



هوش محاسباتی - پروژه نهایی

ددلاین: ۳۰ دی ماه، ساعت ۵۹:۲۳

مدرس: دکتر ملک

نیمسال اول ۱۴۰۰-۱۴۰۱

۳ ..... پروژه تک نفره

۳ ..... هدف پروژه

۳ ..... فاز اول: پیش پردازش داده

۴ ..... فاز دوم: انتخاب و طراحی مدل

۴ ..... فاز سوم: آموزش مدل

۴ ..... پروژه دو نفره

۴ ..... هدف پروژه

۴ ..... فاز اول: جمع آوری و پیش پردازش داده

۵ ..... فاز دوم: انتخاب و طراحی مدل

۵ ..... فاز سوم: آموزش مدل

توضیح: برای پروژه نهایی دو موضوع انتخاب شده است که باید یکی از آن ها را به اختیار خودتان انتخاب کنید. پروژه اول انفرادی و پروژه دوم در قالب گروه های دو نفره انجام می شود.

## پروژه تک نفره

### هدف پروژه

هدف این پروژه تشخیص مورد پیشنهاد بودن / نبودن (recommended/not recommended) کالا با استفاده از نظرات کاربران است. انتظار می رود که در این پروژه از تکنیک های پردازش زبان طبیعی استفاده کنید تا بتوانید مثبت یا منفی بودن نظر کاربر را تشخیص دهید.

دیتاستی که در اختیار شما قرار گرفته است، شامل پنجاه هزار نمونه از نظرات کاربران نسبت به کالاهای متفاوت است. در بخش زیر توضیح مختصری نسبت به ستون داده های این دیتاست داده شده است.

- Likes: تعداد Like های یک کامنت را نشان می دهد.
- Dislike: تعداد Dislike های یک کامنت را نشان می دهد.
- Title: دارای Title کامنت می باشد و ممکن است نظر کلی یک کامنت را دربرگیرد.
- Comment: متن کامنت در داخل این ستون قرار دارد.
- Advantages: لیستی از مزایای کالا
- Disadvantages: لیستی از معایب کالا

با استفاده از ویژگی های ذکر شده، وظیفه شما این است تا مدلی طرح و آموزش دهید که بتواند تشخیص دهد آیا در نهایت یک کالا پیشنهاد می شود یا خیر. در ستون آخر دیتاست، خروجی را می توانید ببینید که به صورت لیبل های زیر است.

- Recommended
- Not Recommended

### فاز اول: پیش پردازش داده

برای استفاده از داده های متنی نیاز است تا آن ها را به فرمی که قابل پردازش است تبدیل کنید. این تبدیل به مدلی که می خواهید آموزش دهید، ارتباط مستقیم دارد.

برای مثال، اگر مدل شما یک بردار به عنوان ورودی دریافت کند، می توانید متن خود را به یک برداری از شمارش تعداد تکرار کلمات تبدیل کنید. همچنین از روش های دیگری هم مانند TF-IDF و Word Embedding می توانید استفاده کنید.

## فاز دوم: انتخاب و طراحی مدل

پیشنهاد می‌شود که از مدل‌های متفاوت استفاده کنید (غیر از شبکه عصبی) و دقت‌ها را مقایسه کنید. این کار را با استفاده از دیتای validation انجام دهید. همچنین با نمودارهای مختلف، نتایج را نمایش دهید و تحلیل کنید. پیشنهاد می‌شود که قبل از شروع به کار، درباره شبکه‌های عصبی بازگشتی (RNN) مطالعه کنید. در این پروژه، فقط مجاز به طراحی به استفاده از کتابخانه‌ی tensorflow و pytorch و scikit-learn می‌باشید.

## فاز سوم: آموزش مدل

برای آموزش مدل و بدست آوردن بهترین نتیجه، نیاز است تا مقادیر ابر پارامترها (تمامی مدل‌ها)، تعداد لایه‌ها و توابع فعال‌سازی (در شبکه عصبی) خود را تغییر دهید و بررسی کنید که با چه تغییراتی توانستید دقت خود را بهبود دهید. در گزارش پروژه تغییرات داده شده و نتیجه‌های خود را ذکر کنید و با استفاده از نمودارهای مناسب، تغییرات دقت بر اساس پارامترهای متفاوت را نمایش دهید.

نکته: استفاده از کتابخانه‌های آماده برای تشخیص احساسات مجاز نیست.

## پروژه دوفره

### هدف پروژه

هدف این پروژه، تشخیص نوع ماشین براساس تصویری از آن ماشین است.

کلاس‌های خروجی مد نظر:

- پراید
- پژو ۲۰۶
- تیبا
- پژو پارس
- سمند
- دنا

## فاز اول: جمع‌آوری و پیش‌پردازش داده

برای آموزش شبکه عصبی نیاز است تا داده جمع‌آوری کنید. این کار را می‌توان با استفاده از یک crawler به راحتی انجام داد. (پیشنهاد می‌شود که به ازای هر کلاس حداقل ۵۰۰ تا تصویر داشته باشید.)

نکاتی که در پیش‌پردازش کمک کننده است.

- تقسیم دیتاست به بخش‌های train و validation و test
- از آنجا که داده‌هایتان از اینترنت گرفته شده است، مطمئن شوید داده‌ها ابعاد یکسانی دارند. (برای مثال ۱۰۰ \* ۱۰۰)
- مطمئن شوید که کلاس‌های متفاوت، عکس‌های برابری داشته باشند تا موقع یادگیری bias رخ ندهد.
- پیشنهاد می‌شود که label‌هایتان را به صورت ترتیبی encode نکنید و از یک encoding غیر ترتیبی استفاده کنید.
- انواع تبدیل‌ها را برای داده‌ها تست کنید مانند تبدیل به grayscale.

### فاز دوم: انتخاب و طراحی مدل

پیشنهاد می‌شود که از مدل‌های متفاوت استفاده کنید و دقت‌ها را مقایسه کنید. این کار را با استفاده از دیتای validation انجام دهید و در نهایت نتایج خود را با کمک نمودار نمایش دهید و در تحلیل کردن استفاده کنید. در این پروژه، فقط مجاز به طراحی با استفاده از کتابخانه‌ی tensorflow و pytorch می‌باشید.

### فاز سوم: آموزش مدل

برای آموزش مدل و بدست آوردن بهترین نتیجه نیاز است تا مقادیر ابر پارامترها (تمامی مدل‌ها)، تعداد لایه‌ها و توابع فعال سازی (در شبکه عصبی) خود را تغییر دهید و بررسی کنید که با چه تغییراتی توانستید دقت خود را بهبود دهید در گزارش پروژه تغییرات داده شده و نتیجه‌های خود را ذکر کنید و با استفاده از نمودار تغییر در دقت بر اساس پارامترهای متفاوت را نمایش دهید.

### نکات:.

۱. دیتاست پروژه تک نفره را می‌توانید از [اینجا](#) دانلود کنید.
۲. استفاده از زبان غیر از پایتون مجاز **نمی‌باشد**.
۳. پروژه نهایی تحویل خواهد داشت.
۴. گزارش پروژه خود را در یک فایل PDF به صورت دقیق بنویسید.
۵. محتویات پروژه خود را به همراه گزارش در یک فایل زیپ با فرمت زیر در سامانه LMS آپلود کنید.

FinalProject-[Full Name]-[Student ID]

**موفق باشید**

**گروه حل تمرین**