



به نام خدا
دانشگاه تهران



دانشکده مهندسی برق و کامپیوتر

درس شبکه‌های عصبی و یادگیری عمیق

تمرین پنجم

توحید عبدی	نام دستیار طراح	پرسش 1
tohid.abdi@ut.ac.ir	رایانامه	
علی خرم فر	نام دستیار طراح	پرسش 2
khoramfar@ut.ac.ir	رایانامه	
۱۴۰۳.۱۰.۱۳	مهلت ارسال پاسخ	

فهرست

- 1.....قوانین
- 1.....پرسش 1. پیش‌بینی نیروی باد به کمک مبدل و تابع خطای Huber
- 1-1.....مقدمه
- 2-1.....آماده‌سازی
- 3-1.....روش‌شناسی و نتایج
- 4.....پرسش ۲ – استفاده از ViT برای طبقه‌بندی تصاویر گلبول‌های سفید
- 4-۱.....مقدمه
- 5-۲.....آماده‌سازی داده‌ها
- 5-۳.....آموزش مدل‌ها
- 7-۳.....تحلیل و نتیجه‌گیری

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، لطفاً 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 حتماً باید در قالب نوت‌بوک با پسوند ipynb تهیه شوند، در پایان کار، تمامی کد اجرا شود و خروجی هر سلول حتماً 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. پیش‌بینی نیروی باد به کمک مبدل و تابع خطای Huber

1-1. مقدمه

(15 نمره)

این تمرین برای هدف آشنایی با مسئله پیش‌بینی نیروی باد فراساحلی به کمک شبکه مبدل که قادر به شناسایی همبستگی‌های داخلی و وابستگی‌های بلندمدت در داده‌های دنباله‌ای است، طراحی شده است. برای انجام تمرین نیاز دارید که [این مقاله](#) را مطالعه کنید.

در مرحله اول، ابتدا بخش مقدمه مقاله را با دقت مطالعه کنید و سپس به پرسش‌های زیر پاسخ دهید. برای پاسخ هر پرسش 3 مورد ذکر کنید.

- روش‌های آماری سنتی مانند ARIMA و GARCH چه محدودیت‌هایی دارند؟
- برخی از مزایای مدل‌های یادگیری ماشین مانند SVM، Random Forest و XGboost نسبت به رویکردهای آماری سنتی در پیش‌بینی نیروی باد چیست؟
- مدل‌های یادگیری عمیق مانند LSTM، GRU و CNN چگونه محدودیت‌های روش‌های یادگیری ماشین سنتی در پیش‌بینی نیروی باد را برطرف می‌کنند؟
- اهمیت مکانیسم خودتوجهی (self-attention) شبکه مبدل را در یادگیری وابستگی‌های بلند مدت و همبستگی‌های محلی بیان کنید.
- تابع خطای Huber چگونه پایداری (stability) و استحکام (robustness) مدل‌های پیش‌بینی را برای داده‌های باد فراساحلی افزایش می‌دهد؟

1-2. آماده‌سازی

(20 نمره)

بخش دوم مقاله را مطالعه کنید. همان‌طور که در این بخش ذکر شده است، این تحقیق از چندین ماژول برای پیاده‌سازی آزمایش استفاده کرده است.

- ساختار Autoencoder برای چه هدفی مورد استفاده قرار گرفته است؟ این ساختار از چه بخش‌هایی تشکیل شده است؟ به کمک کتابخانه‌های مربوط، یک Autoencoder با ابرپارمترهای دلخواه طراحی کنید.
- مکانیزم توجه و Positional Encoding در شبکه مبدل برای چه هدفی مورد استفاده قرار می‌گیرند؟
- با مطالعه رابطه ریاضی تابع خطای Huber، رفتار این تابع را شرح دهید. توضیح دهید هدف از استفاده از این تابع در این آزمایش چیست. این تابع خطا را پیاده‌سازی کنید.
- روش کار الگوریتم Slime mould را بیان کنید و آن را پیاده‌سازی کنید. آرگومان‌های ورودی این الگوریتم را نام ببرید و هر یک را توضیح دهید.

1-3. روش‌شناسی و نتایج

(70 نمره)

با مطالعه بخش سوم و چهارم مقاله، آزمایش مذکور را پیاده‌سازی کنید. پیاده‌سازی شما باید مطابق با مقاله باشد و لازم است موارد زیر را رعایت کنید.

