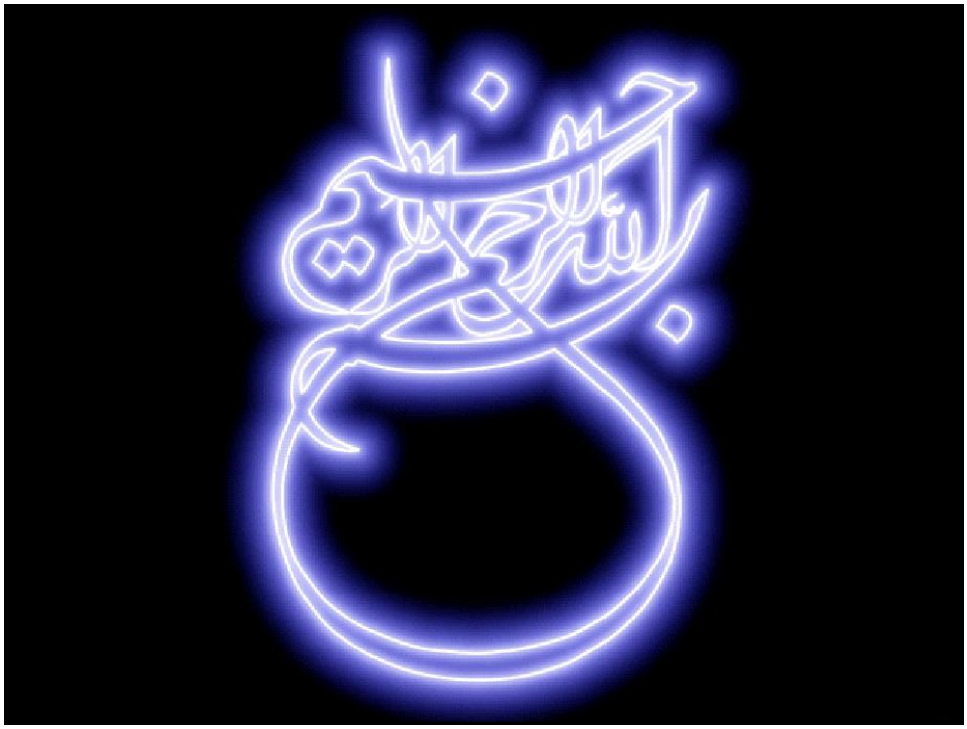
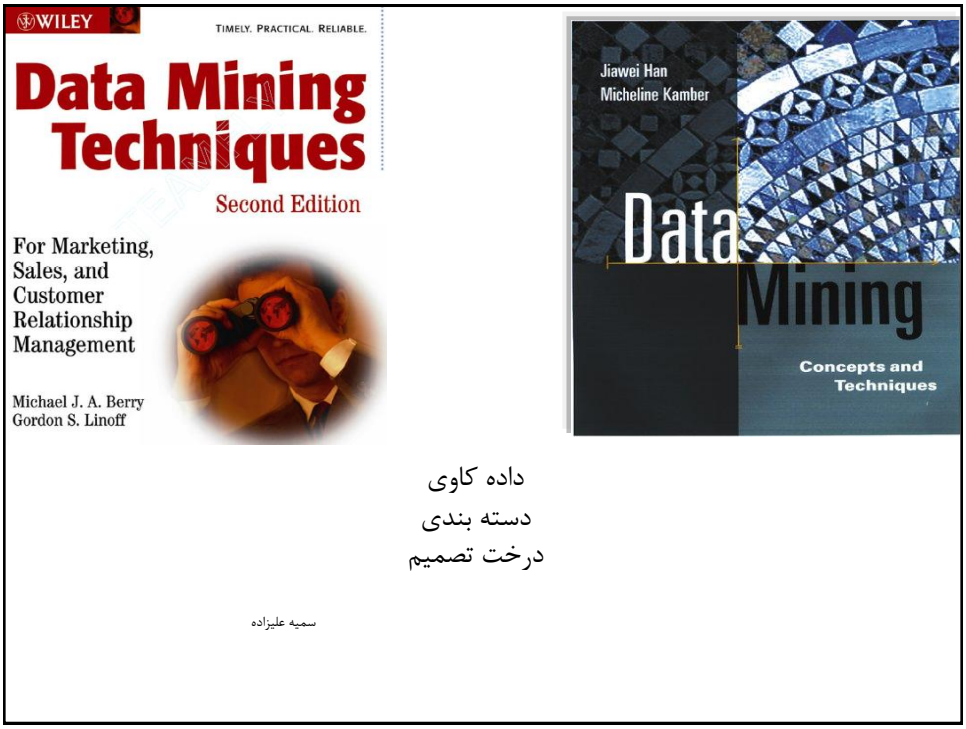



WILEY
TIMELY. PRACTICAL. RELIABLE.
Data Mining Techniques
Second Edition
For Marketing, Sales, and Customer Relationship Management
Michael J. A. Berry
Gordon S. Linoff
Jiawei Han
Micheline Kamber
Data Mining
Concepts and Techniques
داده کاوی
دسته بندی
درخت تصمیم
سمیه علیزاده

