



به نام خدا  
درس یادگیری عمیق  
تمرین سری پنجم  
استاد درس : دکتر محمدرضا محمدی  
دستیاران : فاطمه ستوده، نفیسه احمدی،  
محمد مصطفی رستم خانی، بهداد نادری فرد  
دانشگاه علم و صنعت ایران، دانشکده مهندسی کامپیوتر  
نیمسال دوم تحصیلی ۱۴۰۳ - ۱۴۰۴

مهلت تحویل : ۱۴۰۴/۰۳/۰۶  
لطفا به نکات موجود در سند قوانین انجام و تحویل تمرین ها دقت فرمایید.

## سوالات تئوری

۱. مقاله [Composing Parameter-Efficient Modules with Arithmetic Operations](#) را مطالعه کنید و به سوالات زیر پاسخ دهید (۱۵ نمره):

- (آ) با توجه به مقاله، هدف اصلی استفاده از روش های PEFT در مدل های زبانی چیست؟ همچنین دو روش PEFT بررسی شده در این مقاله کدامند و ویژگی های متمایز آن ها نسبت به روش های سنتی finetuning چیست؟
- (ب) مقاله نشان می دهد که ترکیب ماژول های PEFT نسبت به استفاده ی منفرد از آن ها عملکرد بهتری دارد. دلایل این بهبود عملکرد چیست؟ همچنین توضیح دهید چرا در روش پیشنهادی مقاله نیازی به آموزش مجدد ماژول ها وجود ندارد.
- (ج) طبق یافته های مقاله، چرا ترکیب PEM هایی که با مقداردهی اولیه متفاوت آموزش دیده اند، ممکن است منجر به کاهش کارایی مدل ترکیبی شود؟
- (د) مقاله چگونه نشان می دهد که ترکیب PEM ها در فضای وزن می تواند منجر به عملکردی فراتر از حالت تک وظیفه ای شود؟ با استناد به جدول ها یا نمودارهای مقاله، توضیح دهید که این ترکیب چگونه موجب تعمیم بهتر به وظایف یا داده های جدید می شود.

(ه) با وجود مزایای ترکیب ماژول‌های PEFT، مقاله به چه محدودیت‌هایی در به‌کارگیری این روش در کاربردهای واقعی مدل‌های زبانی بزرگ اشاره می‌کند؟



۲. درمورد روش MOCO به سوالات زیر پاسخ دهید (۱۵ نمره).

(آ) انتخاب ضریب تکانه مناسب برای به روز رسانی کدگذارکلید<sup>۱</sup> اهمیت بالایی دارد. دلیل آن چیست؟ اگر این ضریب بیش از حد کوچک یا بزرگ انتخاب شود چه مشکلاتی ایجاد می‌کند؟  
(ب) نقش صف نمونه‌های منفی را در این الگوریتم شرح دهید.

(ج) فرض کنید می‌خواهیم برای مسئله دسته‌بندی تصاویر MRI، از روش MOCO برای یادگیری بازنمایی استفاده کنیم. با توجه به اینکه تعداد تصاویر بدون برچسب فقط ۲۰۰۰ نمونه است، یادگیری بازنمایی به خوبی انجام نمی‌شود. اگر با استفاده از تکنیک‌های داده‌افزایی، از هر نمونه تعداد ۱۰ نمونه جدید ساخته شود تا در نهایت روش MOCO با ۲۲۰۰۰ نمونه آموزش ببیند، چه تاثیری خواهد داشت؟ در نهایت دقت مسئله دسته‌بندی چه تغییری خواهد کرد؟

## سوالات عملی



۳. در این سوال قصد داریم به قابلیت دسته‌بندی بدون نمونه در مدل Clip بپردازیم. بدین منظور از کتابخانه open-clip استفاده می‌کنیم. (۳۰ نمره)

(آ) ابتدا نسخه ConvNext-Base را فراخوانی کنید. سپس از دیتاست Stanford-dogs (مجموعه آزمون) برای ارزیابی توانایی یادگیری بدون نمونه این مدل بهره‌جوید. برای انتخاب قالب پرامپت متن مجاز هستید از یک قالب دلخواه استفاده کنید.

(ب) چند قالب متن دیگر را به دلخواه امتحان کنید و نتایج آنها را مقایسه کنید. توجه داشته باشید، به علت تعداد بالای کلاس‌های مجموعه داده، انتخاب قالب‌های متفاوت برای کلاس‌های مختلف عملی به نظر نمی‌رسد.

(ج) recall، precision، accuracy و F1 را به تفکیک کلاس گزارش کنید. کلاسی که بدترین F1 را دارد را اعلام کنید. سعی کنید برای این کلاس قالب‌های متفاوتی پیدا کنید تا عملکرد مدل درمورد آن کلاس بهبود یابد. نتایج را مقایسه کنید.


<sup>1</sup>Key Encoder


(د) برای هر کلاس یک جاگذاری قابل آموزش تعریف کنید و به کمک گرادیان کاهشی آنرا آموزش دهید (مجموعه آموزش). سپس دقت را روی مجموعه آزمون محاسبه و گزارش کنید.

(ه) رمزگذار متن را کنار بگذارید. رمزگذار تصویر را منجمد کرده و یک لایه تمام متصل اضافه کنید. لایه اضافه شده را آموزش دهید (مجموعه آموزش). سپس دقت را روی مجموعه آزمون محاسبه و گزارش کنید.

(و) اینبار مدل ConvNext-XXLarge را فراخوانی کرده و آزمایشهای قسمت الف و ب را با آن تکرار کرده، نتایج را مقایسه کنید.

**توجه: برای گزارش این سوال نتایج را مقایسه، تحلیل و سپس توجیه کنید.**

۴.  در این سوال قصد داریم با استفاده از یک مدل زبانی فارسی با soft prompting آشنا شده و سپس به روش‌هایی از Reasoning بپردازیم. قسمت‌های مشخص شده در نوت بوک LLM.ipynb را تکمیل کنید (۲۰ نمره).

۵.  در این تمرین با استفاده از کتابخانه‌های HuggingFace و PEFT، یک مدل Vision Trans-former (ViT) را روی دیتاست Food101 آموزش خواهید داد. هدف اصلی استفاده از روش Low-Rank Adaptation (LoRA) برای آموزش مؤثرتر و کم‌هزینه‌تر مدل است. لطفاً نوت‌بوک peft\_lora.ipynb را کامل کنید (۲۰ نمره).

- داده‌ها را بارگذاری و آماده‌سازی کنید (از دیتاست food101 استفاده شده است).
- مدل ViT را از HuggingFace بارگیری کنید.
- با استفاده از PEFT و LoRA، بخش‌هایی از مدل را قابل آموزش قرار دهید.
- مدل را با استفاده از Trainer آموزش دهید.
- در پایان، از مدل آموزش‌دیده برای پیش‌بینی تصویر جدید استفاده کنید.