موژان میرجلیلی – ۹۸۳۱۱۴۰

بخش اصلی:

برای دریافت دیتاست از ۲ خط کد داخل صورت پروژه استفاده شد و برای مصورسازی داده‌ها و پیش‌پردازش از keras و tensorflow استفاده شد. مراحل ۱ تا ۵ هم در قدم اول طبق دستور کار انجام شد. در مرحلهٔ feed forward طبق دستورکار ۲۰۰ داده را جدا می‌کنیم و خروجی را برای آن‌ها محاسبه می‌کنیم. برای سهولت کار تابع سیگموئید را تعریف کرده و در نهایت دقت را محاسبه می‌کنیم (وزن‌ها و بایاس‌ها هم در ابتدا مقداردهی می‌کنیم). در قدم back propagation الگوریتم gradient descend را پیاده‌سازی می‌کنیم و مشتقات را به دست می‌آوریم. وزن‌ها و بایاس‌ها هم مانند قسمت قبل مقداردهی می‌کنیم و در آخر دقت را به دست می‌آوریم. در قدم vectorization نیز سعی می‌کنیم epochها را افزایش دهیم تا سرعت اجرا را کاهش دهیم. در قدم آخر یعنی تست مدل نیز سعی کرده از همهٔ داده‌های آموزشی خود استفاده کنیم.

بخش امتیازی:

۱. مراحل پیاده‌سازی در نوت‌بوک هستند.

۲. نرمال‌سازی یک تکنیک pre-processing است که برای استانداردسازی داده‌ها استفاده می‌شود. به عبارت دیگر، داشتن منابع مختلف داده در محدوده یکسان. نرمال نکردن داده‌ها قبل از آموزش می‌تواند باعث ایجاد مشکل در شبکه ما شود و آموزش آن را به شدت سخت کند و سرعت یادگیری آن را کاهش دهد. نرمال‌سازی یک ابزار پیش پردازش داده است که برای رساندن داده‌های عددی به یک مقیاس مشترک بدون تغییر شکل آن استفاده می شود. به طور کلی، زمانی که داده‌ها را به یک ماشین یا الگوریتم یادگیری عمیق وارد می کنیم، تمایل داریم مقادیر را به یک مقیاس متعادل تغییر دهیم و متغیرها با مقادیر بزرگتر، تاثیر بیشتری نسبت به متغیرها با مقادیر کوچکتر بر خروجی مدل نمی‌گذارند. نرمال‌سازی لایه (LayerNorm) نیز تکنیکی برای نرمال‌سازی توزیع لایه‌های میانی است و به ما شیب‌های صاف‌تر، آموزش سریع‌تر و دقت تعمیم بهتر را می‌دهد.
نرمال‌سازی دسته‌ای تکنیکی برای آموزش شبکه‌های عصبی بسیار عمیق است که ورودی‌ها را در یک لایه برای هر مینی بچ استاندارد می‌کند. این امر باعث تثبیت فرآیند یادگیری و کاهش چشمگیر تعداد دوره‌های آموزشی (epoch) مورد نیاز برای آموزش شبکه های عمیق می شود. مراحل پیاده‌سازی در نوت‌بوک هستند.

۳. Dropout به طور خلاصه یعنی mute کردن تصادفی برخی از نورون‌ها به منظور خلاصه‌تر کردن شبکه. کاری که dropout انجام می‌دهد این است که به طور تصادفی نورون ها (همراه با اتصالات آن‌ها) از شبکه عصبی در طول آموزش در هر تکرار رها شود. وقتی مجموعه‌های مختلفی از نورون‌ها را رها می‌کنیم، معادل آموزش شبکه‌های عصبی مختلف است. بنابراین، روش dropout مانند میانگین‌گیری اثرات تعداد زیادی از شبکه‌های مختلف است. شبکه‌های مختلف به روش‌های مختلف بیش‌برازش خواهند کرد، بنابراین تأثیر dropout کاهش بیش‌برازش خواهد بود. همچنین، این شبکه‌ها همگی وزن‌های مشترک دارند، یعنی وزن‌ها را جداگانه برای این شبکه‌ها بهینه نمی‌کنیم. در dropout، شبکه نمی‌تواند به یک ویژگی و نورون خاص تکیه کند، بلکه باید ویژگی‌های مفیدی را بیاموزد زیرا ممکن است در این فرآیند خاموش شوند در نتیجه همه نورون‌های دیگر باید در آموزش یاد بگیرند. مراحل پیاده‌سازی در نوت‌بوک هستند.

.۴

- SGD: Keras SGD Optimizer (Stochastic Gradient Descent)

بهینه ساز SGD از gradient descent همراه با تکانه استفاده می‌کند. در این نوع بهینه ساز از زیر مجموعه‌ای از batchها برای محاسبه گرادیان استفاده می‌شود.

