دانشگاه تربیت معلم

دانشکدهٔ علوم ریاضی و کامپیوتر

پایان‌نامه کارشناسی ارشد آمار (گرایش آمار ریاضی)

عنوان

پیش‌بینی بر اساس شبکه‌های عصبی مصنوعی

تدوین

آزاده ابراهیمی

استادراهنما

دکتر عین‌اله پاشا

شهریور ۱۳۹۰

چکیده

یکی از مهمترین کاربردهای شبکه‌های عصبی مصنوعی، طراحی مدلی است که بتوان براساس آن مقدار یک یا چند متغیر وابسته را به کمک متغیرهای مستقل (کمکی) پیش‌بینی کرد. ازاین رو هدف ما در این پایان‌نامه طراحی و ارائه‌ی یک مدل شبکه عصبی (پرسپترون چندلایه) براساس داده‌های واقعی و مقایسه‌ی آن با یک مدل رگرسیون خطی چندگانه است، به همین دلیل در فصول یک و دو به بررسی کلیاتی از شبکه عصبی مصنوعی و رگرسیون خطی چندگانه می‌پردازیم. پس از طرح مسئله بهترین مدل شبکه عصبی و رگرسیون خطی را به داده‌ها برازش می‌دهیم. در نهایت بامقایسه‌ی نتایج این دو مدل ـمقادیر R^2 و SSEـبه این نتیجه می‌رسیم که شبکه عصبی طراحی شده در مقایسه با رگرسیون خطی چندگانه، مدل بهتری را به داده‌ها برازش می‌دهد.

واژه‌های کلیدی: شبکه‌های عصبی مصنوعی، شبکه عصبی پرسپترون چندلایه، رگرسیون خطی ساده و چندگانه.

رده‌بندی موضوعی ریاضی 2010: 92B20

فهرست مطالب

مقدمه . . . . . و

فصل اول کلیات در مورد شبکه‌های عصبی مصنوعی ۱

۱.۱ مقدمه ۱

۲.۱ شبکه‌های عصبی بیولوژیک ۱

۳.۱ شبکه‌های عصبی مصنوعی و تاریخچه‌ی آن ۴

۴.۱ ویژگی‌های شبکه‌های عصبی مصنوعی ۶

۱.۴.۱ قابلیت یادگیری ۶

۲.۴.۱ قابلیت تعمیم ۶

۳.۴.۱ پردازش موازی ۷

۵.۱ ساختار شبکه‌های عصبی مصنوعی ۷

۱.۵.۱ عناصر پردازش (نرون یا گره) ۱۰
۲.۵.۱ لایه‌ی ورودی ۱۱
۳.۵.۱ لایه‌های پنهان ۱۱
۴.۵.۱ وزن‌ها (اتصالات یا ارتباطات نرون‌ها) ۱۳
۵.۵.۱ انواع اتصالات یا ارتباط در شبکه‌های عصبی ۱۴
۶.۵.۱ تابع مجموع ۱۵
۷.۵.۱ توابع تبدیل. ۱۶
۸.۵.۱ لایه‌ی خروجی. ۱۹
۶.۱ انواع شبکه‌های عصبی بر مبنای نوع یادگیری ۱۹
۱.۶.۱ روش یادگیری نظارت شده یا با ناظر. ۱۹
۲.۶.۱ روش یادگیری نظارت نشده یا بدون ناظر. ۲۰
۷.۱ آموزش پس انتشار خطا ۲۱
۸.۱ توقف آموزش و مشکل انطباق بیش از حد ۲۶
۹.۱ شبکه‌ی عصبی پرسپترون چند لایه ۲۸
۱۰.۱ یک نتیجه‌ی مهم ۳۱

فصل دوم مروری بر مباحث رگرسیون خطی چندگانه **۳۳**

۱.۲ مقدمه ۳۳
۲.۲ رگرسیون خطی ساده ۳۴
۳.۲ رگرسیون خطی چندگانه ۳۷

فصل سوم کاربرد شبکه‌های عصبی در داده‌های واقعی **۴۰**

۱.۳ مقدمه ۴۰
۲.۳ طرح مسئله ۴۰

۳.۳ آماده سازی (پیش‌پردازش) داده‌ها ۴۲
۴.۳ تعیین ساختار شبکه ۴۳
۵.۳ تحلیل نتایج (خروجی) ۴۴
۱.۵.۳ خلاصه فرآیند. ۴۵
۲.۵.۳ اطلاعات شبکه. ۴۵
۳.۵.۳ دیاگرام. ۴۶
۴.۵.۳ خلاصه‌ی مدل. ۴۷
۵.۵.۳ برآورد پارامترها (وزن‌ها) ۴۸
۶.۵.۳ نمودار مقادیر پیش‌بینی شده درمقابل مقادیر مشاهده شده. ۴۸
۷.۵.۳ نمودار مقادیر باقی‌مانده در مقابل مقادیر پیش‌بینی شده. ۴۹
۸.۵.۳ تحلیل اهمیت متغیر مستقل. ۵۰

