

۱. (۱۶٪) [مطالعه و بررسی]

- ۱-۱ توضیح دهید الگوریتم t-SNE چیست و مثالی از کاربردهای آن بیاورید.
۲-۱ توضیح دهید PCA چیست و مثالی از کاربردهای آن بیاورید.
۳-۱ PCA و t-SNE را با هم مقایسه کنید.

۴-۱ فرض کنید مجموعه داده‌ای از توییت‌ها که شامل عکس‌هایی به همراه متن می‌باشد، در اختیار داریم. قصد بر این است که مفهوم Hate Detection را مدل کنیم. برای این کار نیاز داریم تا شبکه‌ای طراحی کنیم که رابطه میان عکس و متن را یافته، به نحوی که تشخیص دهد هر داده به خصوص (شامل عکس و متن) دارای مضامین نژادپرستی و یا جنیست زده می‌باشد یا خیر. عنوانی از عکس را که مفهومی نژادپرستانه دارد در زیر مشاهده می‌کنید. الف) نخست توضیح دهید چه چالش‌هایی برای حل و مدل کردن این مسئله وجود دارد. ب) یک معماری پیشنهاد دهید که کار تشخیص این الگو را انجام دهد، با جزئیات توضیح دهید.



۲. (۳۹٪) [پیاده‌سازی: شبکه عصبی بازگشتی] در این پرسش قصد داریم با استفاده از شبکه عصبی

بازگشتی، زیرنویس مناسبی برای تصاویر مورد نظر تولید کنیم. عنوان نویسی تصویر که از آن با عناوینی همچون شرح نویسی و یا زیرنویس تصویر هم یاد می‌شود، توصیف محتوای تصویر با استفاده از نمایش متنی است. در واقع با توجه به موجودیت‌های مشاهده شده در تصویر و همچنین درک صحنه و ارتباط بین بخش‌های مختلف آن، به توصیف خودکار تصویر پرداخته و این توصیف را در قالب جمله‌ای متناسب با معیارهای زبان برمی‌گرداند. این موضوع که ترکیبی از دانش بصری کامپیوتری و پردازش زبان طبیعی می‌باشد، اخیراً به یکی از مهمترین مباحث در بینایی ماشین تبدیل شده است. رویکردهای مختلفی برای این کار پیشنهاد شده است که در بین آنها مدل‌ها مبتنی بر یادگیری عمیق به عنوان پیشرفته‌ترین مدل‌های



ثابت شده‌اند. اصولاً عنوان نویسی مبتنی بر یادگیری عمیق به مدل‌های رمزگشا-رمزگذار متکی است. روش پیشنهادی برای این سوال نیز، متکی بر همین روش می‌باشد.

توضیحات کد و داده: برای این تمرین از مجموعه داده COCO Caption^{۱۱} که در سال ۲۰۱۴ منتشر شده است استفاده می‌کنیم، که شامل ۸۰۰۰۰ تصویر آموزشی و ۴۰۰۰۰ تصویر اعتبارسنجی است. داده‌ها از قبل برای شما پیش‌پردازش و در یک فایل ذخیره شده‌اند که شامل ۱۰۰۰۰ جفت تصویر-متن برای آموزش و ۵۰۰ جفت برای آزمایش می‌باشد. برای پردازش متن‌ها نیز به هر کلمه یک شناسه عدد صحیح اختصاص داده شده است که این امکان را می‌دهد که یک عنوان را با دنباله‌ای از اعداد صحیح نشان دهیم. نگاشت بین شناسه‌های اعداد صحیح و کلمات در ورودی به نام «vocab» (هر دو «idx_to_token» و «token_to_idx») ذخیره می‌شود و از تابع «decode_captions» از «a5_helper.py» برای تبدیل تنسورهای شناسه‌های صحیح به رشته‌ها استفاده می‌کنیم.