شبکه های عصبی

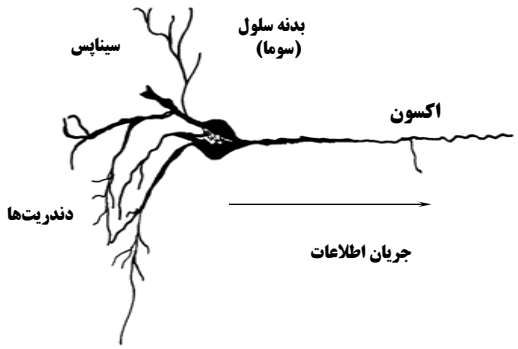
سمیه علیزاده هیات علمی دانشکده صنایع
دانشگاه خواجه نصیر طوسی

تعریف

- سلولهای مغز انسان دارای ساختار متفاوتی از سایر سلولهای بدن انسان می‌باشند به این سلولهای مغزی نرون گفته می‌شود. هر نرون یک بدنه، یک آکسون و چندین دندریت داشته و واسط بین آکسون یک نرون و دندریتهای نرونهای دیگر سیناپس نام دارد. همچنین هر نرون بر اساس یک آستانه تحریک در یکی از دو وضعیت تحریک شده و ساکن قرار می‌گیرند.

سمیه علیزاده هیات علمی دانشکده صنایع
دانشگاه خواجه نصیر طوسی

ساختار سلول عصبی



سمیه علیزاده هیات علمی دانشکده صنایع
دانشگاه خواجه نصیر طوسی

سلولهای عصبی مغز انسان

- این ساختار نرون در مغز انسان به تعداد $۱۰^{۱۱}$ تکرار می‌شود و از آنجا که هر نرون حداقل به ۱۰۰۰۰ نرون دیگر متصل می‌باشد، در مغز انسان $۱۰^{۱۵}$ اتصال سیناپسی وجود دارد که تمامی فعالیتهای ذهنی را به انجام می‌رسانند.

سمیه علیزاده هیات علمی دانشکده صنایع
دانشگاه خواجه نصیر طوسی

استفاده از ویژگیهای سلول عصبی

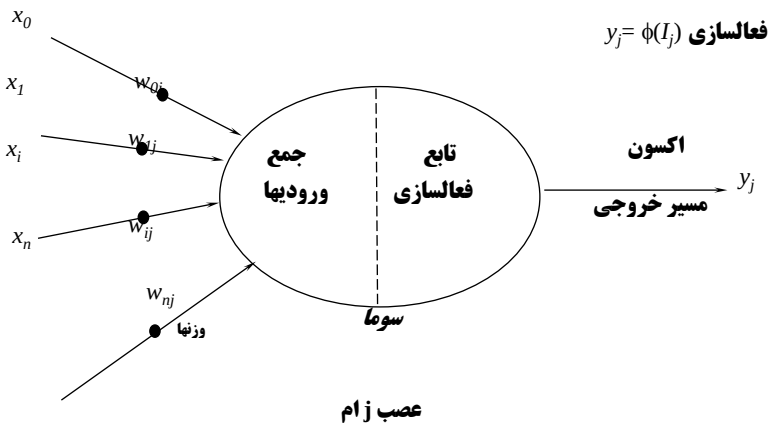
- مجموعه‌ای از گره‌ها (واحدها، نرونها، عناصر محاسباتی)
- هر گره ورودی و خروجی دارد.
- هر گره بر اساس تابعی خاص محاسبه ساده‌ای انجام می‌دهد.
- بین گره‌ها، اتصالات موزون وجود دارد.
- اتصالات بر اساس معماری شبکه مشخص می‌شوند.
- نتیجه یک شبکه تابعی بسیار پیچیده از ارتباطات موزون می‌باشد

سمیه علیزاده هیات علمی دانشکده صنایع
دانشگاه خواجه نصیر طوسی

مدل ریاضی شبکه عصبی

جمع ورودیهای وزندار $I_j =$

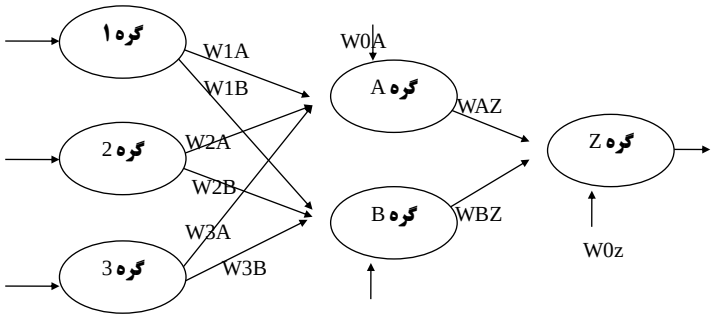
تابع فعالسازی $y_j = \phi(I_j)$



عصب ز ام

سمیه علیزاده هیات علمی دانشکده صنایع
دانشگاه خواجه نصیر طوسی

مثال



$W_{OZ}=۰/۵$	$W_{OB}=۰/۷$	$W_{OA}=۰/۵$	$x_۰=۱/۰$
$W_{AZ}=۰/۹$	$W_{۱B}=۰/۹$	$W_{۱A}=۰/۶$	$x_۱=۰/٤$
$W_{BZ}=۰/۹$	$W_{۲B}=۰/۸$	$W_{۲A}=۰/۸$	$x_۲=۰/۲$
$W_{۳B}=۰/٤$	$W_{۳A}=۰/۶$	$x_۳=۰/۷$	