فصل چهارم مقایسه‌ی رگرسیون و شبکه‌های عصبی **۵۲**

۱.۴ مقدمه ۵۲
۲.۴ مدل رگرسیون برای پیش بینی UCS ۵۲
۳.۴ مدل رگرسیون برای پیش‌بینی E ۵۷
۴.۴ نتیجه‌گیری ۶۰

مراجع **۶۱**

واژه‌نامه فارسی به انگلیسی **۶۲**

نمایه **۶۵**

مقدمه

نحوه‌ی یادگیری و ایجاد ارتباط میان مفاهیم و اشیاء، یکی از رموز پیچیده‌ی دستگاه آفرینش است. آن چه جانوران عالی و در رأس آن‌ها، انسان را به داشتن چنین قدرت شگرفی توانا می‌سازد، ساختار فوق پیچیده‌ی دستگاه عصبی آن‌ها است. این دستگاه از تجمع و آرایش ویژه‌ی سلول‌های تخصص یافته به نام نرون، تشکیل شده است. با تأثیرپذیری و تأثیرگذاری متقابل بر دستگاه‌های دیگر بدن، ارتباط هدف‌مند جاندار را با محیط اطرافش ممکن می‌سازد. دریافت پیام‌های بیرونی، تحلیل و پاسخ‌دهی مناسب به آن‌ها منوط به چینش و عملکرد صحیح نرون‌ها است. [۴]

پیشرفت حاصل شده در زمینه‌های بیولوژیکی در دهه‌های ابتدایی قرن بیستم، دانشمندان را به فکر تقلید از دستگاه عصبی برای حل مسائل ریاضی و آماری انداخت. در اواخر قرن بیستم، سیستم‌های هوشمند ظهور نموده و در سال‌های اخیر رشد چشمگیری داشته است. در سیستم‌های هوشمند، از مدل‌هایی مشابه الگوی رفتاری، چگونگی تصمیم‌گیری، تجربه‌ی انسان، ساختار مغز، شبکه‌های اعصاب و یا ساختار کروموزوم‌های سلول‌ها در حل مسائل استفاده می‌شود.

در واقع منظور از سیستم هوشمند، سیستمی نیست که تمام عملکردهای مغز انسان را انجام دهد، بلکه تنها با تعریف شرایط و قوانین، سیستم در یک محدوده‌ی تعریف شده، هوشمند می‌گردد. سیستم‌های هوشمند برای حالت‌هایی که ابهامات زیادی در مسئله وجود دارد کارایی بیشتری دارند. در سیستم‌های هوشمند اساس کار پویایی سیستم است. این سیستم‌ها می‌توانند آموزش‌پذیر هم باشند.

یک سیستم هوشمند مانند یک جعبه‌ی سیاهی است، که یک سری ورودی‌ها و خروجی‌ها به این جعبه متصل هستند. برحسب این که از چه نوع سیستم هوشمندی استفاده می‌گردد، عملیات ریاضی خاصی در داخل این جعبه صورت می‌گیرد. سیستم‌های هوشمند را در برابر هوش آدمی، هوش مصنوعی نیز می‌نامند.

سیستم‌های هوشمند، یکی از ابزارهای بسیار قوی برای حل مسایل پیچیده می‌باشند. از نقاط قوت یک سیستم هوشمند عصبی آن است که به معادلات ریاضی حاکم بر روابط پدیده‌ها کاری ندارد بلکه تنها به تاریخچه‌ی آن پدیده کار دارد. در واقع یک سیستم هوشمند عصبی با به خاطرسپاری تاریخچه‌ی یک پدیده، می‌تواند آن را پیش‌بینی کند و در صورتی که از اطلاعات جامع و کامل استفاده شده باشد، می‌تواند برای پیش‌بینی رفتار این پدیده در منطقه‌ی دیگر نیز استفاده نمود. [۶]

در این پایان‌نامه به بررسی یکی از مهمترین کاربردهای شبکه‌های عصبی مصنوعی- که شاخه‌ای از هوش مصنوعی است- یعنی پیش‌بینی با استفاده از شبکه عصبی پرسپترون چند لایه MLP[1] پرداخته می‌شود. سپس این مدل را با مدل رگرسیون خطی چندگانه مقایسه می‌کنیم.

