

### گزارشکار ۳ - امیرحسام نخئی - حسن فرهود

سوال (۱)

مدل پرسپترون یک مدل ساده از شبکه عصبی است که برای مسائل دسته‌بندی باینری استفاده می‌شود. این مدل دارای محدودیت‌هایی است که آن را مناسب برای مسائلی با تنوع بالا نمی‌کند. محدودیت‌های مهم مدل پرسپترون عبارتند از:

۱. مسائل خطی: پرسپترون تنها می‌تواند مسائلی را که توسط یک سطح تصمیم‌گیری خطی قابل حل باشند، حل کند. این به این معناست که اگر داده‌ها را نمی‌توان با یک خط از هم جدا کرد، پرسپترون نمی‌تواند آن مسئله را به درستی حل کند.

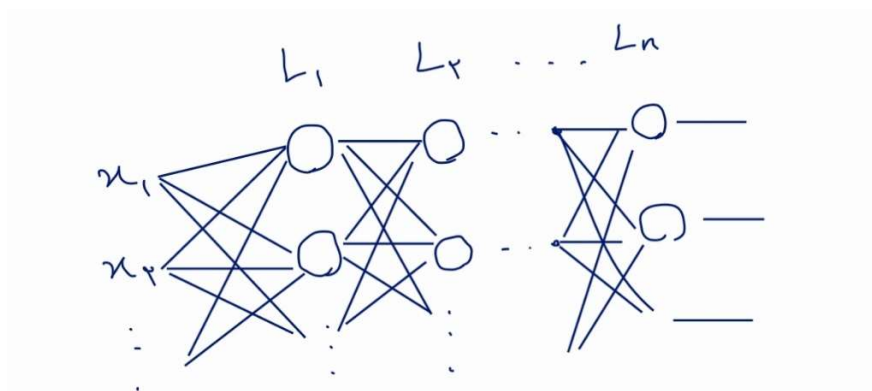
۲. تنها برای باینری: پرسپترون تنها برای مسائل دسته‌بندی دو دسته‌ای قابل استفاده است. این به این معناست که اگر تعداد دسته‌ها بیشتر از دو باشد، نیاز به استفاده از روش‌های پیچیده‌تری مانند شبکه‌های عصبی چندلایه دارید.

۳. عدم توانایی در مدل‌سازی مسائل پیچیده: پرسپترون نمی‌تواند مسائل پیچیده‌تری مانند توالی‌های زمانی یا تنظیم‌های پیچیده را مدل کند. برای مدل‌سازی چنین مسائلی نیاز به مدل‌های مختلفی مانند شبکه‌های عصبی با تعداد لایه‌ها و ویژگی‌های بیشتر دارید.

۴. حساسیت به انتخاب ویژگی‌ها: پرسپترون حساس به انتخاب ویژگی‌ها است. اگر ویژگی‌های مناسبی برای توصیف داده‌ها انتخاب نشود، پرسپترون نمی‌تواند به درستی عمل کند.

در کل، پرسپترون یک مدل ساده است که برای مسائل ساده دسته‌بندی مناسب است، اما برای مسائل پیچیده‌تر و چالش‌برانگیزتر نیاز به مدل‌های پیچیده‌تر و قدرتمندتری دارید.

سوال ۲)



سوال ۳)

الگوریتم پس‌انتشار خطا (Backpropagation) یکی از مهمترین الگوریتم‌های آموزش شبکه‌های عصبی است. این الگوریتم به شبکه عصبی اجازه می‌دهد تا وزن‌های خود را به گونه‌ای تنظیم کند که خطا در پیش‌بینی‌ها کاهش یابد. در ادامه، مراحل الگوریتم پس‌انتشار خطا به تفصیل شرح داده می‌شود:

۱. مقداردهی اولیه وزن‌ها: در ابتدا، وزن‌های شبکه عصبی به صورت تصادفی یا با استفاده از روش‌های خاصی مقداردهی اولیه می‌شوند.