سمیه علیزاده هیات علمی دانشکده صنایع
دانشگاه خواجه نصیر طوسی

فصل هفتم

مبنای شبکه‌های عصبی

۷–۱ مقدمه

در سال ۱۹۵۶، بنیاد راکفلر[1] برگزاری کنفرانسی را به عهده داشت که چشم اندازش به قرار زیر بود:

امکان استفاده از کامپیوترها و شبیه‌سازی در هر زمینه از یادگیری و سایر حوزه‌های هوش.

در همین کنفرانس بود که اصطلاح "هوش مصنوعی" مورد استفادهٔ عمومی قرار گرفت. بطور کلی هوش مصنوعی را به صورت زیر میتوان تعریف نمود:

فرایندهای کامپیوتری که سعی دارند فرایندهای تفکر انسان را تقلید نمایند، این فرایندها با فعالیتهایی که نیاز به استفاده از هوش دارند در ارتباطند.

عموماً این تعریف شامل زمینه‌های یادگیری خودکار، فهم زبان طبیعی، بینایی و تشخیص صدا، بازی کردن، حل مسائل ریاضی، روبوتیک و سیستمهای خبره است. در سالهای اخیر برخی محققان، شبکه‌های عصبی و فناوریهای مرتبط با آن را به عنوان بخشی از هوش مصنوعی در نظر می‌گیرند، در حالی که بعضی دیگر با توجه به سررشتهٔ خود در علوم پزشکی، شبکه‌های عصبی را زیر مجموعه هوش مصنوعی نمی‌دانند. در این بخش، ما شبکه‌های عصبی را به عنوان یکی از زیرمجموعه‌های موجه هوش مصنوعی می‌شناسیم. علاوه بر این ما، الگوریتم‌های ژنتیک، منطق

[1] . Rockefeller Foundation

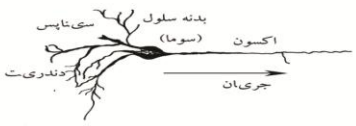
فازی یا سیستمهای فازی، امواج زیر[1]، اتومات سلولی[2] و سیستمهای آشوب[3] را جزئی از زمینهٔ کلی هوش مصنوعی می‌دانیم.

۲-۷ مبنای پزشکی شبکه‌های عصبی

مغز انسان سیستم بسیار پیچیده‌ای است که قادر به تفکر، یادآوری و حل مسائل است. تاکنون تلاشهای زیادی صورت گرفته است تا مدلهای کامپیوتری از عملکرد مغز تقلید نمایند و اگر چه نتایج این تلاشها بسیار تماشایی است ولی در مقایسه با عملکرد پیچیده مغز انسان کاملا کمرنگ هستند.

عصب یک واحد سلولی اساسی از سیستم عصبی مغز است. عصب یک عنصر پردازشگر ساده است که از طریق مسیرهای ورودی بنام **دندریتها**[4] علائمی را از سایر اعصاب دریافت کرده و آنها را با هم ترکیب می‌کند. اگر این علامت ورودی مرکب به اندازه کافی قوی باشد، عصب به اصطلاح "شلیک" می‌کند و یک علامت خروجی را در طول **اکسونی**[5] که به دندریت سایر اعصاب متصل است ارسال می‌کند. شکل ۱-۷ طرحی از قسمتهای مختلف یک عصب است. هر علامتی که در طول دندریت یک عصب فرستاده می‌شود از یک **سیناپس**[6] یا **اتصال سیناپسی** عبور می‌کند. این اتصال، یک شکاف بسیار کوچک در دندریت است که با نوعی مایع هادی عصبی[7] پر شده است که جریان بارهای الکتریکی را تندتر یا کندتر می‌کند. اصولا عملکرد عصبها بر پایه فعل و انفعالات شیمیایی است و این مایع هادی عصبی علائمی الکتریکی تولید می‌کند و به سوی هسته یا سومای[8] عصب می‌فرستد. تنظیم میزان مقاومت یا هدایت الکتریکی این فاصلهٔ سیناپسی، یک فرایند بسیار مهم است. به علاوه همین تنظیم هدایت باعث حفظ کردن مطالب و یادگیری می‌شود. وقتی قدرت سیناپسی اعصاب تنظیم شود مغز "یادگیری می‌نماید" و اطلاعات را ذخیره می‌کند.