1) Multi Layer Perceptron

فصل اول

کلیات در مورد شبکه‌های عصبی مصنوعی[1]

۱.۱ مقدمه

در این فصل جهت آشنایی با شبکه‌های عصبی و کاربرد ویژه‌ی آن‌ها در پیش‌بینی، به تشریح اصول زیر بنایی این شبکه‌ها می‌پردازیم. ابتدا به شرح شبکه‌های عصبی بیولوژیکی موجود در مغز انسان، که منشاء تقلید و شبیه‌سازی در مسائل مهندسی شده‌اند، می‌پردازیم. سپس تاریخچه‌ی کوتاهی از دوره‌های گوناگونی که شبکه‌های عصبی مصنوعی (شبیه‌سازی شده به تقلید از شبکه‌های عصبی بیولوژیک) در طی سالیان گذشته باآن روبه‌رو بوده‌اند، ارائه می‌گردد، در ادامه شبکه‌های عصبی مصنوعی و اصول اولیه‌ی آن تشریح می‌گردد. پس از آن، ساختار شبکه‌های عصبی؛ شامل لایه‌های مختلف، نرون‌ها و توابع تبدیل آن‌ها تشریح و انواع شبکه‌های عصبی مورد بررسی قرار می‌گیرد. در ادامه یکی از مهمترین خطراتی که فرآیند یادگیری شبکه عصبی را تهدید می‌کند یعنی مشکل «انطباق بیش از حد» تشریح شده است و پس از آن به معرفی یکی از مهم‌ترین شبکه‌های عصبی در پیش‌بینی-پرسپترون چند لایه- می‌پردازیم.

۲.۱ شبکه‌های عصبی بیولوژیک

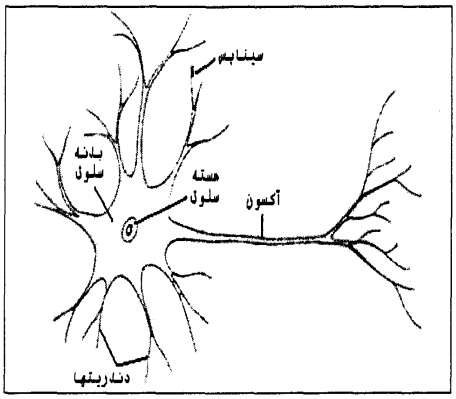
مغز به عنوان یک سیستم پردازش اطلاعات با ساختار موازی و کاملاً پیچیده که دو درصد (٪۲) از وزن بدن را تشکیل می‌دهد، بیش از بیست درصد (٪۲۰) کل اکسیژن بدن را مصرف می‌کند، که برای خواندن، نفس کشیدن، حرکت، تفکر،

1) Artificial Neural Network

تفحص و کلیه اعمال آگاهانه و بسیاری از رفتارهای ناخودآگاه استفاده می‌شود.

نگرش نوین در مورد کار مغز، ماحصل تفکراتی است که در اوایل قرن بیستم رامون سگال[1] در مورد ساختار مغز، به عنوان اجتماعی از اجزای محاسباتی کوچک به نام «نرون»[2] ارائه کرده است. برخی از ساختارهای نرون در هنگام تولد ساخته می‌شود و قسمت‌های دیگر در طول سیر حیات به خصوص در اوایل زندگی به خود می‌آیند و قوام می‌گیرند. دانشمندان به این نتیجه رسیده‌اند که عملکرد بیولوژیکی از قبیل ذخیره سازی و حفظ اطلاعات درخود نرون‌ها و ارتباطات بین نرون‌ها نهفته است. به عبارتی فنی‌تر، یادگیری به عنوان ایجاد ارتباطات جدید بین نرون‌ها و تنظیم مجدد ارتباطات موجود استنباط می‌شود. نرون‌ها نیز ساختاری موازی دارند. یعنی این که همهٔ نرون‌ها به طور هم زمان کار می‌کنند و پاسخ می‌دهند لذا عملکرد سریع‌تر مغز به دلیل ساختار کاملاً موازی نرون‌ها است.

تعداد تخمینی نرون‌ها در مغز انسان بین ۵۰ تا ۱۵۰ بیلیون است. نرون‌ها ساده‌ترین واحد ساختاری سیستم‌های عصبی هستند. (شکل ۱ـ۱) بافت‌هایی که عصب نامیده می‌شوند. اجتماعی از نرون‌ها هستند. این نرون‌ها، اطلاعات و پیام‌ها را از یک قسمت بدن به قسمت دیگر متصل می‌کنند. میلیون‌ها نرون در بدن انسان وجود دارد که حتی ساده‌ترین کارهای روزمره‌ی انسان از قبیل پلک زدن، تنها از طریق همکاری همه جانبه این نرون‌ها میسر است.



