



به نام خدا
دانشگاه تهران
دانشکده مهندسی برق
و کامپیوتر



درس شبکه‌های عصبی و یادگیری عمیق
تمرین پنجم

پرسش ۱	نام دستیار طراح	پرهام بیچرانلو
	رایانامه	Parhambicharanlu1378@gmail.com
پرسش ۲	نام دستیار طراح	جواد سراج
	رایانامه	jvseraj@gmail.com
	مهلت ارسال پاسخ	۱۴۰۳.۰۳.۲۰

قوانین	۱
پرسش ۱. تشخیص اخبار جعلی مبتنی بر مدل‌های ترنسفورمر	۱
۱-۱. آشنایی با BERT و CT-BERT	۱
۱-۲. دادگان	۱
۱-۲. پیاده سازی مدل با رویکرد fine-tuning	۲
۱-۳. پیاده سازی مدل با رویکرد feature-based	۳
۱-۴. تحلیل نتایج	۳
پرسش ۲. به کارگیری مدل‌های ترنسفرمر در طبقه بندی تصاویر	۴
۲-۱. آشنایی با ترنسفورمرهای تصویر	۴
۲-۲. لود و پیش پردازش دیتاست	۴
۲-۳. fine-tuning شبکه کانولوشنی	۵
۲-۴. fine-tuning شبکه ترنسفرمر	۵
۲-۵. مقایسه نتایج	۶

قبل از پاسخ دادن به پرسش‌ها، موارد زیر را با دقت مطالعه نمایید:

- از پاسخ‌های خود یک گزارش در قالبی که در صفحه‌ی درس در سامانه‌ی Elearn با نام **REPORTS_TEMPLATE.docx** قرار داده شده تهیه نمایید.
- پیشنهاد می‌شود تمرین‌ها را در قالب گروه‌های دو نفره انجام دهید. (بیش از دو نفر مجاز نیست و تحویل تک نفره نیز نمره‌ی اضافی ندارد) توجه نمایید الزامی در یکسان ماندن اعضای گروه تا انتهای ترم وجود ندارد. (یعنی، می‌توانید تمرین اول را با شخص A و تمرین دوم را با شخص B و ... انجام دهید)
- **کیفیت گزارش شما در فرآیند تصحیح از اهمیت ویژه‌ای برخوردار است؛** بنابراین، لطفا تمامی نکات و فرض‌هایی را که در پیاده‌سازی‌ها و محاسبات خود در نظر می‌گیرید در گزارش ذکر کنید.
- در گزارش خود مطابق با آنچه در قالب نمونه قرار داده شده، برای شکل‌ها زیرنویس و برای جدول‌ها بالانویس در نظر بگیرید.
- الزامی به ارائه توضیح جزئیات کد در گزارش نیست، اما باید نتایج بدست آمده از آن را گزارش و تحلیل کنید.
- **تحلیل نتایج الزامی می‌باشد، حتی اگر در صورت پرسش اشاره‌ای به آن نشده باشد.**
- **دستیاران آموزشی ملزم به اجرا کردن کدهای شما نیستند؛** بنابراین، هرگونه نتیجه و یا تحلیلی که در صورت پرسش از شما خواسته شده را به طور واضح و کامل در گزارش بیاورید. در صورت عدم رعایت این مورد، بدیهی است که از نمره تمرین کسر می‌شود.
- **کدها حتما باید در قالب نوت‌بوک با پسوند ipynb تهیه شوند، در پایان کار، تمامی کد اجرا شود و خروجی هر سلول حتما در این فایل ارسالی شما ذخیره شده باشد.** بنابراین برای مثال اگر خروجی سلولی یک نمودار است که در گزارش آورده‌اید، این نمودار باید هم در گزارش هم در نوت‌بوک کدها وجود داشته باشد.
- **در صورت مشاهده‌ی تقلب امتیاز تمامی افراد شرکت‌کننده در آن، 100- لحاظ می‌شود.**
- تنها زبان برنامه نویسی مجاز **Python** است.
- **استفاده از کدهای آماده برای تمرین‌ها به هیچ وجه مجاز نیست.** در صورتی که دو گروه از یک منبع مشترک استفاده کنند و کدهای مشابه تحویل دهند، تقلب محسوب می‌شود.

- نحوه محاسبه تاخیر به این شکل است: پس از پایان رسیدن مهلت ارسال گزارش، حداکثر تا یک هفته امکان ارسال با تاخیر وجود دارد، پس از این یک هفته نمره آن تکلیف برای شما صفر خواهد شد.