---
[1] . Wavelets
[2] . Cellular Automata
[3] . Chaotic Systems
[4] . Dendrites
[5] . Axon
[6] . Synapse
[7] . Neurotransmitter Fluid
[8] . Soma



شکل ۱-۷ طرح قسمتهای مختلف یک عصب بیولوژیکی

وقتی انسان متولد می‌شود پوستهٔ دماغی مغز او تقریبا صد بیلیون عصب دارد. خروجی هر یک از این اعصاب از طریق اکسونهایشان (مسیرهای خروجی) به حدود هزار عصب دیگر فرستاده می‌شود. هر یک از این هزار مسیر دارای یک اتصال سیناپسی است که جریان بارهای الکتریکی را توسط یک فرایند عصبی- شیمیایی کنترل می‌کند. بنابراین حدود صد تریلیون اتصال سیناپسی وجود دارند که بر رفتار مغز تأثیر می‌گذارند. واضح است که اگر ما بخواهیم از فرایندهای مغز انسان تقلید نماییم نمی‌توانیم کار بیلیونها عصب و تریلیونها اتصال سیناپسی را در نظر بگیریم. علاوه بر آن، بزرگترین شبکه‌های عصبی ما شامل حدود چند هزار عصب مصنوعی و کمتر از یک میلیون اتصال سیناپسی مصنوعی است.

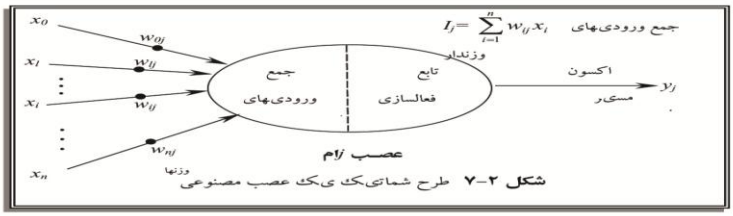
یکی از زمینه‌هایی که شبکه‌های عصبی مصنوعی در آن بسیار مفید واقع می‌شوند سرعت است. وقتی شخصی به سمت یک اتاق قدم برمی‌دارد حدود نیم ثانیه طول می‌کشد تا شخص دیگری را در آن اتاق است تشخیص دهد. ما می‌گوییم این فرایند تشخیص شامل حدود ۲۰۰ تا ۲۵۰ عمل جداگانه در مغز است. از نقطه نظر سرعت، این به آن معناست که مغز انسان با سرعتی حدود ۴۰۰ تا ۵۰۰ هرتز کار می‌کند. کامپیوترهای دیجیتالی جدید با سرعتی حدود ۱۰۰ تا ۲۰۰ مگاهرتز عمل می‌کنند که سرعت بسیار بالاتری نسبت به مغز است. ولی از سوی دیگر این مزیت به نحو محسوسی کاهش می‌یابد زیرا کامپیوترهای دیجیتالی بروش سری کار می‌کنند در حالی که مغز بروش موازی عمل می‌کند. البته در سالهای اخیر تراشه شبکه عصبی ساخته شده که کامپیوترهای عصبی را قادر می‌سازد بروش موازی عمل کنند.

irmgn.ir

فهرست اصطلاحات در زمینهٔ شبکه‌های عصبی هنوز استاندارد نشده است. ممکن است شما کتابها یا مقالات فنی پیدا کنید که از شبکه‌های عصبی مصنوعی به عنوان *سیستمهای ارتباطی* و از اعصاب مصنوعی به عنوان **عناصر پردازشگر**[1] (PEs)، *گره‌ها* و به عبارت ساده‌تر، *اعصاب* یاد شود. در این بخش ما از اصطلاحات *اعصاب* و شبکه‌*های عصبی* استفاده می‌کنیم و از صفت مصنوعی صرف نظر می‌نماییم زیرا فقط با اعصاب مصنوعی در این بخش سروکار داریم.

۷-۳ اعصاب مصنوعی

یک عصب مصنوعی مدلی است که اجزاء آن شباهت مستقیمی به اجزاء واقعی دارند. شکل ۷-۲ نمایی از یک عصب مصنوعی است. علائم ورودی با $x_0, x_1, x_2, \ldots, x_n$ مشخص شده‌اند. این علائم، متغیرهایی پیوسته هستند و نه پالسهای الکتریکی که در مغز رخ می‌دهند. هر یک از این مقادیر ورودی تحت تأثیر وزنی (که گاه وزن سیناپسی نامیده می‌شود) قرار می‌گیرند که تابع این وزن شبیه اتصال سیناپسی در یک عصب واقعی است. بسته به میزان هدایت یا مقاومت جریان علائم الکتریکی، این وزنها می‌توانند مثبت یا منفی باشند. این عناصر پردازشگر از دو قسمت تشکیل شده‌اند. قسمت اول ورودیهای وزن‌دار را با هم جمع می‌زند و کمیتی بنام I بدست می‌آورد؛ قسمت دوم یک صافی غیر خطی است که معمولاً **تابع فعالسازی**[2] نامیده می‌شود و از طریق آن جریانهای علائم ورودی ترکیب می‌شوند.



