

لیست پروژه‌های دوره کارآموزی هوش مصنوعی شرکت روشن - تابستان ۱۴۰۴

شامل ۴ پروژه:

۱- پرسش و پاسخ فارسی

هدف از این پروژه آموزش یک مدل پرسش و پاسخ به زبان فارسی است.

ابتدا عملکرد دو مدل persian_xlm_roberta_large و parsbert_question_answering_PQuAD را بر روی داده PersianQA ارزیابی کرده و سپس هر دو مدل را با استفاده از این داده fine-tune کنید.

لینک دیتاست:

[GitHub - PersianQA: Persian \(Farsi\) Question Answering Dataset \(+ Models\)](#)

۲- تشخیص برچسب‌های لباس از روی تصویر

هدف از این تسک، تشخیص برچسب‌های مربوط به یک لباس از روی تصویر آن است.

لینک دیتاست:

[Fashion Product Images Dataset](#)

در این دیتاست، ۴۴ هزار تصویر کالا‌های مختلف وجود دارد و برای هر کدام برچسب‌هایی نظیر مردانه/زنانه، رنگ، دسته بندی لباس، نوع لباس، فصل مناسب برای لباس و ... مشخص شده. هدف از این تسک پیش‌بینی این برچسب‌ها است.

در ادامه در صورت موفقیت آمیز بودن پیش‌بینی برچسب‌های لباس، می‌توان با استفاده از یک مدل بازگشتی، به پیش‌بینی متن مربوط به لباس پرداخت.

۳- شباهت سنجی تصاویر

برای آموزش چنین مدلی، روش‌های زیادی وجود دارد که دو مورد از معروف ترین آنها، آموزش به کمک triplet loss و centered loss است. از آنجایی که آموزش به کمک triplet loss به داده زیاد و توان پردازشی بالا نیاز دارد، در این تسک می‌خواهیم مدلی را به کمک centered loss آموزش دهیم و دقت آن را در یافتن تصاویر مشابه بسنجیم.

برای آموزش مدل، از دیتاستی شامل تصویر ۱۰۲ گل مختلف استفاده می‌کنیم. این دیتاست باید به دو بخش آموزش و ارزیابی شکسته شود. همچنین دیتاست ارزیابی ما باید مشابه دیتاست LFW در چهره، شامل زوج تصاویر باشد (یک زوج تصویر که لیل آنها شبیه بودن یا شبیه نبودن دو تصویر است) و دقت نهایی به صورت ROC-AUC در داده تست گزارش شود.

لینک دیتاست:

<https://www.robots.ox.ac.uk/%7Evgg/data/flowers/102>

۴- نویسه خوان اعداد دست‌نویس

نویسه خوانی اعداد دست‌نویس که به صورت عادی با دیتاست MNIST انجام می‌شود، بخش اصلی چالش نویسه‌خوانی، یعنی توانایی خواندن یک دنباله از کاراکترها را در نظر نمی‌گیرد. هدف از انجام این تسک، ساخت مدلی است که بتواند با گرفتن تصویر یک عدد چند کاراکتری، آن را تبدیل به متن کند. برای انجام این کار توصیه می‌شود دو روش متفاوت زیر پیاده‌سازی شود و با یکدیگر مقایسه شود:

۱. جدا کردن کاراکترها با روش‌های پردازش تصویر و ساخت مدلی که هر کاراکتر را جداگانه بخواند.
۲. ساخت مدل با استفاده از یک شبکه RNN، که بتواند خروجی را برای کل خط تولید کند و آموزش آن با استفاده از تابع هزینه CTC loss.

https://www.orand.cl/orand_car/ORAND-CAR-2014.tar.gz

لینک دیتاست: