



دانشگاه تهران
دانشکده فنی
دانشکده مهندسی برق و کامپیوتر



پروژه درس یادگیری ماشین

(بخش سوم - مسابقه و گزارش نهایی)

دکتر مصطفی توسلی پور

دکتر محمدرضا ابوالقاسمی

طراحان پروژه:

فردین عباسی، عرفان عسگری، سیدامین موسوی

مقدمه

در پروژه نهایی درس یادگیری ماشین، قصد داریم از طریق تجربه‌ای عملی، شما را با چالش‌ها و جذابیت‌های این حوزه در دنیای واقعی آشنا کنیم. در طول این پروژه، با مسائلی مانند جمع‌آوری و پاک‌سازی داده‌ها، استفاده از مدل‌ها و روش‌های گوناگون، و نیز تحلیل و بررسی نقادانه نتایج آن‌ها روبرو خواهید شد. همانند سایر پروژه‌ها در این حوزه، برای حل مسائل لازم است با مطالعه دانش مرتبط، ویژگی‌های مناسب را استخراج کرده و مدل‌ها و روش‌های مناسب را به کار بگیرید.

در این پروژه، همچنین مسابقه‌ای میان دانشجویان درس برگزار می‌شود که برای کسب نتیجه‌ای مطلوب و برتری نسبت به سایر شرکت‌کنندگان، نیازمند داشتن روحیه تیمی، رویکردی خلاقانه و البته نگرشی نقادانه در مواجهه با مسائل خواهید بود.

همچنین در این پروژه به یکی از شاخه‌های یادگیری ماشین، یعنی یادگیری نیمه نظارت شده می‌پردازیم که از نظر عملی شباهت زیادی به چالش‌های دنیای واقعی دارد؛ چرا که در عمل، دستیابی به داده‌های برچسب‌خورده فرایندی زمان‌بر و پرهزینه است. بنابراین، نیازمند رویکردهای نوآورانه‌ای برای مواجهه با این چالش‌ها هستیم. امید است این پروژه به شما امکان دهد کاربردهای عملی روش‌های یادگیری ماشین که در طول درس فراگرفته‌اید را مشاهده کرده و مهارت‌های ضروری برای به کارگیری آن‌ها در شرایط غیرایده‌آل دنیای واقعی را کسب کنید.

بارم بندی

بارم بندی بخش‌های مختلف پروژه به صورت زیر است:

بخش پروژه	نمره
جمع آوری داده	7.5
گزارش اولیه	27.5
گزارش نهایی	25
عملکرد و مسابقه	40 + 15 (امتیازی)
مجموع	100 + 15 (امتیازی)

* توجه کنید که نمره امتیازی در نظر گرفته صرفاً برای ترمیم نمره پروژه است و در صورتی که نمره ای بیشتر از 100 کسب کنید، نمره اضافی تاثیری در دیگر بخش‌های درس ندارد.

بارم بندی بخش سوم پروژه (گزارش نهایی و مسابقه) در جدول زیر آمده است:

قسمت	گزارش	عملکرد*
1	5	20
2	10	0
3	10	20
مجموع	25	40

*عملکرد مدل‌ها با توجه به معیار ارزیابی هر قسمت در سامانه مسابقه سنجیده می‌شود.

مسابقه

برای شرکت در مسابقه نیاز است ابتدا تمامی اعضای گروه در [سایت کوئرا](#) ثبت نام کرده و از بخش تنظیمات [تشکیل تیم](#) دهند. در ادامه با مراجعه به [صفحه مسابقه پروژه](#) و وارد کردن رمز عبور MLS25 در این مسابقه ثبت نام کنید. پس از تایید ثبت نام قادر به مشاهده قالب بارگذاری نتایج سوالات خواهید بود. سنجش عملکرد مدل‌های شما توسط داوری خودکار این سامانه صورت می‌پذیرد. از بخش جدول امتیازات قادر به مشاهده نتایج تیم‌های دیگر و رقابت با آنها خواهید بود. به منظور تشویق تیم‌ها برای کسب مقام در این مسابقه، نمره امتیازی در نظر گرفته شده است:

- مقام‌های 1 و 2: 15 نمره امتیازی
- مقام‌های 3 و 4: 10 نمره امتیازی
- مقام‌های 5 تا 8: 5 نمره امتیازی