شکل ۷-۲ طرح شماتیک یک عصب مصنوعی

[1] . Processing Elements
[2] . Activation Function

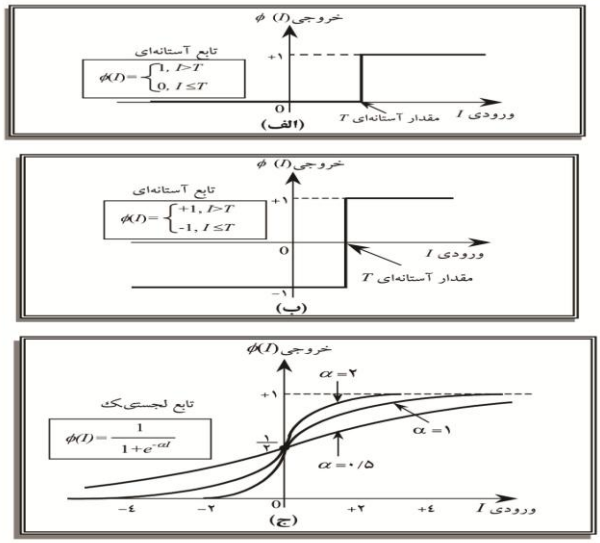
شکل ۷-۳ چند تابع فعالسازی ممکن را نشان می‌دهد. تابع فعالسازی می‌تواند مانند شکل ۷-۳ (الف) یک تابع آستانه‌ای[1] باشد که فقط زمانی اطلاعات را عبور می‌دهد که خروجی I که مربوطه به قسمت اول عصب مصنوعی است از مقدار آستانه‌ای T تجاوز کند (خروجی پایانی عصب معمولاً a+1 علامت است). همچنین این تابع می‌تواند مانند شکل ۷-۳ (ب) یک تابع علامت[2] (که گاه تابع اندازه‌گیری نامیده می‌شود) باشد که وقتی خروجی کمتر از مقدار آستانه‌ای T باشد اطلاعات منفی فرستاده و وقتی خروجی بیشتر از مقدار آستانه‌ای T باشد اطلاعات مثبت می‌فرستد. در اغلب موارد، تابع فعالسازی یک تابع پیوسته است که بطور تدریجی بین دو مقدار مجانب مثلاً صفر و ۱ یا ۱- و ۱+ تغییر می‌کند و **تابع S مانند**[3] یا تابع سیگموئید نام دارد. پُر استفاده‌ترین تابع فعالسازی، تابع لجستیک[4] است که یکی از انواع توابع فعالسازی S مانند است که در شکل ۷-۳ (ج) نشان داده شده است و از رابطه زیر محاسبه می‌شود

$$\phi(I) = \frac{1}{1+e^{-\alpha I}} \qquad (۷-۳-۱)$$

که α ضریبی است که در تغییرات تابع بین دو مقدار مجانب، شیب تابع را مشخص می‌نماید.

"تابع فشار"[5] نام گویاتری برای تابع فعالسازی است که نشان می‌دهد این تابع، مقادیر خروجی یک عصب مصنوعی را بین مقادیر مجانب، فشرده یا محدود می‌کند. این محدودسازی باعث می‌شود تا خروجی عناصر پردازشگر در یک محدودهٔ پویا و قابل قبول قرار گیرد. البته در برخی موقعیتهای مشخص و قطعی می‌توان از یک رابطه خطی، گاهی فقط در نیم فضای سمت راست به عنوان تابع فعالسازی استفاده کرد. ولی باید به این نکته توجه کرد که استفاده از یک تابع خطی، خصوصیت غیرخطی بودن عصب مصنوعی را از بین می‌برد. یک شبکه عصبی بدون ویژگی غیرخطی بودن نمی‌تواند پدیده‌های غیر خطی را مدلسازی کند.

[1] . Threshold Function
[2] . Signum Function
[3] . Sigmoidal Function
[4] . Logistic Function
[5] . Squashing Function



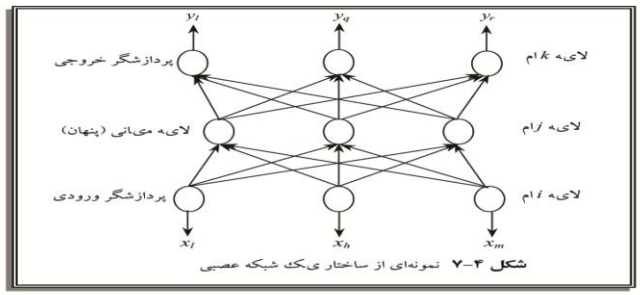
شکل ۳-۷ توابع انتقال برای اعصاب: (الف) تابع فعالسازی آستانه‌ای (وقتی $T=۰$)، تبدیل به یک تابع فعالسازی صفر و یک خواهد شد). (ب) تابع فعالسازی علامت (که گاه تابع اندازه‌گیری نامیده می‌شود). (ج) توابع فعالسازی لجستیک برای مقادیر $\alpha=۰/۵$ و $\alpha=۱$ و $\alpha=۲$.

