

شبکه‌های چند لایه پرسپترون

مجید نصیری منجیلی

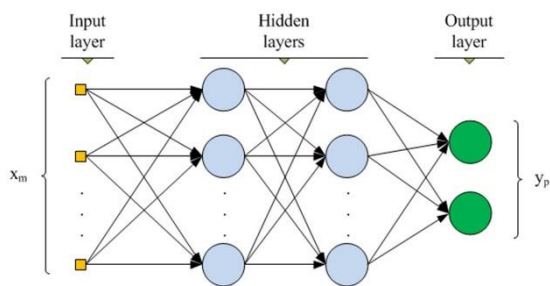
دانشکده مهندسی کامپیوتر، دانشگاه تربیت دبیر شهید رجایی

majid.nasiri@srttu.edu

چکیده

در این گزارش با استفاده از شبکه چندلایه پرسپترون به کلاس‌بندی دیتاست‌های دو کلاسه iris و satimage پرداخته می‌شود. در این گزارش به بررسی تاثیر تعداد نورون‌ها و لایه‌های مخفی بر کارایی شبکه می‌پردازیم. این گزارش بر مبنای تمرین درس یادگیری ماشین می‌باشد.

به کلاس‌بندی دیتاست‌های گاسین دو کلاسه، iris و satimage پرداخته می‌پردازیم.



شکل ۱: شبکه چندلایه پرسپترون

در ادامه در بخش ۲ توضیحات پیاده سازی و شرایط تست، در بخش ۳ نتایج بدست آمده برای تست‌های مختلف و در نهایت در آخرین بخش نتیجه گیری را خواهیم داشت.

۲- پیاده‌سازی

شبکه چندلایه پرسپترون (شکل ۱) از تعدادی ورودی $X = \{X_1, X_2, \dots, X_m\}$ ، لایه‌های مخفی و خروجی $Y = \{y_1, y_2, \dots, y_p\}$ تشکیل شده است. در این شبکه ورودی‌ها حاوی m ویژگی می‌باشند و خروجی‌ها نیز به p کلاس متفاوت تقسیم می‌شود. ضمناً برای آموزش این شبکه و تنظیم وز-ن‌های آن از الگوریتم پس انتشار^۴ بهره گرفته شده است. با توجه به [1] به روش‌های زیر می‌توان تعداد نورون‌ها را در لایه مخفی بطور تخمینی محاسبه کرد.

۱- مقدمه

در اواخر سال ۱۹۴۰ میلادی پیشگامان علم شبکه‌های عصبی، مک کلوچ و پیتس در رابطه با توانایی ارتباط درونی یک نورون مطالعاتی را انجام دادند. آنها به یک مدل محاسباتی بر مبنای یک عنصر شبکه نورونی ساده ارائه نمودند. در سال ۱۹۵۸ روزبلا^۱ پرسپترون^۲ را مطرح کرد و همچنین در سال ۱۹۷۱ وربوس^۳ یک الگوریتم پس انتشار را در رساله دکتری خود منتشر کرد.

کلاس‌بندی یکی از مهمترین مسائلی است که در اکثر در موضوعات هوش مصنوعی با آن مواجه می‌شویم، روش‌های مختلفی برای کلاس‌بندی دیتاها وجود دارد که یکی از مهمترین آنها استفاده از شبکه‌های چندلایه پرسپترون می‌باشد. در این گزارش ما با استفاده از شبکه چندلایه پرسپترون

^۳ Werbos

^۴ Back-Propagation

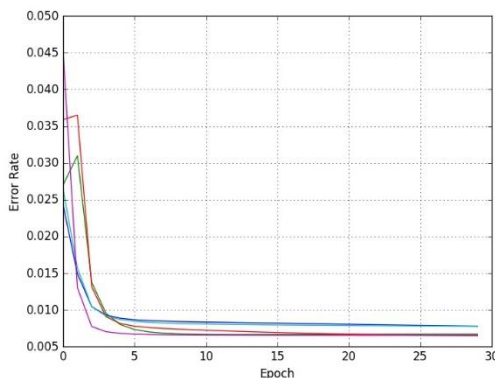
^۱ Rosenblatt

^۲ Perceptron

۳-۲- دیتاست iris

ما با کلاس‌بندی این دیتاست با شبکه‌های پرسپترون متناوبی که طراحی کرده‌ایم و با تعداد ۱۰۰ تکرار در تست‌های مختلف به خطای کلاس-بندی حدود ۱٪ رسیده‌ایم.

در یک مرحله یک شبکه پرسپترون با تعداد یک لایه مخفی و تعداد نورونهای ۲،۶،۱۰ و ۱۴ نورون



شکل ۲: نمودار خطای کلاس‌بندی دیتای دوکلاسه با توزیع گاسین

برای این لایه طراحی و تست شد، نتایج کلاس-بندی در شکل ۳ آورده شده است. در مرحله بعد با افزایش تعداد لایه‌های مخفی به تعداد ۱،۲ و ۳ شبکه‌ها را آموزش داده و نتایج بدست آمده در شکل ۴ آمده است.