- بخشی از داده‌ها که در مقاله توضیح داده شده است را انتخاب کنید و شکل 6 مقاله را تولید کنید. (5 نمره)
- به کمک روشی که در مقاله ذکر شده، داده‌های پرت را شناسایی کرده و شکل 7 مقاله را تولید کنید. (5 نمره)

- به کمک اتوانکودری که در بخش 1-2 طراحی کرده‌اید، داده‌ها را دینویز کنید. داده‌های اولیه و دینویز شده را مانند شکل 8 مقاله مقایسه کنید. (5 نمره)
- پیش‌پردازش‌های مورد نیاز شامل تقسیم داده‌ها به آموزش و تست، نرمال‌سازی، انتخاب متغیرهای مستقل و وابسته و آماده‌سازی داده‌ها به کمک روش پنجره متحرک را انجام دهید. به نظر شما ترتیب مناسب برای انجام این مراحل چگونه است و با بیان یک مثال نقض توضیح دهید که اعمال این پیش‌پردازش‌ها با ترتیب نامناسب چه مشکلاتی را به وجود می‌آورد. (15 نمره)
- معماری‌های RNN، MLP و Transformer مذکور در مقاله را با توابع خطای MSE و Huber برای پیش‌بینی single-step پیاده‌سازی کرده و آموزش دهید (شامل 6 حالت). (15 نمره)
- پیش‌بینی مدل‌ها برای داده‌های تست را به کمک معیارهای MAE، RMSE و MAPE ارزیابی کرده و جدولی مشابه با جدول 3 مقاله نمایش دهید. توجه کنید این جدول باید به کمک کد چاپ شود و محل نمایش سطرها و ستون‌ها مشابه مقاله باشد. (10 نمره)
- (امتیازی) برای ابرپارامترهای معماری مبدل چندین مقدار در نظر بگیرید و به کمک الگوریتم Slime mould، مقادیر بهینه برای این ابرپارامترها را به دست آورید. (5 نمره)
- معماری مبدل را برای پیش‌بینی multi-step در زمان‌های $t+4$ ، $t+8$ و $t+16$ به کمک 2 تابع خطای مذکور آموزش دهید و بر روی داده‌های تست ارزیابی کنید. نتایج را به کمک جدولی مانند جدول 4 مقاله نمایش دهید. با توجه به اطلاعات دیتاست، هر یک از این زمان‌ها چند دقیقه بعد را نشان می‌دهند؟ (10 نمره)

پرسش ۲ – استفاده از ViT برای طبقه‌بندی تصاویر گلبول‌های سفید

۲-۱. مقدمه

در بدن انسان، گلبول‌های سفید خون نقش مهمی در دفاع از بدن در برابر عوامل بیماری‌زا دارند. این گلبول‌ها به چندین نوع تقسیم شده که در این تمرین **چهار نوع** از انواع مختلف آن‌ها را بررسی می‌کنیم: **نوتروفیل (Neutrophil)**: سلول‌هایی برای مقابله سریع با عفونت‌ها.

ائوزینوفیل (Eosinophil): سلول‌هایی که در مقابله با انگل‌ها و واکنش‌های آلرژیک نقش دارند.

لنفوسیت (Lymphocyte): سلول‌هایی که در تولید آنتی‌بادی و ایمنی بدن نقش دارند.

مونوسیت (Monocyte): سلول‌هایی که به پاکسازی بدن از بقایای سلولی کمک می‌کنند.

تشخیص و شمارش این سلول‌ها معمولاً توسط متخصصان آزمایشگاه با میکروسکوپ انجام می‌شود. اما این فرآیند دستی، وقت‌گیر و مستعد خطای انسانی است. به همین دلیل، از تکنیک‌های یادگیری عمیق برای طبقه‌بندی تصاویر گلبول‌های سفید استفاده می‌کنیم.

در این تمرین، می‌خواهیم دو نوع شبکه عصبی را بررسی کنیم:

ویژن ترانسفورمر (ViT): مدلی که با استفاده از مکانیسم توجه (Attention) می‌تواند به بخش‌های مختلف تصویر توجه کند و ویژگی‌های محلی و کلی تصویر را به‌خوبی استخراج کند.

شبکه عصبی کانولوشنی (CNN): شبکه‌ای که با استفاده از لایه‌های کانولوشن می‌تواند ویژگی‌های تصاویر مثل لبه‌ها، رنگ‌ها و اشکال را شناسایی کند.

هدف مقایسه معماری و عملکرد این دو مدل در طبقه‌بندی چهار نوع گلبول سفید در مجموعه داده **BCCD** است که مجموعه‌ای کوچک شامل تصاویر **نویزدار** و **کم‌کیفیت** است.

توجه داشته باشید که هدف اصلی این تمرین، آشنایی مقدماتی شما با شبکه‌های ViT و مقایسه آن با CNN است. **نیازی نیست** که دقت مدل شما دقیقاً منطبق با نتایج مقاله باشد. همچنین دیتاستی که در این تمرین استفاده می‌کنید، تفاوت‌های جزئی با دیتاست مقاله دارد و برای اهداف آموزشی تهیه شده است.