۲. فاز فرستادن (Forward Pass): داده ورودی از طریق شبکه عصبی منتقل می‌شود. این مرحله شامل محاسبه خروجی‌های شبکه برای هر نمونه ورودی است.

۳. محاسبه خطا: برای هر نمونه ورودی، خطای مدل از معیار میانگین مربعات خطا محاسبه می‌شود. خطا نشان‌دهنده تفاوت بین خروجی واقعی و پیش‌بینی شده توسط مدل است.

۴. فاز پس‌انتشار (Backward Pass): در این مرحله، خطا به عقب منتقل می‌شود. این به این معناست که خطای محاسبه شده به هر لایه وزن در شبکه عصبی منتقل می‌شود. این انتقال خطا با استفاده از قاعده زنجیره‌ای انجام می‌شود.

۵. تنظیم وزن‌ها: وزن‌های هر لایه براساس خطایی که به آن لایه منتقل شده است تنظیم می‌شوند. این تنظیم با استفاده از یک الگوریتم بهینه‌سازی (مانند نسبت یادگیری) انجام می‌شود.

۶. تکرار مراحل ۲ تا ۵: این مراحل به تکرار انجام می‌شوند تا وزن‌ها به گونه‌ای تنظیم شوند که خطا به حداقل برسد. این تکرارها به عنوان (epochs) شناخته می‌شوند.

۷. ارزیابی و تست: پس از آموزش مدل، آن را بر روی داده‌های ارزیابی یا تست اجرا می‌کنیم تا عملکرد نهایی مدل را ارزیابی کنیم.

این مراحل به صورت تکراری ادامه می‌یابند تا مدل به یک عملکرد مطلوب برسد. الگوریتم پس‌انتشار خطا به شبکه‌های عصبی این امکان را می‌دهد تا ویژگی‌های مناسب را از داده‌ها استخراج کرده و مسائل مختلف را حل کنند.

## شرح آزمایش :

ابتدا کتابخانه numpy را اضافه می‌کنیم. سپس به تعریف کلاس MLP می‌پردازیم. مقادیر سائز ورودی، سائز لایه پنهان، سائز خروجی، و Learning Rate گرفته می‌شود. ماتریس‌های  $w_1$  و  $w_2$  ابتدا به صورت مقادیر تصادفی و بردارهای  $bias_1$  و  $bias_2$  به صورت صفر تعیین می‌شوند. سپس به تعریف توابع سیگموئید و مشتق سیگموئید می‌پردازیم. در ادامه تابع fit را تعریف می‌کنیم که به عنوان ورودی، ماتریس ورودی، ماتریس خروجی (label)، Learning Rate، و تعداد epochs را می‌گیرد. در این تابع، در قسمت feedforward، ماتریس ورودی و  $w_1$  در همدیگر ضرب می‌شوند و به تابع فعالساز لایه اول (sigmoid) داده می‌شوند و خروجی لایه اول معین می‌شود. سپس خروجی لایه اول در  $w_2$  ضرب می‌شود و به تابع فعالساز لایه دوم داده می‌شود و خروجی لایه دوم که خروجی کل شبکه است محاسبه می‌شود. در ادامه مرحله backpropagation فرا میرسد. مقدار ارور و تغییرات پارامترها محاسبه می‌شوند و در مرحله بعدی آپدیت می‌شوند. سپس تابع predict تعریف می‌شود که ورودی آن ماتریس ورودی است و در حقیقت وزن‌ها و بایاس‌های بهینه شده را در feedforward قرار می‌دهد. بعد از تعریف کلاس MLP، ماتریس ورودی و ماتریس خروجی (label) را معین می‌کنیم و یک اینستنس از کلاس MLP تعریف می‌کنیم و متد fit و predict را روی آن اجرا می‌کنیم. در انتها صحت شبکه مورد بررسی قرار می‌گیرد که میزان accuracy ۱ می‌باشد که نشان دهنده این است که مقادیر وزن‌های شبکه صحت خوبی داشته اند.