

درس یادگیری ماشین

گزارش تکلیف Backpropagation

استاد درس:

دکتر افتخاری

نگارش:

امیرحسین ابوالحسنی

شماره دانشجویی: ۴۰۰۴۰۵۰۰۳

پاییز 1403

فهرست مطالب

۱	مقدمه	۲
۲	پیاده سازی لایه خطی	۲
۳	پیاده سازی توابع فعال سازی	۲
۱.۳	سیگموید	۲
۲.۳	واحد یک سو شده خطی	۲
۳.۳	Softmax	۲
۴	پیاده سازی توابع هزینه	۲
۵	آموزش شبکه عصبی	۲
۶	ارزیابی و نتایج	۲

۱ مقدمه

الگوریتم پس انتشار خطا^۱، الگوریتمی برای یادگیری با نظرات در شبکه‌های عصبی با استفاده از گرادیان کاهشی است. در این روش، برای یک شبکه عصبی مصنوعی و تابع خطای مشخص، گرادیان تابع خطا نسبت به وزن‌های شبکه عصبی محاسبه می‌شود. در این تکلیف به پیاده سازی بلوک‌های سازنده یک شبکه عصبی پراخته می‌شود، و در هر بلوک، متدهای مورد نیاز برای انجام الگوریتم پس انتشار خطا پیاده سازی می‌شود.

۲ پیاده سازی لایه خطی

هر لایه از شبکه عصبی متشکل از تعدادی نورون می‌باشد که تعداد بعد ورودی را به تعداد بعد خروجی نگاشت می‌کند. پارامترهای مهمی که باید برای هر لایه ذخیره شود وزن‌های لایه و بایاس می‌باشد. همچنین گرادیان‌ها نسبت به وزن و بایاس نیز باید نگه داشته شود.

۳ پیاده سازی توابع فعال سازی

۱.۳ سیگموید

تابع سیگموید به فرمول :

$$f(x) = \frac{1}{1 + e^{-x}}$$

ورودی را به بازه $[0, 1]$ نگاشت می‌کند و برای انحام دسته‌بندی دو کلاسه مورد استفاده قرار می‌گیرد. یکی دیگر از دلایل استفاده از سیگموید، سادگی در محاسبه مشتق آن است.

$$f'(x) = f(x) \cdot (1 - f(x))$$

۲.۳ واحد یک سو شده خطی

تابع ReLU با فرمول

$$f(x) = \max(0, x)$$

سعی در ایجاد روابط غیر خطی در شبکه دارد. همچنین مشتق این تابع به سادگی محاسبه می‌گردد:

$$f'(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

۳.۳ Softmax

۴ پیاده سازی توابع هزینه

۵ آموزش شبکه عصبی

۶ ارزیابی و نتایج

^۱ Back Propagation