مجموعه داده

پس از جمع آوری و تجمیع تصاویر بارگذاری شده، مجموعه داده نهایی جهت بارگیری بر روی سامانه قرار داده شده است. این مجموعه دارای سه بخش است که در جدول زیر آمده است:

test آزمون	validation ارزیابی	train آموزش	
			تصویر
			کادر محاطی
			عبارت ریاضی

*: ناقص

👁: مخفی

✅: در دسترس

همانطور که مشاهده می‌شود بخش آموزش شامل تصاویر و کادر محاطی کاراکترها است. همچنین در این بخش تعداد کمی از برچسب‌های عبارات ریاضی مشخص شده‌اند. بخش ارزیابی دارای تصاویر و سایر برچسب‌ها شامل کادر محاطی و عبارات ریاضی است که می‌توانید عملکرد مدل‌های خود را روی آن بسنجید. از داده‌های بخش آزمون فقط تصاویر در دسترس هستند.

برچسب‌ها

متناظر با هر تصویر دارای برچسب، یک فایل **json** با همان نام در پوشه labels قرار دارد. در این فایل خصیصه‌های زیر آمده است:

- **نمادها: (annotations)**

لیستی از کادرهای محاطی کاراکترهای موجود در تصویر که با فرمت x,y, width, height و برحسب پیکسل می‌باشند.

- **عبارت: (expression)**

عبارت ریاضی متناظر با آن تصویر که به صورت رشته‌ای از کاراکترها می‌باشد و برای تعداد کمی از داده‌های آموزشی و تمام داده‌های ارزیابی موجود است.

قسمت 1: مکان یابی کاراکترها¹

در قسمت ابتدایی این پروژه، می بایست با پیاده سازی معماری یکی از روش های خانواده شبکه های کانولوشنی مبتنی بر ناحیه² (R-CNN)، مدلی آموزش دهید که کاراکتر های درون عبارات را مکان یابی کند. این کار معادل یادگیری پیش بینی کادرهای محاطی روی داده های آموزش است که شامل یافتن مکان و ابعاد هر کاراکتر می شود. هایپرپارامترها و معیارهای ارزیابی مربوط به این وظیفه را روی داده های آموزش و ارزیابی گزارش کنید. در آموزش شبکه فرقی بین برچسب کاراکترها وجود ندارد؛ در واقع با یک مسئله دو کلاسه "نماد" و "پس زمینه" مواجه هستید. توجه شود که در این قسمت می بایست معماری و توابع شبکه عصبی را خودتان پیاده سازی کرده و آموزش دهید و مجاز به استفاده از معماری های آماده نیستید.

 **گزارش:** در گزارش خود موارد زیر را ارائه دهید:

- روش و معماری شبکه به کار رفته (در صورت عدم ارائه توضیحات آن روش یا اصلاح آن در پیش گزارش)
- توضیح مختصر درباره ساختار کد پیاده سازی شده
- فرآیند آزمون و خطا در مسیر آموزش و تعیین پارامترها
- مثال مصورسازی شده از نتیجه رسم کادرهای محاطی
- توضیح در مورد معیار ارزیابی mAP³
- ارزیابی عملکرد مدل با استفاده از معیار ارزیابی mAP

 **مسابقه:** پس از آماده شدن مدل، کادرهای محاطی کاراکترها در تصاویر مجموعه داده آزمون را پیش بینی کنید و نتایج را مطابق با قالب بیان شده در سوال "مکان یابی کاراکترها" در صفحه مسابقه پروژه، بارگذاری کنید. معیار ارزیابی عملکرد مدل و رقابت در این قسمت mAP می باشد.

 **نکته:** با از دست دادن نمره این قسمت در عملکرد و مسابقه می توانید از کتابخانه های آماده استفاده کنید تا قادر به پیاده سازی قسمت های بعدی پروژه باشید. اما همچنان باید آموزش مدل خود را با وزن های تصادفی آغاز کنید و استفاده از وزن های از قبل آموزش داده شده⁴ مجاز نیست. توجه کنید که در این صورت نباید در این بخش از مسابقه شرکت نمایید و در صورت عدم رعایت کردن این موضوع علاوه بر حذف از مسابقه با کسر نمره از پروژه مواجه خواهید شد.

