



پروژه هشتم

هدف: آشنایی با شبکه ترنسفورمر و مفاهیم مرتبط با آن

کد: کد این فعالیت را به زبان پایتون بنویسید .

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، ورودی، خروجی و توضیحات کامل و جامعی تهیه کنید.

تذکره ۱: مطابق قوانین دانشگاه هر نوع کپی برداری و اشتراک کار دانشجویان غیر مجاز بوده و شدیداً برخورد خواهد شد. استفاده از کدها و توضیحات اینترنت به منظور یادگیری بلامانع است، اما کپی کردن غیرمجاز است.

تذکره ۲: مجموعه های داده مورد استفاده را به جز در مواردی که صریحاً در صورت سوال ذکر شده باشد، حتماً قبل از استفاده بصورت تصادفی به سه بخش آموزش (۷۰ درصد داده ها)، آزمون (۲۰ درصد داده ها) و اعتبارسنجی (۱۰ درصد داده ها) تقسیم نمایید.

راهنمایی: در صورت نیاز میتوانید سوالات خود را در خصوص پروژه از تدریس‌یارهای درس، از طریق ایمیل زیر یا در گروه تلگرامی بپرسید.

Email: ann.ceit.aut@gmail.com

ارسال: فایل های کد و گزارش خود را در قالب یک فایل فشرده با فرمت StudentID_HW08.zip تا تاریخ ۱۴۰۱/۱۱/۲۰ ارسال نمایید.

ساده سازی (simplification)^۱ و یکپارچه سازی (unification)^۲ دو مفهوم بسیار مهم در فلسفه علم هستند. آلبرت انیشتین، فیزیک دان معروف، معتقد بود که "هدف نهایی همه علوم، توجیه بیشترین تعداد ممکن پدیده ها با استنتاج از حداقل فرضیات و قوانین ممکن است"^۳. در راستای همین باور، او پس از معرفی نظریه های نسبیت، سه دهه پایانی زندگی خود را به تلاشی ناموفق برای یکپارچه سازی نیروهای گرانش و الکترومغناطیس گذراند. پس از انیشتین نیز، تلاش اکثر محققین فیزیک نظری در راستای تحقق این هدف و رسیدن به نظریه همه چیز^۳ بوده است.

¹ <https://plato.stanford.edu/entries/simplicity/>

² [https://en.wikipedia.org/wiki/Unification_\(physics\)](https://en.wikipedia.org/wiki/Unification_(physics))

³ theory of everything

ساده‌سازی و یکپارچه‌سازی در علم، تنها محدود به فیزیک نمی‌شود و در زمینه‌های دیگر مانند ریاضیات، علوم کامپیوتر و ... تلاش‌های زیادی به خود اختصاص داده‌است. تیغ اکام^۴ به عنوان یک اصل کلی در حمایت از ساده‌سازی علوم، پیش از این مورد توجه محققین هوش مصنوعی نیز بوده‌است.^۵ از طرفی هدف نهایی علم هوش مصنوعی، هوش مصنوعی عمومی (artificial general intelligence)، حل تمام مسائل با یک عامل (agent)، ناظر بر اهمیت یکپارچه‌سازی در این علم است.

