AutoAugment برای خودکارسازی جنبه‌های پیچیده طراحی مدل توسعه یافته‌اند. اصول اصلی این روش‌ها را توضیح دهید و موارد زیر را بررسی کنید. برای هریک از ۲ روش ذکر شده، موارد ذیل را به صورت جداگانه مورد بحث قرار دهید: (۲۰ نمره)

- (آ) نحوه تعریف مسئله‌ی خودکارسازی توسط هر تکنیک
- (ب) سازوکارها یا ساختارهای کنترلی مورد استفاده (مثل Controller-Trainer، sub-policies)
- (ج) چالش‌های مرتبط با پیچیدگی فضای جستجو و هزینه محاسباتی
- (د) نقش یادگیری تقویتی (RL)، بهینه‌سازی بیزی و تطابق چگالی در حل این چالش‌ها
- (ه) ذکر نمونه‌ها یا نسخه‌های خاص هر روش (در صورت وجود)



(آ) ماتریس  $W$  زیر را در نظر بگیرید. ابتدا تعریف کنید که مرتبه (rank) یک ماتریس چیست و چه مفهومی دارد سپس مرتبه ماتریس  $W$  را به صورت دستی محاسبه کنید. در ادامه ماتریس

W را به دو ماتریس تجزیه کنید به گونه ای که حاصل ضرب آن‌ها برابر با W باشد ( از مرتبه ماتریس استفاده کنید). در نهایت محاسبه کنید که تعداد پارامترها در این تجزیه، نسبت به تعداد پارامترهای اولیه ماتریس W، چقدر کاهش پیدا کرده است. تاثیر مرتبه ماتریس را در تعداد پارامترها بررسی کنید. ( ۵ نمره)

$$W = \begin{bmatrix} 1 & 3 & 0 & 2 & 3 \\ 0 & 1 & -2 & 1 & 2 \\ 2 & 5 & 1 & 3 & 4 \\ 1 & 2 & 1 & 1 & 1 \\ 3 & 7 & 2 & 4 & 5 \end{bmatrix}$$

(ب) ( با توجه به مقاله LoRA به سوالات زیر پاسخ دهید. مدل پیش‌آموزش دیده BERT با مشخصات زیر داده شده است:

Model dimension: 1024, number of blocks: 24, number of attention heads: 16,  
vocabulary size: 30000, Rank r of LoRA: 16

هدف ما آموزش دقیق این مدل پیش‌آمخته بر روی دو تسک Question Answering و Senti-ment Analysis می‌باشد. تعداد پارامترهای قابل آموزش و تعداد پارامترهای ذخیره‌سازی برای inference را در دو حالت زیر بدست آورید: ( ۵ نمره)

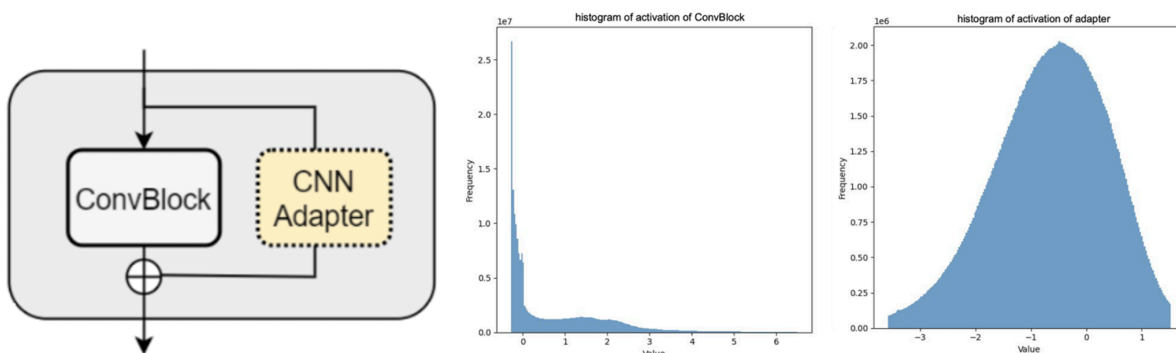
• آموزش با LoRA

• تنظیم دقیق معمولی ( بدون LoRA)

توجه: فقط پارامترهای بخش attention را در نظر بگیرید.