البته با توجه به روش اول، تعداد نورون‌های لایه مخفی از رابطه‌ی ۱ بدست می‌آید.

$$\begin{aligned} \text{Neuron Number} &= \left(\frac{2}{3} * 4 + 3 \right) \\ &= 5.66 \\ &\approx 6 \end{aligned} \quad (2)$$

که در این پیاده‌سازی برای حالتی که تنها یک لایه مخفی داریم ۶ نورون و همین تعداد نورون برای لایه‌های مخفی اضافه شده (تا سه لایه ۶ نورونی) هم قرار داده می‌شود.

- تعداد نورون‌ها برابر با ۲/۳ (بین ۷۰٪ تا ۹۰٪) تعداد نورون‌های لایه ورودی، در صورت کافی نبودن اضافه کردن به تعداد نورون‌های خروجی
- تعداد نورون‌ها باید کمتر از دو برابر نورون‌های لایه ورودی باشد.
- تعداد نورون‌ها مابین تعداد نورون‌های ورودی و خروجی

که در این گزارش برای تعیین تعداد نورون‌ها از معیار اول استفاده شده است.

۳- نتایج

کلاس‌بندی دیتاست‌های نامبرده با لایه‌های مختلف و تعداد نورون‌های متفاوت در لایه‌ها، پیاده‌سازی و نتایج متفاوتی بدست آمد.

۳-۱- دیتای دوکلاسه با توزیع گاسین

با توجه به روش اول، تعداد نورون‌های لایه مخفی از رابطه‌ی ۱ بدست می‌آید.

$$\begin{aligned} \text{Neuron Number} &= \left(\frac{2}{3} * 2 + 1 \right) \\ &= 2.33 \\ &\approx 2 \end{aligned} \quad (1)$$

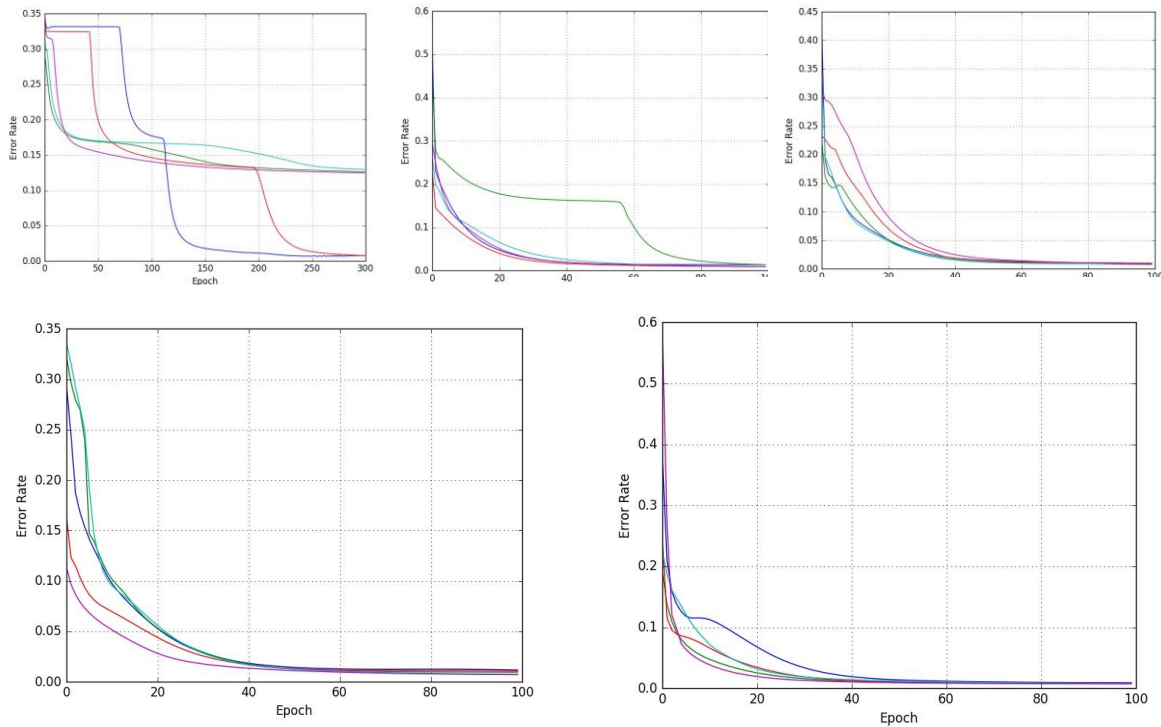
بنابراین برای کلاس‌بندی این دیتا یک شبکه پرسپترون با یک لایه مخفی شامل دو نورون مخفی کفایت می‌کرد. شکل ۲ میزان دقت بدست آمده برای تعداد تکرار^۵ دیتا نشان می‌دهد. همانطور که در شکل ۲ دیده می‌شود با حدود ۲۰ بار تکرار دیتا جهت آموزش شبکه، شبکه به خطای کلاس‌بندی کمتر از ۱٪ رسیده است. در شکل ۲ نمودارهای با رنگ‌های متفاوت تست‌های بیشتر و رسیدن به میزان دقت مشابه را نشان می‌دهد.

