

شبکه‌های عصبی پیچشی

سوال ۱

هدف این تمرین، بررسی تاثیر پارامترهای مختلف معماری شبکه‌های عصبی پیچشی^۱ در مسئله دسته‌بندی تصاویر است. برای این منظور، مدل‌های مختلف با ساختارها و تنظیمات متفاوت طراحی شده و عملکرد آن‌ها در دسته‌بندی مجموعه‌داده CIFAR-10 مورد مقایسه و تحلیل قرار می‌گیرند.

مجموعه‌داده

در این تمرین از مجموعه‌داده CIFAR-10 استفاده می‌شود. این مجموعه شامل ۶۰۰۰۰ تصویر رنگی ۳۲ در ۳۲ پیکسلی در ۱۰ کلاس مختلف است که ۵۰۰۰۰ تصویر برای آموزش و ۱۰۰۰۰ تصویر برای آزمون در آن در نظر گرفته شده‌اند.

برای هر یک از موارد زیر، مدل مورد نظر را پیاده‌سازی نموده، آموزش دهید، و نتایج نهایی شامل accuracy، precision و recall را گزارش نمایید. همچنین منحنی‌های آموزش و اعتبارسنجی (loss) را رسم کرده و تحلیل نمایید.

۱ - مدلی پایه شامل ۳ لایه کانولوشن (هر کدام به دنبال یک لایه Pooling) طراحی کرده و آموزش دهید. نتایج نهایی را گزارش نمایید.

۲ - حساسیت مدل به عمق شبکه را با انتخاب حداقل ۳ مقدار برای تعداد لایه‌ها بررسی نمایید. در هر حالت علاوه بر مقادیر دقت، میانگین گرادیان لایه اول را طی آموزش گزارش کرده و مسئله vanishing gradient را تحلیل کنید.

۳ - مدل پایه را با Dropout و Weight Decay آموزش دهید. مقادیر مختلفی برای این پارامترها انتخاب نموده و تأثیر آن‌ها را بر فاصله بین دقت آموزش و اعتبارسنجی بررسی و تحلیل نمایید.

۴ - تأثیر Batch Normalization را بررسی کنید. با افزودن این لایه به مدل پایه، عملکرد نهایی و را با حالت بدون آن مقایسه نمایید.

Convolutional Neural Networks^۱

۵ - شبکه ResNet18 را با وزن‌های از پیش آموزش دیده، بارگذاری کرده و آن را وزن‌های آن را تنظیم دقیق^۲ کنید. نتایج عملکرد بهترین مدل خودتان را با مدل ResNet18 قبل و بعد از تنظیم دقیق وزن‌ها مقایسه و نتایج آزمایشات را تحلیل کنید.

۶ - با استفاده از کتابخانه Data Augmentation (مانند torchvision.transforms)، حداقل سه روش Color Jitter، Horizontal Flip، Random Crop عملکرد مدل پایه بررسی و تحلیل نمایید.

۷ - سه استراتژی Learning Rate Scheduling را پیاده‌سازی کرده و مقایسه نمایید:

- نرخ یادگیری ثابت

- کاهش مرحله‌ای

Cosine Annealing -

عملکرد و سرعت همگرایی مدل در هر حالت را بررسی و تحلیل نمایید.

۸ - با استفاده از ابزار Grad-CAM، نقشه فعال‌سازی برای پنج تصویر تصادفی از مجموعه تست تولید کرده و تحلیل نمایید که مدل روی کدام بخش‌های تصویر تمرکز کرده و آیا این تمرکز منطقی است یا خیر.

در انجام تمرینات به نکات زیر توجه فرمایید.

۱ - پیاده‌سازی‌های کامپیوترا برای زبان برنامه‌نویسی پایتون و با بهره‌گیری از چارچوب کاری PyTorch انجام دهید.

۲ - بخش عمده‌ای از نمره تمرینات به گزارش تمرين اختصاص دارد و ارسال برنامه‌ها بدون گزارش فاقد ارزش است. در تهیه گزارش دقت نمایید که تمام اطلاعات، تصاویر و نمودارهای مورد نیاز برای اثبات پاسخ‌ها مبتنی بر آزمایشات خواسته شده در تمرین، بطور کامل و دقیق ذکر شده باشند.

۳ - مطابق قوانین دانشگاه هر نوع کپی‌برداری و اشتراک کار دانشجویان غیرمجاز است. در صورت مشاهده چنین مواردی، با طرفین شدیداً برخورد خواهد شد.

۴ - استفاده از کدها و توضیحات اینترنت یا کدها و توضیحات تولید شده با مدل‌های هوش مصنوعی

به منظور یادگیری بلامانع است، اما کمپیوتداری و انجام تمرینات توسط این ابزارها غیرمجاز است. در صورتی که از چنین ابزارهایی بهره می‌گیرید، حتماً به تمام جزئیات و نکات مرتبط با پاسخ‌ها مسلط باشید، در غیر اینصورت نمره کل تمرین را از دست خواهید داد.

۵ - مجموعه‌داده‌های مورد استفاده را به جز در مواردی که صریحاً در صورت سوال ذکر شده باشد، حتماً قبل از استفاده بطور تصادفی به سه بخش آموزشی (۷۰ درصد)، آزمون (۲۰ درصد) و اعتبارسنجی (۱۰ درصد) تقسیم نمایید.

۶ - در صورت نیاز می‌توانید سوالات خود را در خصوص پروژه از تدریسیارهای درس، به ایمیل زیر ارسال نموده یا در گروه «بله» مطرح نمایید.

ann.ceit.aut@gmail.com

۷ - فایل‌های کد و گزارش خود را در قالب یک فایل فشرده با فرمت StudentID_HW04.zip تا تاریخ ۱۴۰۴/۰۲/۱۹ فقط در بخش مربوطه در سایت درس بارگذاری نمایید. توجه نمایید، هر روز تاخیر منجر به کسر ۱۰ درصد از نمره پروژه می‌شود.