¹ Character Localization

² Region-based convolutional neural network

³ mean Average Precision

⁴ Pretrained weights

قسمت 2: خوشه‌بندی و مصورسازی

در این قسمت از هر کاراکتر به عنوان یک داده استفاده می‌کنیم. با استفاده از کادرهای محاطی مجموعه داده آموزش، تصویر کاراکترها را از عبارات جدا کرده و با استفاده از چندین روش سنتی به استخراج ویژگی از آنها پردازید. در ادامه با هرکدام از این ویژگی‌ها یا ترکیباتی از آنها و همچنین با استفاده از پیکسل‌های تصویر خام، کاراکترها را خوشه‌بندی و مصورسازی⁵ کنید.

 **گزارش:** در گزارش خود موارد زیر را ارائه دهید:

- روش‌های استخراج ویژگی به کار رفته (در صورت عدم ارائه توضیحات آن روش یا اصلاح آن در پیش‌گزارش)
- روش‌های خوشه‌بندی به کار رفته و پارامترهای آن
- نمایش اعضای نزدیک به مرکز و همچنین دورافتاده خوشه‌ها
- بحث درباره کیفیت نتیجه خوشه‌بندی‌ها بر اساس مصورسازی و استفاده از دو روش ارزیابی خوشه‌بندی دلخواه

⁵ Visualize

قسمت 3: تشخیص کاراکترهای عبارات

با توجه به اینکه در مجموعه داده آموزش تنها برچسب عبارت ریاضی برای تعداد کمی از داده‌ها موجود است و بیشتر آنها بدون برچسب می‌باشند، با شرایط یک مسئله یادگیری نیمه نظارت شده⁶ روبرو هستیم. در این قسمت باید با استفاده از روش‌های یادگیری نیمه نظارت شده روی داده‌های آموزش، کاراکترهای هر عبارت ریاضی را طبقه بندی کنید. لزومی به استفاده از ویژگی‌های استخراج شده در قسمت قبل نیست و هر روشی با ارائه توضیحات کافی قابل قبول است.

 **گزارش:** در گزارش خود موارد زیر را ارائه دهید:

- روش نیمه نظارت شده و مدل به کار رفته
- توضیح مختصر درباره ساختار کد پیاده سازی شده
- فرآیند آزمون و خطا در مسیر آموزش و تعیین پارامترها
- مثال مصورسازی شده از نتیجه طبقه بندی
- ارزیابی عملکرد مدل با استفاده از معیارهای مختلف طبقه بندی

 **مسابقه:** پس از آماده شدن مدل، عبارات ریاضی درون تصاویر مجموعه داده آزمون را پیش بینی کنید و نتایج را مطابق با قالب بیان شده در قسمت "تشخیص کاراکترها" در صفحه مسابقه پروژه، بارگذاری کنید. معیار ارزیابی عملکرد مدل و رقابت در این قسمت Levenshtein distance بین عبارت پیش بینی شده و حقیقی می‌باشد.

 **نکته:** نتایج بارگذاری شده در این قسمت باید خروجی پیش بینی مستقیم مدل روی کاراکترهای مکانیابی شده باشند. هرگونه اصلاح یا تصحیح عبارات پس از پیش بینی مدل غیر مجاز است.

⁶ Semi-supervised learning

نکات پایانی

لطفاً در ارسال به موارد زیر توجه نمایید، در صورت عدم رعایت هر یک از موارد زیر پروژه شما تصحیح نخواهد شد.