○ سه روز اول: بدون جریمه

○ روز چهارم: ۵ درصد

○ روز پنجم: ۱۰ درصد

○ روز ششم: ۱۵ درصد

○ روز هفتم: ۲۰ درصد

- حداکثر نمره‌ای که برای هر سوال می‌توان اخذ کرد ۱۰۰ بوده و اگر مجموع بارم یک سوال بیشتر از ۱۰۰ باشد، در صورت اخذ نمره بیشتر از ۱۰۰، اعمال نخواهد شد.

○ برای مثال: اگر نمره اخذ شده از سوال ۱ برابر ۱۰۵ و نمره سوال ۲ برابر ۹۵ باشد، نمره نهایی تمرین ۹۷.۵ خواهد بود و نه ۱۰۰.

- لطفا گزارش، کدها و سایر ضمایم را به در یک پوشه با نام زیر قرار داده و آن را فشرده سازید، سپس در سامانه‌ی Elearn بارگذاری نمایید:

HW[Number]_[Lastname]_[StudentNumber]_[Lastname]_[StudentNumber].zip

(مثال: HW1_Ahmadi_810199101_Bagheri_810199102.zip)

- برای گروه‌های دو نفره، بارگذاری تمرین از جانب یکی از اعضا کافی است ولی پیشنهاد می‌شود هر دو نفر بارگذاری نمایند.

پرسش ۱. تشخیص اخبار جعلی مبتنی بر مدل‌های ترنسفورمر

در این تمرین از مدل‌های مبتنی بر ترنسفورمرها برای تشخیص اخبار جعلی استفاده خواهید کرد.

۱-۱. آشنایی با BERT و CT-BERT

(۵ نمره)

از اولین و مشهورترین مدل‌های مبتنی بر ترنسفورمر در پردازش زبان طبیعی می‌توان به BERT اشاره کرد. این مدل از ترنسفورمرها و مکانیسم توجه^۱ برای یادگیری ارتباط معنایی بین کلمات استفاده می‌کند. و چون هدف آن تولید یک مدل زبانی^۲ است، از معماری ترنسفورمر رمزگشا^۳ استفاده می‌کند. برای جزئیات بیشتر [مقاله](#)^۴ آن را بخوانید. درباره این مدل تحقیق کنید و به سوالات زیر پاسخ دهید.

(۱) یکی از مزیت‌های مدل BERT امکان استفاده از تکنیک یادگیری انتقالی^۵ است. این تکنیک را توضیح دهید و شرح دهید که چه زمانی از آن استفاده می‌شود؟

(۲) دو رویکرد feature-based و fine-tuning برای یادگیری انتقالی از مدل‌های از پیش آموزش‌دیده وجود دارد. این دو رویکرد را با هم مقایسه کنید.

مدل BERT یک مدل زبانی عمومی است و اطلاعات کمی درباره دامنه‌های خاص مثل مالی، پزشکی یا کووید دارد. برای اینکه یک مدل نتایج بهتری در وظایف^۶ خاص بگیرد یکی از راه‌ها استفاده از رویکرد domain-specific pretraining است. که مدل CT-BERT هم از همین رویکرد استفاده کرده است. یعنی وزن‌های مدل BERT را یکبار دیگر روی دادگان مرتبط به دامنه کوید آموزش داده است. این مدل روی ۲۲/۵ میلیون توییت که حداقل یکی از کلید واژه‌های “coronavirus”، “ncov”، “wuhan”، “covid”، “sars-cov-2” بودند آموزش داده شده است.

در ادامه از این دو مدل استفاده خواهید کرد.

۱-۲. دادگان

(۱۵ نمره)

مجموعه‌ای از پست‌ها، کامنت‌ها و اخبار مرتبط به کووید را شامل می‌شوند که بر اساس صحتشان به واقعی و جعلی دسته بندی شده‌اند. این دادگان از پلتفرم‌هایی مثل توییتر و یوتیوب جمع آوری شده‌اند.

برای اینکه از این دادگان بتوانید برای آموزش مدل‌ها استفاده کنید باید ابتدا آن‌ها را پیش پردازش کنید.

- (۱) باید متن‌ها را توکنایز کنید. برای این کار از توکنایزر مدل استفاده کنید.
- (۲) طول هر متن بیشتر از ۱۲۸ نباشد. اگر طول متنی بیشتر بود آخر آن را دور بریزید.
- (۳) ایموجی‌ها را با کمک [این](#) کتابخانه^۷ به متن تبدیل کنید.