۷-۴ شبکه‌های عصبی مصنوعی

یک شبکه عصبی مصنوعی را می‌توان به صورت زیر تعریف کرد:

یک سیستم پردازش داده‌ها که از تعداد زیادی عناصر پردازشگر ساده و بسیار مرتبط با هم (یعنی همان اعصاب مصنوعی) تشکیل شده است و در ساختار آن از پوستهٔ دماغی مغز الهام گرفته شده است.

این عناصر پردازشگر معمولا در لایه‌ها یا صفحات منظمی قرار گرفته‌اند بطوری که بین لایه‌ها ارتباطات کامل و یا تصادفی وجود دارد. شکل ۴-۷ این نظم و ترتیب را نشان می‌دهد بطوری که لایه ورودی به منزلهٔ پردازشگری است که پس از پردازش داده‌های ورودی، آنها را به شبکه ارائه می‌دهد. این لایهٔ ورودی یک لایه عصبی محاسباتی نیست زیرا گره‌های آن نه وزن ورودی دارند و نه تابع فعالسازی (برخی مؤلفان در تشریح شبکه‌های عصبی، این لایه را به حساب نمی‌آورند. ما این لایه را به حساب می‌آوریم ولی از علائم متفاوتی برای گره‌های این لایه استفاده می‌کنیم زیرا باید بین انواع مختلف اعصاب، تفاوت قائل شویم). لایهٔ فوقانی، لایهٔ خروجی است که خروجی



شکل ۴-۷ نمونه‌ای از ساختار یک شبکه عصبی

irmgn.ir

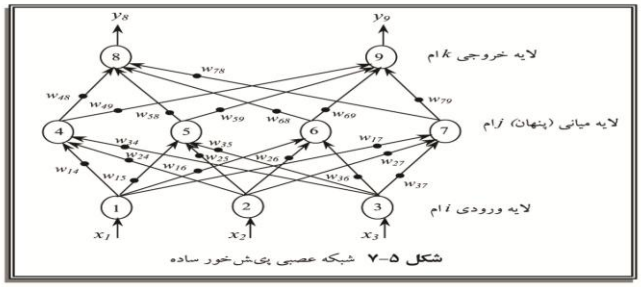
شبکه را در پاسخ به یک ورودی مشخص نشان می‌دهد. سایر لایه‌ها، لایه‌های میانی یا پنهان نامیده می‌شوند زیرا بین آنها و دنیای خارجی ارتباطی وجود ندارد. نوعاً لایه‌های ورودی، لایه‌های پنهان و لایه‌های خروجی به ترتیب لایه‌های iام، jام، kام نامیده می‌شوند.

در حال حاضر شبکه‌های عصبی مورد استفاده بر دو نوعند: شبکه‌های عصبی *غیرشرکت‌پذیر*[1] که در آنها بردار خروجی با بردار ورودی متفاوت است، و شبکه‌های عصبی *خود شرکت‌پذیر*[2] که در آنها بردار ورودی و بردار خروجی یکسانند. بجز مواردی که تصریح شده‌اند همه شبکه‌های عصبی این کتاب، غیر شرکت‌پذیرند.

یک شبکه عصبی نوعی، شبکه ایست با "ارتباطات کامل[3]" و به این معناست که بین هر یک از عصبهای هر لایه با هریک از عصب‌های لایهٔ بعد، یک ارتباط وجود دارد که نمونه‌ای از آن در شکل ۵-۷ نشان داده شده است. شبکه‌ای که در آن هیچگونه ارتباط افقی بین اعصاب یک لایه نباشد و ارتباطی بسوی لایه‌های قبلی نیز وجود نداشته باشد، شبکه پیش‌خور[4] نامیده می‌شود. شبکه‌هایی که ارتباطات بازخور[5] دارند (یعنی شبکه‌هایی که ارتباطاتی از یک لایه بسوی لایه قبلی دارند) نیز شبکه‌های مفیدی هستند و در قسمتهای بعد مطرح می‌شوند. همچنین ارتباطات افقی بین اعصاب یک لایه نیز ارتباطات بازخور نامیده می‌شود. در موارد خاصی از خروجی یک عصب به ورودی همان عصب بازخور وجود دارد. در کلیه موارد، این ارتباطات دارای وزنی هستند که باید تحت آموزش قرار گیرد.