به طور کلی، ابتدا این شبکه ویژگی‌های تصویر را استخراج کرده و به عنوان ورودی لایه مخفی به شبکه RNN می‌دهد و کلمه به کلمه جمله‌ای برای توصیف تصویر ساخته می‌شود. مراحل این عملیات در غالب توابع سازنده شبکه RNN توسط شما پیاده‌سازی خواهند شد. توابعی که توسط شما پیاده‌سازی خواهند شد شامل موارد زیر می‌شوند:

- ۱- تابع rnn_step_forward که گامی از عملیات پیشخور و تابع rnn_forward که تمامی دنباله پیشخور به شبکه را مدیریت می‌کنند.
- ۲- تابع rnn_step_backward که گامی از عملیات پس انتشار خطا و تابع rnn_backward که تمامی دنباله پس انتشار خطا را در شبکه را مدیریت می‌کنند.
- ۳- تابع WordEmbedding که وظیفه آن تعبیه‌سازی هر کلمه در یک بردار عددی در فضای چندبعدی (Embedding Space) است. این نمایش برداری به مدل کمک می‌کند تا مفاهیم معنایی یا ارتباطات بین کلمات را بهتر درک کند.
- ۴- تابع temporal_softmax_loss که وظیفه آن محاسبه‌ی خطای پیش‌بینی مدل در طول توالی‌های زمانی است، جایی که هر عنصر توالی یک پیش‌بینی برای یک کلمه از واژگان است.

¹ <https://cocodataset.org>



۵- کلاس CaptioningRNN که در برگیرنده تمامی توابع قبلی است و در آن caption برای تصاویر تولید می‌شود.

برای پیش‌برد تمرین و تکمیل کدهای مورد نظر به شکل زیر عمل کنید:

- گام اول: هنگامی که محتوای دانلود شده را از حالت فشرده خارج کردید، لطفاً پوشه را در Google Drive خود آپلود کنید. سپس، با کلیک راست روی فایل *.ipynb، آن را با Google Colab باز کنید. توصیه می‌کنیم فایل *.py خود را در Google Colab ویرایش کنید. توجه شود، نوت بوک rnn_captioning.ipynb و فایل rnn_captioning.py را در کنار هم بگذارید (به طور کلی به ساختار قرار گیری فایل‌هایی که در اختیار شما قرار داده شده است، دست نزنید!)
- گام دوم: سلول‌های نوت بوک وابستگی اجرایی به تکمیل کدهای داخل فایل *.py دارند و برای اجرای هر سلول، توابعی داخل فایل *.py. نیازمند تکمیل می‌باشند. در انتها نتیجه کار خود را در هر دو فایل *.ipynb و *.py، در Google Drive ذخیره کنید (روی «فایل» - «ذخیره» کلیک کنید) و کل فولدري که در اختیار شما قرار داده شده است حاوی نوت بوک و فایل *.py. ای که توسط شما به روزرسانی شده اند را برای ما ارسال کنید.

هنگام کار موارد زیر را در نظر داشته باشید:

- به کدهای داخل فایل *.ipynb. نباید دست بزنید و فقط کدهای فایل *.py. باید توسط شما تکمیل شوند.
- در فایل *.py. بلوک‌هایی برای نوشتن کد توسط شما مشخص شده است. هیچ کدی را خارج از این بلوک‌ها ننویسید یا تغییر ندهید.
- سلول‌های نوت بوک را نیز اضافه یا حذف نکنید.

نکته: قبل از ارسال، همه سلول‌های فایل *.ipynb. را اجرا کنید و خروجی‌ها را پاک نکنید.



۳. (۴۵٪) [پیاده‌سازی: شبکه LSTM]

هدف این تمرین آشنایی شما با وظیفه تشخیص اخبار جعلی است. در این سوال شما از مقاله Fake news detection: A hybrid CNN-RNN based deep learning approach استفاده خواهید کرد که به پیوست برای شما وجود دارد. در این بخش شما با یک binary classification رو به رو هستید و نیاز است که متن خبر موجود را به عنوان خبر صحیح یا جعلی دسته‌بندی کنید. در این مقاله نتایج روی دادگان گزارش شده‌اند، اما شما باید برای دادگان FA-KES که به پیوست برای شما ارسال شده است نتایج را گزارش کنید. در صورتی که تمایل به استفاده از پارامترهای متفاوتی از موارد گفته شده در مقاله دارید، لطفاً آن را در گزارش ذکر کنید.

۳-۱ [۵٪] [تئوری - توضیحات مدل‌ها] در ابتدا تفاوت معماری RNN و LSTM را توضیح دهید. توضیح دهید که در داده‌های متنی ویژگی بازگشتی بودن چرا مؤثر است. در مورد مدل Hybrid که در مقاله گفته شده نیز توضیح دهید و به تفاوت آن با مدل‌های بازگشتی عادی اشاره کنید.