۱-۲. پیاده سازی مدل با رویکرد fine-tuning

(۳۰ نمره)

در این بخش باید بخشی از مدل‌های مقاله را پیاده سازی کنید و سپس آن‌ها را آموزش دهید. برای این کار ابتدا [مقاله](#)^۸ را کامل بخوانید و از معماری و هایپرپارامترهای پیشنهادی آن استفاده کنید. معماری کلی آن در شکل ۲ آمده است. به این صورت که ابتدا مدل زبانی مثل BERT آمده است و سپس از یکی از ساختارهای رایج شبکه عصبی مثل LSTM بر روی آن اضافه شده است.

باید این چند مدل را پیاده سازی کنید و سپس آموزش دهید:

توجه کنید که در این بخش کل لایه‌ها را آموزش می‌دهید یعنی لایه‌های مدل BERT یا CT-BERT هم علاوه بر لایه‌های روی این مدل‌ها باید در حین آموزش بروز بشوند. برای جزئیات بیشتر به بخش 3.2 مقاله مراجعه کنید.

مدل اول: در این مدل از BERT استفاده کنید و بازنمایی توکن [CLS] در لایه آخر آن را به عنوان ورودی لایه خروجی که یک لایه dense است بدهید. حال این مدل را fine-tune کنید.

مدل دوم: در این مدل از BERT استفاده کنید و بازنمایی آخرین لایه ترنسفورمر آن را به یک لایه BiGRU بدهید. و خروجی hidden state را به لایه خروجی بدهید. در نهایت این مدل را fine-tune کنید.

مدل سوم: مانند مدل دوم فقط بجای BERT از CT-BERT استفاده کنید.

برای هر کدام از مدل‌ها نمودار تغییرات خطا و دقت در هنگام آموزش بر روی دادگان آموزشی و ارزشیابی رسم نمایید. همچنین دقت، امتیاز F1 و ماتریس آشفتگی را روی دادگان تست گزارش کنید. تحلیل نتایج را در بخش نتایج انجام خواهید داد.

۳-۱. پیاده سازی مدل با رویکرد feature-based

(۳۰ نمره)

در این بخش از رویکرد feature-based برای استفاده از مدل های PLMs^۹ استفاده خواهید کرد. یعنی لایه های قسمت PLMs را در هنگام آموزش فریز می کنید. به عبارت دیگر در هنگام آموزش مدل وزن لایه های مدلی مثل BERT یا CT-BERT ثابت خواهند بود.

در این بخش باید سه مدل بخش قبل را با این رویکرد آموزش دهید.

برای مثال فرض کنید مدل دوم را می خواهید آموزش دهید. برای این کار بعد از پیاده سازی آن در هنگام آموزش وزن های لایه BERT را فریز می کنید و وزن های لایه BiGRU و لایه خروجی را باز می گذارید. پس در حین آموزش مدل فقط وزن های لایه BiGRU و لایه خروجی تغییر می کند. و مدل BERT اینجا فقط نقش feature extractor را دارد.

مانند بخش قبل نمودارهای دقت و خطا در هنگام آموزش را رسم کنید. همچنین دقت، امتیاز F1 و ماتریس آشفتگی را روی دادگان تست گزارش کنید.

تحلیل نتایج را در بخش نتایج انجام خواهید داد.

۴-۱. تحلیل نتایج

(۲۵ نمره)

در این بخش به مقایسه مدل هایی که در دو بخش قبلی پیاده سازی کردید خواهید پرداخت. صرفاً خروجی را بیان نکنید. علت هایی که باعث تفاوت شده اند را هم بیان کنید.

(۱) در مدل CT-BERT را با مدل BERT مقایسه کنید. کدام بهتر عمل کرد؟

(۲) مدل های اول و دوم بخش دوم را با هم مقایسه کنید.

(۳) مدل های مبتنی بر رویکرد fine-tuning را با مدل های مبتنی بر رویکرد feature-based مقایسه کنید.

(۴) دو نمونه از متن هایی که هر کدام از مدل های بخش دوم اشتباه پیش بینی کردند را نشان دهید. و حدس خود را برای این اشتباهات بیان کنید.

پرسش ۲. به کارگیری مدل‌های ترنسفرمر ۱۰ در طبقه بندی تصاویر

در این پرسش به بررسی مدل‌های ترنسفرمر و به کارگیری آنها در طبقه بندی تصاویر می‌پردازیم.

۲-۱. آشنایی با ترنسفرمرهای تصویر^{۱۱}

هدف این بخش، آشنایی شما با ترنسفرمر تصاویر است. مقاله ViT که در زیر آورده شده است را مطالعه کنید. سپس به سوالات زیر پاسخ دهید.

[An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale](#)

الف) برداشت خود از ساختار و نحوه کارکرد شبکه ViT را تشریح کنید. (۵ نمره)

ب) در مورد بخش‌های مختلف معماری ViT توضیح دهید و نحوه ایجاد ورودی را توضیح دهید. (۵ نمره)

ج) به نظر شما چه ایراداتی به ترانسفرمر ViT وارد است و چگونه می‌توان این معماری را بهبود داد؟ برای این پاسخ می‌توانید ایده‌های مقالات دیگر در این حوزه بهره بگیرید. (۵ نمره)

پس از آشنایی با ساختار ترنسفرمر تصاویر، در ادامه به پیاده سازی یک مدل ترنسفرمر و آموزش آن بر روی دادگان CFAR10 می‌پردازید. در مقاله زیر با unfreeze کردن لایه‌هایی از شبکه‌های ترنسفرمری و کانولوشنی، به fine-tune کردن این شبکه‌ها پرداخته است. در ادامه به پیاده سازی مقاله و ارزیابی نتایج خواهید پرداخت.

- Investigating Transfer Learning Capabilities of Vision Transformers and CNNs by Fine-Tuning a Single Trainable Block ([link](#))

برای پیاده سازی در محیط کولب، با دستور pip install کتابخانه transformers و datasets را نصب کنید.

۲-۲. لود و پیش پردازش دیتاست

(۱۰ نمره)

ابتدا داده‌های CFAR10 را لود کرده. پیش پردازش‌های لازم را بر روی دیتاست انجام دهید. برای لود دیتاست می‌توانید از torchvision.datasets استفاده کنید.

۲-۳. fine-tuning شبکه کانولوشنی

(۲۰ نمره)

الف) یک شبکه تماماً کانولوشنی از مقاله را در نظر بگیرید. سپس مدل را با وزن‌های pretrained با دیتاست imagenet1k را لود کنید. لایه‌هایی از مدل را که در مقاله ذکر شده unfreeze کنید.

ب) مشخص کنید چه تعداد از پارامترهای مدل trainable هستند.

ج) مدل را بر روی دادگان CFAR10 فاین-تیون کنید.

د) نمودار تابع هزینه و دقت را برای دادگان آموزش و اعتبارسنجی رسم کنید و دقت بر روی داده‌های آموزش و اعتبارسنجی را گزارش کنید.

ه) (امتیازی) مدت زمان آموزش و اعتبارسنجی را به ازای هر اپیک را محاسبه کنید و میانگین مدت زمان آموزش و میانگین مدت زمان اعتبارسنجی را گزارش کنید. (۲ نمره)

۲-۴. fine-tuning شبکه ترنسفرمر

(۴۰ نمره)

الف) شبکه ترنسفرمری DeiTBaseDistilled را لود کنید. برای لود کردن مدل می‌توانید از کتابخانه timm استفاده کنید. لایه‌هایی که در مقاله گفته شده است را unfreeze کنید.

ب) مشخص کنید که چه تعداد از پارامترهای مدل trainable هستند.

ج) مدل ترنسفرمر را بر روی دادگان Fine-Tune CFAR10 کنید.

د) نمودار تابع هزینه و دقت را برای دادگان آموزش و اعتبارسنجی رسم کنید و دقت بر روی داده‌های آموزش و اعتبارسنجی را گزارش کنید.

ه) (امتیازی) - مدت زمان آموزش و اعتبارسنجی را به ازای هر اپیک را محاسبه کنید و میانگین مدت زمان آموزش و میانگین مدت زمان اعتبارسنجی را گزارش کنید و این زمان را با قسمت-ه از بخش ۲-۳ مقایسه کنید. (۳ نمره)

توجه کنید برای پیاده‌سازی بخش ۲-۴ می‌توانید از هر معماری ترنسفرمری که در مقاله به آن اشاره شده به جای DeiTBaseDistilled استفاده کنید و لایه‌هایی که در مقاله ذکر شده را unfreeze کنید و به خواسته‌های این بخش پاسخ دهید.

۵-۲. مقایسه نتایج

(۱۵ نمره)

دقت داده‌های اعتبارسنجی مدل کانولوشنی و ترنسفرمری را با هم مقایسه کنید و نتایج به دست آمده خود را با نتایج مقاله مقایسه کنید. در صورت وجود اختلاف، علت اختلاف دقت های شما با دقت های گزارش شده مقاله در چیست؟

¹ Attention

² Language model

³ Encoder

⁴ <https://arxiv.org/pdf/1810.04805>

⁵ Transfer learning

⁶ Task

⁷ <https://pypi.org/project/emoji/>

⁸ <https://www.sciencedirect.com/science/article/pii/S0950705123003921>

⁹ Pre-Trained Language Models

¹⁰ Transformers

¹¹ Vision Transformers