

پس انتشار خطا (Backpropagation)

هیوا ابوالهادی زاده ۴۰۰۴۰۵۰۰۴

درس یادگیری ماشین

۲۷ آذر ۱۴۰۳

۱ مقدمه

پس انتشار خطا یکی از الگوریتم‌های اساسی در یادگیری عمیق است که برای آموزش شبکه‌های عصبی استفاده می‌شود. هدف از این الگوریتم به حداقل رساندن خطا و به‌روزرسانی وزن‌های شبکه به گونه‌ای است که خروجی شبکه به مقدار مطلوب نزدیک‌تر شود. هدف این گزارش بررسی پیاده‌سازی کامل از شبکه عصبی چندلایه (MLP) با استفاده از NumPy است که الگوریتم پس انتشار خطا را برای آموزش مدل بر روی مجموعه داده Fashion-MNIST پیاده و تست می‌کند.

۲ شرح الگوریتم

الگوریتم پس انتشار خطا شامل دو مرحله اصلی است:

۱. مرحله پیش‌رو (Forward Pass): در این مرحله، ورودی‌ها به شبکه داده می‌شوند و خروجی محاسبه می‌شود.

۲. مرحله پس انتشار (Backward Pass): در این مرحله، خطای خروجی محاسبه و از طریق شبکه به‌طور معکوس انتشار می‌یابد. سپس وزن‌های شبکه به‌وسیله گرادیان نزولی به‌روزرسانی می‌شوند.

فرمول محاسبه گرادیان برای هر وزن به‌صورت زیر است:

$$\frac{\partial L}{\partial w} = \frac{1}{N} \sum_{i=1}^N \frac{\partial L}{\partial y_i} \cdot \frac{\partial y_i}{\partial w}$$

که در آن L تابع هزینه است و w وزن‌های شبکه می‌باشد.

۳ اجزای اصلی کد

۱.۳ لایه‌ها

مدل شامل سه نوع لایه اصلی است:

□ لایه خطی (Linear): تبدیل ورودی با ضرب ماتریسی وزن‌ها

□ لایه فعال‌سازی ReLU: معرفی غیرخطی به مدل

□ لایه Softmax: تبدیل خروجی به احتمالات

۲.۳ تابع هزینه

از تابع هزینه Cross-Entropy برای مسائل طبقه‌بندی استفاده شده است.

۴ نتایج آموزش

□ کلاس‌های انتخابی: لباس، شلوار و کفش ورزشی

□ دقت نهایی: 98.48%

۵ نکات کلیدی پیاده‌سازی

□ استفاده از معماری شبکه عصبی با دو لایه پنهان

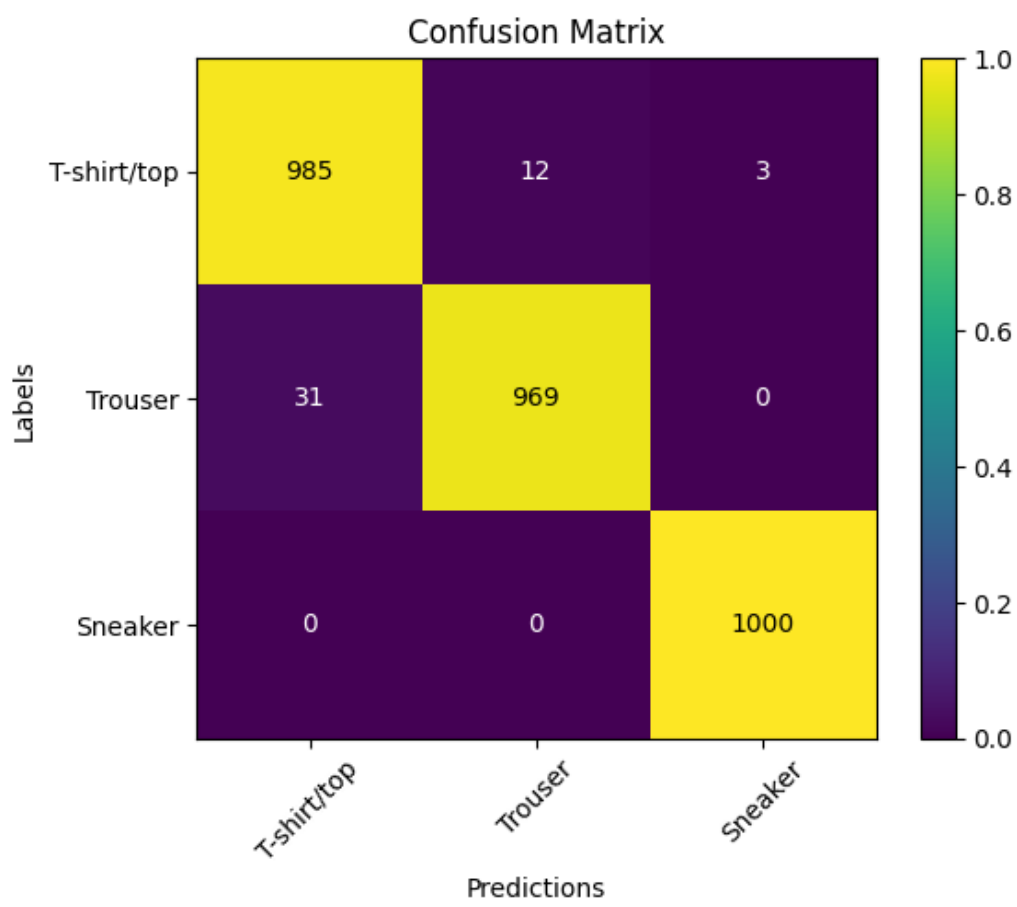
□ نرمال‌سازی داده‌ها

□ استفاده از هات کدینگ برای برچسب‌ها

□ محاسبه گرادیان با استفاده از قاعده زنجیره‌ای

۶ نتیجه‌گیری

پس‌انتشار خطا به‌عنوان یکی از روش‌های اصلی در آموزش شبکه‌های عصبی، نقش مهمی در بهبود عملکرد مدل‌ها دارد. این الگوریتم به کمک روش‌هایی مانند گرادیان نزولی، وزن‌های شبکه را بهینه می‌کند و باعث می‌شود که مدل در پیش‌بینی‌ها دقیق‌تر عمل کند.



شكل ١: confusion matrix