همانطور که شکل ۵-۷ نشان می‌دهد هر یک از ارتباطات بین اعصاب، دارای یک وزن قابل تنظیم است. این شبکه عصبی ساده، یک شبکه پیش‌خور با ارتباطات کامل است که سه عصب در لایه ورودی، چهار عصب در لایه میانی یا پنهان و دو عصب در لایهٔ خروجی دارد. هریک از وزنها بصورت نقطه توپری بر روی ارتباط رسم شده و با علامت w_{ij} نشان داده شده است. به عنوان مثال وزن w_{34} نشان دهندهٔ وزنی است که روی ارتباط بین عصب‌های ۳ و ۷ قرار دارد.

[1] . Heteroassociative Neural Network
[2] . Autoassociative Neural Network
[3] . Fully Connected
[4] . Feedforward Network
[5] . Feedback Connections



شکل ۵-۷ شبکه عصبی پی‌ش‌خور ساده

حال اجازه دهید که شبکه عصبی شکل ۵-۷ را در نظر بگیریم که در آن بردار ورودی X از عناصر x_1، x_2، x_3 تشکیل شده و بردار خروجی Y دارای عناصر y_8 و y_9 است. وقتی علامت x_1 به عصب شماره ۱ در لایه ورودی می‌رود خروجی x_1 به کلیه اعصاب مصنوعی در لایه میانی یا پنهان می‌رود و از وزنهای w_{14}، w_{15}، w_{16} و w_{17} عبور می‌کند. علائم ورودی x_2، x_3 نیز مسیر مشابهی را طی می‌کنند یعنی از طریق وزنهای مربوطه که در شکل ۵-۷ نشان داده شده‌اند، به عصب‌های شماره ۴، ۵، ۶ و ۷ فرستاده می‌شوند.

حال اجازه دهید رفتار عصب شماره ۴ را در نظر بگیریم. از سه عصب موجود در لایه ورودی، سه ورودی به این عصب فرستاده می‌شود که تحت تاثیر وزنهای ارتباطی w_{34}، w_{24}، w_{14} قرار می‌گیرند. قسمت اول این عصب، مجموع این سه ورودی وزن‌دار را محاسبه می‌کند. سپس حاصل جمع بدست آمده به قسمت دوم فرستاده می‌شود که یک تابع غیر خطی است و معمولا یک منحنی لجستیک بین صفر و یک است که در شکل ۳-۷ (ج) نشان داده شده است. سپس خروجی این تابع فعالسازی یا تابع فشار با عبور از وزنهای w_{48} و w_{49} به عصب‌های شماره ۸ و ۹ فرستاده می‌شود. عصب‌های شماره ۵، ۶ و ۷ نیز روش مشابهی را طی می‌کنند. عصب‌های ۸ و ۹

مجموع ورودیهای وزن‌دار را که از عصبهای ۴، ۵، ۶ و ۷ دریافت کرده‌اند بدست می‌آورند و این حاصل جمعها از توابع فعالسازی عبور میکنند تا y_8 و y_9 یعنی عناصر بردار خروجی Y ایجاد شود.

علائم برداری و ماتریسی

استفاده از علائم بردار و ماتریس برای ورودیها، خروجی‌ها و وزنها بسیار مناسب است. اجازه دهید شبکه عصبی شکل ۵-۷ را از بالای لایه پنهان برش دهیم، این برش در شکل ۶-۷ نشان داده شده است. خروجی‌های عصب‌های ۴، ۵، ۶ و ۷ در بردار V_j نشان داده شده‌اند و عناصر V_j عبارتند از v_4، v_5، v_6 و v_7. اگر ما توابع فعالسازی را به توابع خطی محدود کنیم روابط ریاضی را که در قسمت قبلی تشریح شد می‌توان به شکل ماتریس نوشت و بنابراین بردار ستونی V_j معادل است با ضرب نقطه‌ای ماتریس وزنها یعنی W_{ij} در بردار ورودی X_i. این رابطه را بصورت زیر می‌توان نشان داد.

$$\begin{bmatrix} v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} = \begin{bmatrix} w_{14} & w_{24} & w_{34} \\ w_{15} & w_{25} & w_{35} \\ w_{16} & w_{26} & w_{36} \\ w_{17} & w_{27} & w_{37} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad (۷-۴-۱)$$

$$V_j = W_{ij} \,.\, X_i \qquad (۷-۴-۲)$$

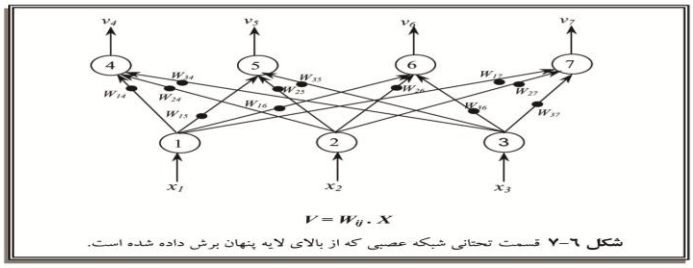
به روشی مشابه با استفاده از نیمهٔ فوقانی شبکه عصبی که در شکل ۷-۷ رسم شده می‌توان نشان داد که بردار خروجی Y_k با ضرب نقطه‌ای ماتریس اوزان W_{jk} و بردار ورودی V_j معادل است. این رابطه را به صورت زیر می‌توان نشان داد

$$\begin{bmatrix} y_8 \\ y_9 \end{bmatrix} = \begin{bmatrix} w_{48} & w_{48} & w_{68} & w_{78} \\ w_{49} & w_{59} & w_{69} & w_{79} \end{bmatrix} \cdot \begin{bmatrix} v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} \qquad (۷-۴-۳)$$