^۵ epoch

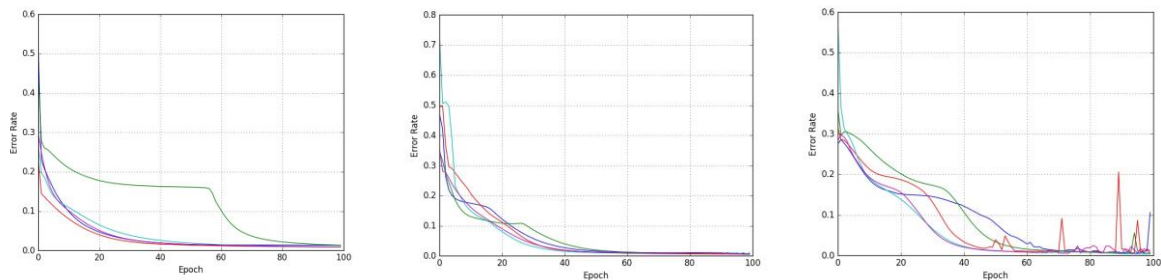
۳-۳- دیتاست satimage

همچنین برای این دیتاست شبکه چندلایه پرسپترون طراحی و تست شد که نتایج بدست

آمده برای تست های متوالی با رنگهای متفاوت در شکل ۵ آورده شده است.



شکل ۳: نمودار خطای کلاس‌بندی دیتای iris با یک لایه مخفی. (بالا چپ) تعداد ۲ نورون. (بالا وسط) تعداد ۶ نورون. (بالا راست) تعداد ۱۰ نورون. (پایین چپ) تعداد ۱۴ نورون. (پایین راست) تعداد ۱۸ نورون.



شکل ۴: نمودار خطای کلاس‌بندی دیتای iris با یک تعداد لایه‌های مخفی متفاوت. (چپ) تعداد ۱ لایه مخفی. (وسط) تعداد ۲ لایه مخفی. (راست) تعداد ۳ لایه مخفی.

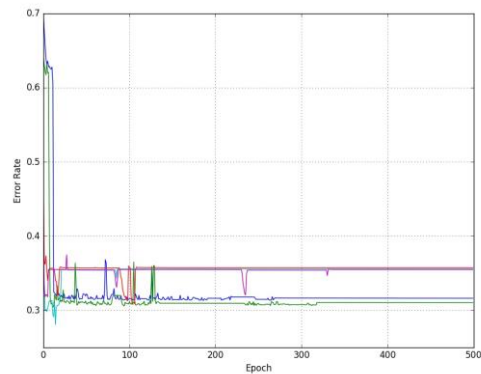
۴- نتیجه‌گیری

در صورتی که دیتای ورودی به شبکه، به صورت خطی تفکیک پذیر باشد، شبکه چند لایه پرسپترون به خطای صفر همگرا می‌شود. ولی در مواردی دیگر که دیتای ورودی به صورت خطی

تفکیک پذیر نیست می‌توان با تنظیم تعداد لایه‌های مخفی و تعداد نورون‌های هر لایه به کمترین خطای کلاس‌بندی دست یافت. در حقیقت هیچ فرمول بسته‌ای برای انتخاب تعداد لایه‌های مخفی و تعداد نورون‌های هر لایه وجود ندارد.

مراجع

[1] Karsoliya, Saurabh. "Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture." *International Journal of Engineering Trends and Technology* 3.6 (2012): 714-717.



شکل ۵: نمودار خطای کلاسیک دیتای iris با یک لایه مخفی در تست‌های مختلف

با توجه به نتایج تست‌های انجام شده می‌توان نتیجه گرفت که در حالتی که تنها یک لایه مخفی داریم انتخاب تعداد نورون‌های آن از روش اول ذکر شده در [1] متناسب و بهینه می‌باشد. در واقع انتخاب تعداد نورون‌های کمتر باعث افت راندمان و افزایش آن جز سربار پردازشی بیشتر فایده‌ای ندارد. همچنین اضافه کردن لایه‌های بیشتر تا حدی باعث کاهش خطا می‌شود، ولی بیش از آن کاهش زیادی در خطا نخواهیم داشت.

از دیگر مواردی که دیده می‌شود این است که در شبکه‌های با تعداد لایه و نورون بیشتر، نمودار میزان خطای شبکه در تکرارهای مختلف نوسان بیشتری نسبت به موقعی که تعداد لایه‌ها و نورون‌ها کمتر است، دارد.

همچنین همان‌طور که در شکل ۳ دیده می‌شود، در شبکه‌ای که تنها یک لایه مخفی دارد وقتی که تعداد نورون‌های این لایه کمتر از معیار ذکر شده باشد (۲ نورون کمتر از حداقل تعداد نورون)، نیاز به تکرار بیشتر آموزش برای همگرایی خطای خروجی به حداقل را دارد. و بالعکس با افزایش تعداد نورون همگرایی سریعتر اتفاق می‌افتد.