۳-۲ [۵٪] [تئوری - ورودی مدل] با توجه به اینکه ورودی شما به عنوان اخبار در قالب متن به مدل داده خواهد شد، توضیح دهید word embedding به چه منظور استفاده می‌شود، در مورد نحوه استفاده آن برای ورودی‌های متنی تحقیق کنید و راه‌های ایجاد embedding کلمات را توضیح دهید. روشی را که برای انجام این تمرین استفاده می‌کنید نیز ذکر کنید.

۳-۳ [۱۰٪] [پیاده‌سازی - پیش‌پردازش] جهت پیاده‌سازی پیش‌پردازش‌های لازم را انجام دهید و موارد استفاده‌شده را ذکر کنید.

۳-۴ [۱۵٪] [پیاده‌سازی - آموزش مدل‌ها] پس از انجام پیش‌پردازش، شما باید برای ایجاد مدل برای انجام تسک نیاز دارید در این بخش شما یک مدل RNN، یک مدل LSTM و یک مدل hybrid (CNN-RNN) ایجاد کنید و به آموزش مدل‌ها بپردازید. در نمودار loss و Accuracy را در طول زمان یادگیری رسم کنید. معیارهای Precision، recall، F1-score را در دو مدل بدست آورید و به مقایسه دو مدل بپردازید.

۳-۵ [۱۰٪] [تحلیل نتایج] تحلیل خود را در مورد نتایج به دست آمده بیان کنید. نظر شما در مورد علت دقت‌های بدست آمده چیست، چگونه می‌توان دقت را بهتر کرد؟ آیا به نظر شما ضعفی در مدل‌ها وجود

دارد که بتوان با برطرف کردن آن به بهبود دقت کمک کرد؟
****نکته:** برای انجام این تمرین دو گزینه دارید. هم می‌توانید فایل *Istm.ipynb* را کامل کنید. یا می‌توانید خودتان مقاله را خوانده و پیاده‌سازی کنید

۴. (۳۰٪) **سوال امتیازی/اختیاری:** تولید تصویر صورت گربه با خودرمزگذار متغیر [مدل‌های VAE از جمله مدل‌های مولد هستند که با افزودن یک مؤلفه‌ی احتمالاتی به ساختار خودرمزگذارها، قابلیت تولید داده‌های جدید را فراهم می‌کنند. در این بخش به پیاده‌سازی این مدل می‌پردازیم. هدف این سوال پیاده‌سازی مدل VAE برای تولید تصاویر گربه می‌باشد و لینک دادگان در ادامه قرار داده شده است:

[لینک دادگان](#)



شکل ۱ نمونه‌ای از مجموعه دادگان تصاویر صورت گربه‌ها

۴-۱ [۳٪] **[تحقیق]** توضیح دهید که Loss یک VAE از چه اجزایی تشکیل شده است و رابطه‌ی آن را

بنویسید.

۴-۲ [۲٪] **[تحقیق]** توضیح دهید reparameterization trick چیست و چرا لازم است؟



۳-۴ [۱۰٪] [پیاده‌سازی] یک مدل VAE آموزش دهید. برای آموزش از شبکه‌ی آورده شده و ابرپارامترهای در جدول ۱ و ۲ استفاده کنید. مدل را آموزش دهید و مقادیر بازسازی (reconstruction loss) و KL و جمع کل تابع هزینه را در طول آموزش ذخیره و در یک نمودار نشان دهید.

جدول (۱) لایه‌های پیشنهادی برای شبکه

Encoder	Decoder
Conv2d(3,32)	Linear(32*32*128,h_dim)
Relu()	Relu()
Conv2d(32,64)	Conv2dT(128,64)
Relu()	Relu()
Conv2d(64,128)	Conv2dT(64,32)
Relu()	Relu()
Flatten()	Conv2dT(32,3)
Linear(32*32*128,h_dim)	Sigmoid()

جدول (۲) ابرپارامترهای پیشنهادی

Parameter	Value
Epochs	50
Optimizer	Adam
Initial learning rate	0.001
Image dimension	(256, 256)

۴-۴ [۵٪] [بازسازی] پس از اتمام آموزش ۸ نمونه تصادفی از دادگان را از مدل vae عبور دهید و نسخه بازسازی شده و اولیه را با هم مقایسه کرده و عملکرد مدل را بررسی کنید.

۴-۵ [۵٪] [تولید نمونه] با استفاده از decoder. ۱۰ تصویر جدید ایجاد کنید.

۴-۶ [۵٪] [تحقیق] توضیح دهید beta-VAE چیست و اثر تغییر β را روی بازسازی و فشردگی‌سازی

بنویسید.