در طول این ترم با چند حوزه پرکاربرد یادگیری عمیق از جمله تصویر، متن، صوت و سری‌های زمانی آشنا شدید. همانطور که دیدید در هر حوزه معمولاً یکی از دو معماری کانولوشنی یا بازگشتی به عنوان معماری پیشفرض استفاده می‌شد. اگرچه تلاش‌هایی برای استفاده از شبکه‌های کانولوشنی برای همه‌ی حوزه‌ها صورت گرفته‌است، اما هیچ یک از این تلاش‌ها به اندازه معماری ترنسفورمر در سال‌های اخیر موفق نبوده‌است. شبکه ترنسفورمر پس از معرفی در سال ۲۰۱۷ به سرعت به استاندارد اصلی حوزه متن تبدیل شد و پس از آن به تدریج به حوزه‌های دیگر راه یافت. در سال‌های اخیر تلاش‌های زیادی برای یکپارچه‌سازی مسائل یادگیری ماشین^۶ و ساده‌سازی^۷ آن‌ها به وسیله شبکه ترنسفورمر صورت گرفته که روز به روز به محبوبیت این شبکه افزوده است. به عنوان مثال شبکه [Unified-IO](#) از یک معماری ثابت مبتنی بر ترنسفورمر برای حل همزمان ده‌ها مسئله از حوزه متن و تصویر استفاده می‌کند. هدف از این تمرین، آشنایی با شبکه ترنسفورمر و بعضی از مفاهیم رایج مرتبط با آن است.

(۱) اصول اولیه

مهم‌ترین بخش شبکه ترنسفورمر را می‌توان مکانیزم توجه موجود در آن دانست. این مکانیزم بیش از ۵ سال پس از معرفی شبکه ترنسفورمر، تقریباً بدون تغییر مانده است.

الف) مکانیزم توجه مورد استفاده در ترنسفورمر را به صورت رسمی ارائه و به طور خلاصه توضیح دهید.

ب) تفاوت اصلی مکانیزم خود-توجه^۸ و مکانیزم توجه متقاطع^۹ با توجه به تعریف بالا چیست؟

(۲) پیش آموزش بدون نظارت

شبکه ترنسفورمر به عنوان یک شبکه با تعداد پارامترهای نسبتاً زیاد، مانند سایر شبکه‌های مشابه، نیاز به داده‌های آموزشی فراوان دارد. برای حل این مشکل، فرآیند [BERT^{۱۰}](#) از نخستین کارها برای پیش آموزش بدون نظارت شبکه ترنسفورمر بود. در حقیقت، موفقیت اولیه ترنسفورمرها تا حدی زیادی مدیون BERT بوده‌است.

^۴ https://en.wikipedia.org/wiki/Occam%27s_razor

^۵ <https://towardsdatascience.com/what-occams-razor-means-in-machine-learning-53f07effc97c>

^۶ Lu, Jiasen, et al. "Unified-io: A unified model for vision, language, and multi-modal tasks." arXiv preprint arXiv:2206.08916 (2022).

^۷ Carion, Nicolas, et al. "End-to-end object detection with transformers." Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16. Springer International Publishing, 2020.

^۸ self attention

^۹ cross attention

^{۱۰} Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." arXiv preprint arXiv:1810.04805 (2018).

ج) با استفاده از [این لینک](#) یا سایر توضیحات موجود در اینترنت، ایده‌های اصلی BERT برای پیش آموزش شبکه ترنسفورمر را توضیح دهید.

د) در تمرین ششم از شبکه LSTM برای تحلیل احساسات دیتاست IMDB استفاده کردید. همین مسئله را به وسیله یک شبکه پیش آموزش دیده BERT حل کنید ([راهنمایی](#)). نمودار تابع هزینه و دقت دادهای آموزش و اعتبارسنجی را نمایش دهید و تحلیل کنید.

۳) شبکه ترنسفورمر برای مسائل دیگر (امتیازی)

[ماژول Transformers](#) یک فضای یکپارچه برای استفاده از مدل‌های ترنسفورمر در حوزه‌های مختلف ایجاد کرده‌است.

ه) دو مسئله مختلف از حوزه‌های دلخواه (متن، تصویر و ...) انتخاب و بوسیله شبکه‌های پیش آموزش دیده موجود در این ماژول ترنسفورمر، از ورودی‌های دلخواه برای ایجاد خروجی مناسب استفاده کنید. (نیازی به آموزش و ارزیابی شبکه‌ها نیست. هدف، صرفاً تجربه گستردگی مدل‌های موجود در حوزه‌های مختلف است).