

نحوه ی عملکرد شبکه عصبی:

کار با هر شبکه عصبی شامل سه مرحله می باشد: آموزش، ارزیابی و تست.

در مرحله آموزش، شبکه الگوهای موجود در داده های ورودی را یاد می گیرد. هر شبکه عصبی برای یادگیری از قانون خاص استفاده می کند.

ارزیابی، قدرت شبکه عصبی در ایجاد پاسخ های قابل قبول برای ورودی هایی است که عضو مجموعه آموزشی نبوده اند.

در مرحله تست نیز شبکه عصبی برای عملکردی که به آن منظور طراحی گردیده است، استفاده می شود. هدف از آموزش شبکه، رسیدن به شرایطی است که شبکه قادر به پاسخگویی صحیح به داده های ارائه شده در آموزش شبکه (به خاطر سپردن) و همچنین داده های مشابه و متفاوت از ورودی هایی که از آنها برای آموزش شبکه استفاده شده است (تعمیم دادن) باشد.

برتری عمده شبکه های عصبی آموزش داده شده بر محاسبات کلاسیک این است که نتایج مورد نیاز با تلاش کمتر و در زمان کمتری قابل حصول است. در نتیجه این مزایا خصوصاً برای مسائلی که مستلزم محاسبات طولانی هستند بسیار مفید و موثر واقع گردد.

برای آنکه بتوانیم از شبکه عصبی استفاده کنیم باید یک سری نمونه یا دیتاست داشته باشیم این دیتاست ها شامل یک سری پارامترهای ورودی هستند. به طور مثال ما می خواهیم دمای هوای تهران را پیش بینی کنیم. یک سری پارامترها مانند میزان رطوبت، میزان تابش خورشید، دمای هوا در روز قبل می توانند عوامل تاثیر گذار و مربوط به میزان دمای هوای امروز تهران باشند. هر نمونه شامل مقادیر برای این ۳ پارامتر مذکور و مقدار خروجی مورد نظر هستند. خروجی مورد نظر همان دمای هوای تهران هستند. به طور مثال یک نمونه به این صورت است که میزان رطوبت برابر با ۴۵ درصد، میزان تابش خورشید ۳۵ درصد و دمای دیروز تهران ۱۸ بوده است و دمای امروز تهران ۲۰ است. مقدار ۲۰ خروجی این نمونه می باشد. شبکه عصبی با دیدن این نمونه ها آموزش می بیند اما در طی فرآیند آموزش چه اتفاقی رخ می دهد. و نتیجه ی آموزش چه باید باشد؟ نتیجه ی آموزش این است که ما یک نمونه یعنی برای پارامترهای میزان رطوبت، میزان تابش خورشید و دمای دیروز تهران به شبکه عصبی مقدار می دهیم و این شبکه بعد آموزش باید به

ما بگوئید دمای هوای تهران با توجه به این پارامترها چند است. اما اینکه در طی فرآیند آموزش چه اتفاقی رخ می دهد در ادامه صحبت می کنیم.

همانطور که قبلا گفته ایم، تعداد ورودی های شبکه عصبی برابر با تعداد پارامترهای مساله هستند و لایه ی ورودی را تشکیل می دهند. لایه ی خروجی هم مربوط به خروجی مساله می باشد. وقتی شبکه عصبی نمونه ای را می بیند مقدار پارامترها در هر نرون با یکدیگر جمع می شوند. البته مقدار هر پارامتر در وزن مربوط به اتصال آن نرون با آن ورودی، ضرب می شود. این مقدار عددی را با w بر روی اتصالات نشان داده ایم. بنابراین در نرون خواهیم داشت:

$$y = x_1 \times w_1 + x_2 \times w_2 + x_3 \times w_3$$

در فرمول بالا، x_i ها همان مقدار پارامتر ورودی هستند، و w_i ها وزن ها. وزن مشخص کننده ی میزان تاثیر آن ورودی بر خروجی می باشد. البته در فرمول بالا در حالت کلی مقدار بایاس هم باید اضافه شود.. بعد از اینکه این مقدار بدست آمد یک مرحله با نام تبدیل بر روی آن اعمال می شود که این مرحله را هم باید در یک مبحث کامل توضیح بدهیم، اتفاق می افتد. در نهایت مقداری برای خروجی که همان دمای هوای تهران است بدست می آید.

اگر بخواهیم فرآیند آموزش شبکه عصبی را دقیق تر بررسی کنیم ابتدا یک نرون از شبکه عصبی را در نظر بگیرید. از هر ورودی (هر ورودی یک ویژگی از نمونه است) یک مسیر به نرون وجود دارد. بنابراین با در نظر گرفتن ورودی ها و یک نرون از لایه ی پنهان، مسیری از تمام ورودی ها به نرون مورد نظر خواهیم داشت. همانطور که در بخش های قبلی بیان شد، در شبکه عصبی به هر مسیر وزنی نسبت داده می شود که این وزن قدرت اتصال آن مسیر را نشان می دهد. بنابراین هر ورودی همراه با وزنی که به آن نسبت داده می شود وارد این نرون می شوند و همراه با بایاس با یکدیگر جمع می شوند. سپس با استفاده از یک تابع تبدیل مقدار این حاصل جمع به فرم دیگری تبدیل می شود. در حقیقت اتفاقی که در یک نرون می افتد به صورت زیر است:

$$a = f(W * X + b)$$

در فرمول بالا a خروجی نرون و F تابع تبدیل مورد استفاده در نرون هستند. با توجه به فرمول بالا شبکه عصبی ورودی های را می بیند و به خروجی تبدیل می کند.