(ج) با توجه به مقاله LoRA، توضیح دهید که چه میزان تأخیر (latency) در مرحله inference به مدل اضافه می‌شود. درستی جواب خود را اثبات کنید. همچنین بررسی کنید که مرتبه (rank) در کدام بخش از شبکه تأثیرگذار است. ( ۵ نمره)

(د) در تصویر زیر، یک معماری آداپتور برای شبکه‌های کانولوشنی نمایش داده شده است که ماژول "آداپتور" به طور موازی با یک بلوک کانولوشنی (ConvBlock) عمل می‌کند و خروجی آن به خروجی بلوک کانولوشنی اضافه می‌شود. در تصویر بعدی، هیستوگرام‌های feature map های خروجی هر دو ماژول نشان داده شده‌اند. این هیستوگرام‌ها چه تفاوت‌هایی دارند و دلیل این تفاوت‌ها چیست؟ این تفاوت‌ها چه مشکلاتی می‌توانند ایجاد کنند و برای رفع این مشکلات چه راه‌حلی پیشنهاد می‌دهید؟ ( ۵ نمره)



۳. فرض کنید در حال کار روی یک پروژه طبقه‌بندی تصاویر پزشکی برای تشخیص یک بیماری نادر هستید. شما با دو چالش اصلی روبرو هستید:

- تعداد تصاویر برچسب‌خورده بسیار محدود است.
  - مشخص نیست که چه نوع معماری شبکه‌ای برای این داده‌های خاص بهترین عملکرد را خواهد داشت.
- با الهام از روش‌های معرفی شده در درس یک راهبرد<sup>۱</sup> جامع برای ساخت یک مدل طبقه‌بندی با کارایی بالا پیشنهاد دهید. در پاسخ خود به موارد زیر بپردازید (۱۵ نمره):
- برای انتخاب معماری، کدام رویکرد را انتخاب می‌کنید؟ دلیل انتخاب خود را توضیح دهید.
  - مزایا و معایب انتخاب شما در این شرایط خاص چیست؟
  - چالش‌های اصلی در راهبرد پیشنهادی شما کدامند؟

۴. Fast AutoAugment به عنوان یک راهکار برای رفع مشکل اصلی AutoAugment، یعنی هزینه محاسباتی بالا، معرفی شده است. با توجه به الگوریتم و توضیحات ارائه شده در درس، به سوالات زیر پاسخ دهید (۱۵ نمره):

(آ) ایده کلیدی پشت Fast AutoAugment که آن را سریع‌تر می‌کند، چیست؟ مفهوم تطابق چگالی<sup>۲</sup> را توضیح دهید.

(ب) استراتژی تقسیم داده‌ها به K-fold و استفاده از مجموعه‌های  $D_{\mathcal{A}}^{(k)}$ ,  $D_{\mathcal{M}}^{(k)}$  چگونه به کاهش هزینه محاسباتی کمک می‌کند؟

<sup>1</sup>Strategy

<sup>2</sup>Density Matching

(ج) با استفاده از جدول مقایسه هزینه محاسباتی ، تفاوت سرعت این روش با AutoAugment را برای مجموعه داده ImageNet به صورت کمی بیان کرده و اهمیت این بهبود را تحلیل کنید.

## سوالات عملی



۵. در این تمرین، قصد داریم فرآیند بهینه‌سازی یک شبکه عصبی را به صورت عملی تجربه کنیم. این فرآیند با ساخت یک مدل پایه آغاز شده و با پیاده‌سازی و مقایسه طیف وسیعی از تکنیک‌های بهینه‌سازی هایپرپارامتر، از روش‌های کلاسیک مانند جستجوی شبکه‌ای تا الگوریتم‌های پیشرفته مانند Optuna با قابلیت هرس کردن و الگوریتم ژنتیک، ادامه می‌یابد. وظیفه اصلی شما، تکمیل بخش‌های مشخص‌شده در نوت‌بوک `HyperparameterOptimization.ipynb` است. در نهایت، از شما خواسته می‌شود تا نتایج روش‌های مختلف را تحلیل کرده و کارایی و عملکرد آن‌ها را با یکدیگر مقایسه نمایید (۳۰ نمره).