۲-۲. آماده‌سازی داده‌ها

(25 نمره)

مراحل آماده‌سازی به صورت زیر است:

- **نمایش نمونه تصاویر:** از هر کلاس یک تصویر نمونه نمایش دهید. (5 نمره)
- **بررسی تعداد داده‌ها در هر کلاس:** تعداد تصاویر موجود در هر کلاس را شمارش کنید. بررسی کنید که آیا تعداد داده‌ها در کلاس‌های مختلف متوازن است یا خیر. (5 نمره)
- اگر داده‌ها نامتوازن هستند، می‌توانید از تقویت داده‌ها (Data Augmentation) استفاده کنید. با اعمال تغییرات مختلف می‌توانید تعداد تصاویر کلاس‌های کم‌تعداد را افزایش دهید. همچنین نیاز است پیش‌پردازش لازم برای تغییر اندازه تصاویر متناسب با ورودی مدل ViT انجام شود. با توجه به محدودیت منابع پردازشی، نیازی نیست دقیقاً تعداد تصاویر ذکر شده در مقاله (مثلاً ۴۰۰ تصویر برای هر کلاس) را استفاده کنید. می‌توانید از تعداد کمتری استفاده کنید. (10 نمره)
- **ایجاد مجموعه داده نهایی:** داده‌ها را به دو بخش آموزش (Training) و اعتبارسنجی (Validation) تقسیم کنید. پیشنهاد می‌شود ۹۰ درصد داده‌ها برای آموزش و ۱۰ درصد داده‌ها برای اعتبارسنجی استفاده شود. (5 نمره)

۲-۳. آموزش مدل‌ها

(50 نمره) + 5 نمره امتیازی

ابتدا مدل ViT را بارگذاری کنید. کلاس خروجی را برای ۴ کلاس داده تنظیم کنید و معماری آن را خروجی گرفته و به‌طور کلی توضیح دهید که مدل چگونه کار می‌کند. در اینجا می‌توانید از مدل ViT گوگل استفاده کنید.

در هر حالت، نمودارهای دقت (Accuracy) و خطا (Loss) را برای داده‌های آموزش و اعتبارسنجی رسم کنید.

آموزش مدل ViT در حالت‌های مختلف:

حالت ۱: فقط دسته‌بند Classifier قابل آموزش باشد

(10 نمره)

- تمام لایه‌های مدل را فریز کنید و فقط Classifier را آزاد کنید.
- تعداد کل پارامترها و تعداد پارامترهای قابل آموزش را محاسبه کنید.
- مدل را آموزش دهید و نتایج آن را ارزیابی کنید.

حالت ۲: آموزش دو لایه اول Encoder (Fine-tune First 2 Layers)

(10 نمره)

- همه لایه‌ها را ابتدا فریز کنید. سپس دو لایه اول رمزگذار (Encoder) و دسته‌بند را آزاد کنید.
- تعداد کل پارامترها و تعداد پارامترهای قابل آموزش را محاسبه کنید.
- مدل را آموزش دهید و نتایج آن را ارزیابی کنید.

حالت ۳: آموزش دو لایه آخر Encoder (Fine-tune Last 2 Layers)

(10 نمره)

- همه لایه‌ها را ابتدا فریز کنید. سپس دو لایه آخر رمزگذار (Encoder) و دسته‌بند را آزاد کنید.
- تعداد کل پارامترها و تعداد پارامترهای قابل آموزش را محاسبه کنید.
- مدل را آموزش دهید و نتایج آن را ارزیابی کنید.

حالت ۴: آموزش تمام لایه‌ها (Full Fine-Tune)

(10 نمره)

- همه لایه‌های مدل ViT را آزاد کرده و مدل را به‌طور کامل آموزش دهید.
- تعداد کل پارامترها و تعداد پارامترهای قابل آموزش را محاسبه کنید.
- مدل را آموزش دهید و نتایج آن را ارزیابی کنید.

آموزش مدل CNN:

(10 نمره)

- یک مدل CNN مثل مدل DenseNet-121 را بارگذاری کنید.
 - مدل را به طور کامل (Full Fine-Tune) آموزش دهید.
 - تعداد پارامترها و نمودارهای دقت و خطا را مانند قسمت ViT محاسبه و رسم کنید.
- مرحله امتیازی:** فقط لایه Classifier مدل CNN را آموزش داده و نتایج را با حالت مشابه در ViT مقایسه کنید. در تحلیل نتایج خود معماری هر مدل را مد نظر قرار دهید. (5 نمره)

۲-۳. تحلیل و نتیجه گیری

(25 نمره)

- در پایان، تمام نتایج مراحل قبل CNN و ViT را با هم مقایسه کرده و سپس به سؤالات زیر پاسخ دهید:
- کدام مدل در مجموعه داده شما عملکرد بهتری داشت؟
 - آیا ViT در شرایط موجود (مثلاً داده‌های نویزدار یا کم حجم) توانسته جایگزین مناسبی برای CNN باشد؟