تنظیم وزن ها و بایاس به سمتی است که شبکه عصبی بتواند خروجی نزدیک به خروجی واقعی تولید کند. بنابراین در حین فرآیند آموزش مقادیر وزن ها و بایاس باید تنظیم شوند. شبکه عصبی در حین فرآیند آموزش با مقایسه ی خروجی خود و خروجی واقعی و با کاهش فاصله ی این دو مقدار، به سمت مقدار بهینه ی وزن ها و بایاس ها پیش می رود. در حقیقت شبکه عصبی با استفاده از فرمول خطای زیر این مقادیر بهینه را جستجو می کند:

$$c = \frac{1}{N} \sum_{i=1}^N (f(x_i) - y_i)^2$$

در فرمول بالا y مقدار خروجی واقعی و $f(x)$ مقدار خروجی شبکه عصبی برای هر نمونه را نشان می دهند. در حقیقت فرآیند آموزش شبکه عصبی را می توان به اینگونه توصیف کرد که با کم کردن میزان خطا، مقدار بهینه ای برای مقادیر وزن ها و بایاس ها تعیین می شود. مطالب بالا ساده ترین شکل فرآیند آموزش و ساختار شبکه عصبی بوده است. در ادامه به بررسی جزئیات بیشتری می پردازیم.

برای تشریح ساختار شبکه عصبی و تعداد وزن ها و بایاس ها در ادامه توضیحاتی ارائه می دهیم. در ابتدا یک لایه ی پنهان با یک نرون در نظر می گیریم و تعداد ویژگی های مربوط به داده ها را R در نظر می گیریم. برای هر نرون در لایه ی اول، به تعداد ویژگی های ورودی، پارامتر وزن به همراه یک پارامتر بایاس باید تنظیم شود. با این توصیف برای یک نرون در لایه ی اول و برای یک دیتا ست با R ویژگی، از آنجاییکه یک مسیر از هر ویژگی ورودی ها به نرون وجود دارد، پس در اینجا R مسیر خواهیم داشت، علاوه بر این یک بایاس نیز به این نرون وارد می شود. بنابراین به تعداد $R+1$ پارامتر باید تنظیم شود. اگر در لایه ی اول C نرون داشته باشیم، تعداد پارامترهایی که شبکه عصبی باید برای این لایه تنظیم کند، برابر با $C*R+C$ خواهد بود.

برای لایه ی دوم نیز اتصالات کاملی از هر نرون لایه ی اول به هر نرون لایه ی دوم وجود خواهد داشت. بنابراین اگر در لایه ی اول C نرون و در لایه ی دوم D نرون داشته باشیم، تعداد پارامترهایی که باید در لایه ی دوم تنظیم شوند برابر با $C*D+D$ پارامتر خواهند بود. این روال تا انتها ادامه پیدا می کند. بنابراین تعداد پارامترهایی که باید تنظیم شوند بر اساس تعداد لایه ها و نرون های شبکه عصبی و تعداد ویژگی های داده های مورد استفاده خواهند بود.

یک مساله و شبکه عصبی با ویژگی های زیر را در نظر بگیرید:

- داده ها شامل ۲۰ ویژگی هستند
- شبکه عصبی دارای دو لایه ی پنهان
- در لایه ی اول هفت نرون و در لایه ی دوم چهار نرون در نظر گرفته شده است
- لایه ی خروجی دارای یک نرون می باشد

تعداد متغیرهای مربوط به اتصال ورودی های به لایه ی اول برابر خواهند بود با:

تعداد نرون + تعداد نرون های لایه ی اول * تعداد ویژگی ها = تعداد متغیرهای لایه ی ورودی به لایه ی اول
ها در لایه ی اول

با توجه به فرضیات مساله تعداد متغیرهای مربوط به اتصالات ورودی به لایه ی اول برابر با $7+7*20$ ،
برابر با ۱۴۷ متغیر، خواهد بود.

تعداد متغیرها مربوط به اتصالات لایه ی اول به لایه ی دوم برابر خواهد بود با:

تعداد نرون های * تعداد نرون های لایه ی اول = تعداد متغیرها مربوط به اتصال لایه ی اول به لایه ی دوم
تعداد نرون های لایه ی دوم + لایه ی دوم

با توجه به فرضیات مساله برای مسیر دوم یعنی اتصالات لایه ی اول به لایه ی دوم تعداد متغیرها برابر با
 $4+4*7$ ، برابر با ۳۲ متغیر، خواهد بود.

تعداد متغیرهای مربوط به اتصال لایه ی دوم به لایه ی خروجی:

تعداد * تعداد نرون های لایه ی دوم = تعداد متغیرهای مربوط به اتصالات لایه ی دوم به لایه ی خروجی
تعداد نرون های لایه ی خروجی + نرون های لایه ی خروجی

با توجه به فرضیات مساله تعداد متغیرهای مسیر آخر برابر با $1+1*4$ ، برابر با ۵ متغیر، خواهد بود. البته

تعداد بایاس قابل تنظیم هست و می توانیم برای هر لایه ی مورد نظر بایاس را در نظر گرفت یا آن را

حذف کرد. با توجه به محاسبات بالا برای دیتا ست مفروض با شبکه عصبی توصیف شده برابر با

$6+32+147$ برابر با ۱۸۵ متغیر باید بهینه یابی شوند.

در جلسه ی بعد بحث انواع آموزش با ناظر و بی ناظر و توابع انتقال را بررسی خواهیم کرد.

www.matlabproject.ir