و یا

$$Y_k = W_{jk} \,.\, V_j \qquad (۷-۴-۴)$$

با ترکیب معادلات (۱-۴-۷) و (۳-۴-۷) مشخص می‌شود که بردار خروجی Y_k با ضرب نقطه‌ای دو ماتریس و بردار ورودی X_i بدست می‌آید.



$$V = W_{ij} \,.\, X$$

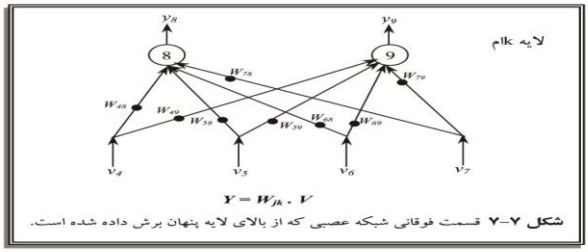
شکل ٦-٧ قسمت تحتانی شبکه عصبی که از بالای لایه پنهان برش داده شده است.

$$\begin{bmatrix} y_8 \\ y_9 \end{bmatrix} = \begin{bmatrix} w_{48} & w_{48} & w_{68} & w_{78} \\ w_{49} & w_{59} & w_{69} & w_{79} \end{bmatrix} \cdot \begin{bmatrix} w_{14} & w_{24} & w_{34} \\ w_{15} & w_{25} & w_{35} \\ w_{16} & w_{26} & w_{36} \\ w_{17} & w_{27} & w_{37} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad (۷-۴-۵)$$

چون از ضرب دو ماتریس می‌توان یک ماتریس واحد بنام W_{ijk} بدست آورد پس نتیجه می‌گیریم که بردار خروجی Y با ضرب نقطه‌ای این ماتریس ترکیبی و بردار ورودی معادل است:

$$Y_k = W_{ij} \,.\, W_{jk} \,.\, X_i = W_{ijk} \,.\, X_i \qquad (۷-۴-۶)$$

محدودیت خطی بودن توابع فعالسازی به این معناست که روابط موجود درمعادلات (۱-۴-۷) تا (۶-۴-۷) به شدت محدود خواهند شد. به علاوه این نشان می‌دهد که یک پرسپترون سه لایه که توابع فعالسازی لایه میانی و خروجی آن خطی باشند را می‌توان با یک شبکه دو لایه جایگزین نمود که تابع فعالسازی لایه خروجی آن خطی است. معهذا این فرآیند مفهوم ماتریسهای اوزان را معرفی می‌کند که کاربردهای فراوانی دارد.



شکل ۷-۷ قسمت فوقانی شبکه عصبی که از بالای لایه پنهان برش داده شده است.

شبکه‌های عصبی و بازخور

شبکه عصبی پیش‌خوری که در شکل ۵-۷ نشان داده شده با یک روش ساده مستقیم کار می‌کند. وقتی بردار X_i روی لایه ورودی اعمال می‌شود، محاسبات مربوط به ورودیهای وزندار و جمع ورودیها و اعمال تابع فعالسازی بسرعت برای هر عصب انجام شده و فرآیند از لایه میانی و سپس لایه خروجی حرکت می‌کند. ولی اگر ارتباطات بازخور وجود داشته باشد چه این ارتباط بین عصبهای هم‌تراز در یک لایه باشد و چه از یک لایه به لایه ما قبل، فرایند بسیار پیچیده‌تر خواهد شد. در یک شبکه عصبی که عملیات ریاضی تقریبا بطور فوری انجام می‌شوند، اطلاعات در سرتاسر شبکه منتشر شده از یک لایه به لایه بعد و نیز در داخل لایه‌ها جریان یافته تا اینکه نوعی حالت همگرایی و تقارب حاصل شود. وقتی عملیات ریاضی بطور مسلسل‌وار انجام شوند فرایند پیچیده‌تر خواهد شد. ابتدا برای روابط پیش‌خور (رو به جلو) خروجی‌ها محاسبه می‌شوند، سپس این محاسبات برای روابط بازخور انجام می‌شود و بعد با استفاده از نتایج محاسبات قبلی، محاسبات برای ارتباطات پیش‌خور دوباره انجام می‌شود و این فرایند ادامه می‌یابد تا مقادیر موازنه شده بدست آیند. در بیشتر موارد شبکه‌های عصبی با روابط بازخور می‌توانند بسیار مفید واقع شوند. ولی با این حال در ۸۰ درصد کاربردها از شبکه‌های عصبی پیش‌خور استفاده می‌شود.