- RMSprop: Keras RMSProp Optimizer (Root Mean Square Propagation)

در بهینه ساز RMSProp، هدف اطمینان از حرکت ثابت میانگین مربع گرادیان‌ها است و هم چنین اینکه تقسیم گرادیان بر ریشه میانگین نیز انجام می‌شود.

- Adam: Keras Adam Optimizer (Adaptive Moment Estimation)

بهینه ساز آدام از الگوریتم آدام استفاده می‌کند که در آن از روش نزولی گرادیان تصادفی برای انجام فرآیند بهینه سازی استفاده می شود. استفاده از آن کارآمد است و حافظه بسیار کمی مصرف می‌کند. در مواردی که حجم عظیمی از داده ها و پارامترها برای استفاده در دسترس هستند مناسب است.Keras Adam Optimizer محبوب ترین و پرکاربردترین بهینه ساز برای آموزش شبکه های عصبی است.

- Adadelta: Keras Adadelta Optimizer

در بهینه ساز Adadelta از نرخ یادگیری تطبیقی با روش stochastic gradient descent استفاده می‌کند. Adadelta برای مقابله با دو اشکال مفید است؛ کاهش نرخ یادگیری مداوم در طول آموزش و همچنین مشکل نرخ یادگیری global را حل می‌کند.

- Adagrad: Keras Adagrad Optimizer

بهینه ساز Keras Adagrad دارای نرخ های یادگیری است که از پارامترهای خاصی استفاده می‌کند. بر اساس فراوانی به روز رسانی‌های دریافت شده توسط یک پارامتر، کار انجام می‌شود. حتی نرخ یادگیری با توجه به ویژگی های فردی تنظیم می‌شود. این بدان معناست که برای برخی از وزن‌ها نرخ‌های یادگیری متفاوتی وجود دارد.

- Adamax: Optimizer that implements the Adamax algorithm
- Nadam: Optimizer that implements the NAdam algorithm
- Ftrl: Optimizer that implements the FTRL algorithm

۵.

- حساسیت[۱]: یادآوری که در اصطلاح کلی به‌عنوان حساسیت شناخته می شود را می‌توان به‌عنوان نسبت موارد مثبت به درستی تعیین شده به همه مشاهدات تعریف شود و به‌عنوان معیاری برای اثربخشی سیستم در پیش بینی موارد مثبت و تعیین هزینه ها دیده شود.

$$\text{حساسیت} = \frac{\text{مثبت واقعی}}{\text{(مثبت واقعی + منفی کاذب)}}$$

- صحت[۲]: درجه صحت در تعیین نتایج مثبت ممکن است به‌عنوان صحت تعریف شود. این اساساً نسبت بین مثبت های واقعی و مجموعه کلی مثبت است. این ظرفیت مدیریتی سیستم را برای مقادیر مثبت نشان می‌دهد، اما بینشی نسبت به مقادیر منفی ارائه نمی‌کند.

$$\text{صحت} = \frac{\text{مثبت واقعی}}{\text{(مثبت واقعی + مثبت کاذب)}}$$

- امتیاز F۱ [۳]: میانگین وزنی صحت و حساسیت است. بنابراین، این اندازه‌گیری، هر دو نوع مقادیر نادرست را در نظر می‌گیرد. امتیاز F۱، زمانی که در «یک» باشد کامل در نظر گرفته می‌شود و زمانی که در «صفر» باشد شکست کامل است. در اصل میانگینی هارمونیک از ۲ مورد حساسیت و صحت است.

$$\text{امتیاز F۱} = \frac{\text{۲} \times \text{(صحت} \times \text{حساسیت)}}{\text{(صحت + حساسیت)}}$$

بخش دوم:

Data augmentation استراتژی است که به ما این امکان را می‌دهد تا به طور قابل توجهی تنوع داده های موجود برای مدل های آموزشی را افزایش دهند، بدون اینکه واقعاً داده های جدید جمع آوری کنند. تکنیک‌های افزایش داده‌ها مانند cropping، padding و horizontal flipping معمولاً برای آموزش شبکه‌های عصبی بزرگ استفاده می‌شوند. Data augmentation در بازی یک شبکه عصبی کانولوشن که می‌تواند به شدت اشیاء را طبقه بندی کند، حتی اگر گفته شود که در جهت گیری‌های مختلف قرار گرفته باشد دارای خاصیتی به نام تغییر ناپذیری است. هنگامی که تصاویر از دسته‌ای کافی نباشد، برای این که بتوانیم بهتر طبقه‌بندی کنیم با Data augmentation میزان آن‌ها را افزایش دهیم و مدل خود را بهبود بخشیم.

---

۱ Recall

۲ Precision

۳ F1 Score

Data augmentation techniques: Cropping، Flipping، Rotation Translation، Brightness، Contrast، Color Augmentation، Saturation.