

سوالات امتیازی

سوال اول

مجموعه داده اعتبار سنجی (validation set) عملاً بخشی از مجموعه داده آموزشی (training set) بوده که برای اعتبار سنجی مدل حین یادگیری استفاده میکنیم، در صورتی که مجموعه داده تست (test set) در زمان آموزش مدل در اختیار ما نیست.

حال برای cross validation که معمولاً در روش k-fold استفاده میشود، n بار داده‌ها را به k بخش تقسیم کرده که از k-1 بخش برای آموزش و از بخش مانده باری validation استفاده میکنیم. این تکنیک زمانی استفاده میشود که داده‌گان موجود به میزانی نباشند که بتوانیم تو مجموعه اعتبار سنجی و آموزش جدا داشته باشیم.

سوال دوم

در Batch Gradient Descent ما تمام مجموعه داده آموزشی را در هر epoch استفاده میکنیم تا اول آموزش دهیم و بعد وزن‌های فعلی را آپدیت کنیم که باعث میشود در مجموعه داده بزرگ مثل ۱۰ میلیون داده، زمان آموزش را در هر epoch به نحو قابل ملاحظه‌ای زیاد کند. این روش، حالت optimal بوده اما معمولاً در سناریو واقعی، منابع مورد نیاز برای این کار را نداریم.

پس به جای آن میتوانیم از Stochastic Gradient Descent استفاده کنیم که در هر مرحله تنها یک داده آموزشی را استفاده میکند. در این روش به دلیل برداشتن تنها یک داده در هر epoch، نمودار آموزش با نوسانات نسبتاً زیادی (که بستگی به داده‌ها دارد) مواجهه میشود. از طرفی هیچ‌گاه به نقطه مینیمم محلی نمیرسد، اما در این حین روند کاهشی خواهد بود.

مشکل روش قبلی آن است که به دلیل برداشتن تنها یک داده آموزشی در هر مرحله، نمیتوان فرایند را vectorize کرده و در نتیجه سرعت کاهش میابد. برای حل این مشکل از Mini-Batch Gradient Descent استفاده میکنیم که به جای تنها یک داده، از مجموعه کوچک تری از داده استفاده میکند تا بتوان آنها را vectorize کرد و سرعت را افزایش داد. این روش مزیت‌های هر دو روش قبل را با نسبت خوبی دارد.

سوال سوم

الگوریتم Batch Normalization یک الگوریتم برای افزایش سرعت و متعادل نگه داشتن شبکه عصبی است به گونه‌ای که قبل انتقال داده‌گان به تابع فعال سازی، آنها را نرمال میکند (mean & variance). این روش کمک میکند تا اگر در وزن‌ها نوسانات زیاد (برای مثال برخی وزن‌ها بسیار بزرگ و برخی بسیار کوچک باشند) را تعدیل کند تا از overfit زود هنگام جلوگیری شود. به عبارتی خروجی این الگوریتم داده‌های با توزیع نرمال خواهند بود.

سوال چهارم

چرا از pooling استفاده کنیم؟ معمولاً تعداد زیادی از عکس‌ها (برای مثال عکس سیاه و سفید را فرض میکنیم تا توضیح آن ساده‌تر باشد...) بخش‌هایی دارد که اطلاعات زیادی به ما نمیدهند، مثل عکس سیاه و سفید یک عدد انگلیسی که بخش‌های سیاه خیلی مفید نیستند (هرچند میتوان اطلاعات موقعیت نسبی را از آنها گرفت، معمولاً از این ویژگی چشم پوشی میشود) و به علت تعداد زیاد داده‌ها معمولاً فضای زیادی اشغال کرده و باعث میشوند مدل سنگین شود. پس از pooling استفاده میکنیم تا اطلاعات بخشی از عکس را خلاصه کنیم به این صورت که هر n در n پیکسل را یک pool در نظر گرفته و با یک استراتژی مثل max و min و avg و یا یک mapping خاص ماتریس n در n آن را به یک پیکسل

سوالات امتیازی

خلاصه میکنیم. با این کار میتوان در یک یا چند لایه اندازه عکس را کوچک کرد که فواید زیادی برای ما دارد مثل سرعت پردازش و کاهش احتمال یافته شدن الگوهای ناخواسته توسط مدل.

در ادامه، شبکه CNN میتواند روابط بین پیکسل های نزدیک را درک کند در صورتی که در یک شبکه ساده MLP این کار امکان پذیر نیست. از طرفی عکس یک مفهوم دو بعدی است که شبکه های MLP نمیتوانند آنرا پردازش کرده و نیاز دارند تا آنها یک بعدی شوند که باعث میشود بخشی از اطلاعات عکس از بین برود. همچنین CNN سریع تر همگرا میشود زیرا بهتر میتواند روابط را پیدا کند، در صورتی که MLP ساده قادر به درک برخی روابط نیست و میتواند الگوهای ناخواسته دیگری که ما دنبالش نیستیم را پیدا کند.