• استاندارد محیط

- در صورتی که کد را بر روی سرور شخصی خود اجرا می کنید، پیشنهاد می شود برای تفکیک پکیج های این پروژه از دیگر پروژه هایتان، از محیط مجازی⁷ مانند miniconda استفاده کنید. برای اطلاعات بیشتر این [لینک](#) را مطالعه کنید.
- در صورتی که کد را روی سرور شخصی اجرا می کنید، می بایست همه پکیج های استفاده شده در کد را به همراه ورژن در یک فایل `requirements.txt` ذخیره کنید. برای اطلاعات بیشتر این [لینک](#) را مطالعه کنید.
- از hard-code کردن آدرس فایل ها خودداری کنید. همچنین برای اینکه آدرس دهی شما در سیستم عامل های مختلف قابل اجرا باشد، از پکیج های مرتبط مانند `pathlib` استفاده کنید. برای اطلاعات بیشتر این [لینک](#) را مطالعه کنید.
- کد شما باید نتایج اعلامی را بازتولید کند؛ در نتیجه با تنظیم **SEED** برای کتابخانه های مورد استفاده و همچنین cuda از این موضوع اطمینان حاصل کنید. برای اطلاعات بیشتر این [لینک](#) را مطالعه کنید.
- مجاز به تغییر دادگان به صورت دستی نمی باشید و هرگونه تغییر فقط درون کد مجاز است.

⁷ Virtual Environment

• ساختار کد

- کد پروژه باید به زبان پایتون تحویل داده شود.
- کد پروژه را باید به صورت **Multi-File** تحویل دهید و اکیداً از زدن تمام کد در یک دفترچه خودداری کنید.
- کد تحویلی حداکثر دارای یک دفترچه **ipynb** است و باقی کدها می‌بایست **py** تحویل داده شود.
- درون دفترچه تنها مجاز به فراخوانی توابع آموزش و ارزیابی و نمایش نتایج معیارها و تصاویر می‌باشید.
- تمامی نتایج نهایی اعم از تصاویر و پرچسب‌های تولیدی، ابتدا در پوشه **results** ذخیره شود. برای نمایش آن‌ها در دفترچه **ipynb** نیز از این پوشه بخوانید.
- ساختار پوشه بارگذاری شده باید **دقیقاً** مطابق زیر باشد. کد شما باید **dataset** را از مسیر مشخص شده بخواند و تمامی کدهای **py** درون پوشه **src** و تمامی نتایج تولید شده توسط کد درون پوشه **results** قرار بگیرد. در صورت نیاز به دسته‌بندی بیشتر، مجاز به افزودن پوشه‌های بیشتر درون **src** و **results** می‌باشید.

```
proj/
├── report.pdf
├── visualizations.ipynb
├── requirements.txt
├── src/
├── results/
└── dataset/
    ├── train/
    │   ├── images/
    │   └── labels/
    ├── val/
    │   ├── images/
    │   └── labels/
    └── test/
        └── images/
```

• اصول تحویل

- هیچگونه شباهتی در انجام پروژه بین گروه‌های مختلف پذیرفته نمی‌شود. در صورت کشف هرگونه تقلب، مطابق قوانین درس با افراد خاطی برخورد خواهد شد.
- رونویسی و شباهت زیاد نسبت به گزارش دیگر گروه‌ها و همچنین استفاده مستقیم از چت بات در نوشتن گزارش موجب در نظر گرفتن نمره صفر برای کل پروژه خواهد شد.
- استفاده از مراجع با ارجاع به آن‌ها بلامانع می‌باشد.
- گزارش شما مهم‌ترین معیار ارزیابی خواهد بود؛ در نتیجه لطفاً زمان کافی را برای تکمیل آن اختصاص دهید. از انواع نمودارها برای مصورسازی نتایج استفاده کنید.
- مهلت ارسال گزارش نهایی به همراه کد و شرکت در مسابقه ساعت 23:59 روز سه شنبه 17 تیر خواهد بود. این بخش از پروژه مهلت ارسال با تاخیر ندارد و بعد از آن امکان بارگذاری در سامانه درس نیست.
- در ابتدای گزارش نام سرگروه و اعضای گروه با ذکر شماره دانشجویی مشخص شود. گزارش با نام **report.pdf** ذخیره شود.
- کد پروژه به همراه گزارش را در یک فایل **zip** جمع و با نام زیر توسط سرگروه بارگذاری شود:

proj_<StdID1>_<StdID2>_<StdID3>.zip

• پرسش و ارتباط

پس از مطالعه کامل پروژه، در صورت ابهام پیشنهاد می‌شود از طریق گروه مجازی درس سوال خود را مطرح کنید. در غیر این صورت، با افراد زیر در ارتباط باشید:

- عرفان عسگری: erfanasgari21@gmail.com
- فردین عباسی: fardinabbasi1381@gmail.com