شکل(۱-۱): نمای کلی یک نرون حقیقی

دندریت‌ها[3] به عنوان مناطق دریافت سیگنال‌های الکتریکی، شبکه‌های تشکیل یافته از فیبرهای سلولی هستند که دارای سطح نامنظم و شاخه‌های انشعابی بیشماری هستند؛ به همین جهت به آن‌ها شبکه‌های دریافتی درخت‌گونه نیز می‌گویند.

دندریت‌ها سیگنال‌های الکتریکی را به هسته سلول منتقل می‌کنند. بدنه سلول، انرژی لازم را برای فعالیت نرون

1) Segal 2) Neuron 3) Dendrit

فراهم نموده و بر روی سیگنال‌های دریافتی عمل می‌کند که با یک عمل ساده جمع و مقایسه با یک سطح آستانه مدل می‌گردد. آکسون[1] برخلاف دندریت‌ها از سطحی هموارتر و تعداد شاخه‌های کمتر برخوردار است. آکسون طول بیشتری دارد و سیگنال الکتروشیمیایی دریافتی از هسته سلول را به نرون‌های دیگر منتقل می‌کند. محل تلاقی یک آکسون از یک سلول به دندریت‌های سلول دیگر را سیناپس[2] گویند که ارتباط بین نرون‌ها را برقرار می‌سازد. سیگنال‌ها می‌توانند بدون تغییر انتقال داده شوند یا این که توسط سیناپس‌ها تغییر داده شوند. یک سیناپس قادر است قدرت ارتباط را افزایش یا کاهش دهد و باعث تحریک یا بازدارندگی نرون مرتبط شود.

بیشترین تعداد نرون‌ها در مغز و مابقی در نخاع و سیستم‌های عصبی جانبی متمرکز شده‌اند. گرچه همگی نرون‌ها کارکرد یکسانی دارند، ولی اندازه و شکل آن‌ها بستگی به محل قرارگیری آن‌ها در سیستم عصبی دارد. با وجود این همه تنوع، بیشتر نرون‌ها از سه قسمت اساسی تشکیل شده‌اند:

۱) بدنهٔ سلول؛ که شامل هسته و قسمت‌های حفاظتی دیگر است.

۲) دندریت

۳) اکسون

دو مورد آخر، عناصر ارتباطی نرون‌ها را تشکیل می‌دهند.

نرون‌ها، بر اساس ساختارهای هدایت پیام در بین خود، به سه دسته تقسیم می‌شوند:

۱) نرون‌های حسی؛ که اطلاعات را از ارگان‌های حسی به مغز و نخاع می‌فرستند.

۲) نرون‌های محرک؛ که سیگنال‌های فرمان را از مغز و نخاع به ماهیچه‌ها و غدد هدایت می‌کنند.

۳) نرون‌های ارتباطی؛ که دو نرون اول (حسی و محرک) را به هم متصل می‌کنند.

ارتباط بین نرون‌های ارتباطی موجبات انجام کارهای پیچیده‌ی روزمره از قبیل؛ تفکر، احساسات، ادراک و محفوظات را فراهم می‌آورد، در انسان این نرون‌ها خیلی بیشتر از تعداد نرون‌های حسی و محرک است. [۳]

لازم به ذکر است که پیام‌های عصبی تنها به صورت یک‌طرفه حرکت می‌کنند، از دندریت‌ها به بدنه سلول و سپس به آکسون. زمانی که پیام به انتهای آکسون می‌رسد، فیبرهای ماهیچه‌ای، سلول‌های غددی و دیگر نرون‌هایی را

1) Axon 2) Synapse

که به آکسون متصل هستند تحریک می‌کند و بدین ترتیب انتقال پیام صورت می‌گیرد.

توانایی یادگیری و واکنش نسبت به تغییرات محیط اطراف، به هوش نیازمند است. فکرکردن و رفتار هوشمندانه توسط مغز و سیستم اعصاب مرکزی کنترل می‌شود. برای مثال کسانی که دارای آسیب مغزی هستند مشکلاتی در یادگیری و واکنش نسبت به تغییرات محیط دارند، ولی به هر حال قسمت‌های سالم مغز می‌توانند به کمک یادگیری جدید این نقص را جبران کنند.

۳.۱ شبکه‌های عصبی مصنوعی و تاریخچه‌ی آن

ساختار کلی شبکه‌های عصبی مصنوعی از شبکه عصبی بیولوژیکی انسان، الهام گرفته شده است. تحقیقات پیرامون شبکه‌های عصبی، با شناخت و بررسی کار یادگیری مغز انسان همراه بوده است. شبکه‌های عصبی مصنوعی، سیستم‌هایی هستند که قادر به انجام عملیاتی همانند سیستم‌های طبیعی عصبی هستند. یا به عبارت بهتر، می‌توانند ویژگی‌هایی شبیه به عملکرد مغز آدمی را تقلید نمایند.

تاریخچه‌ی شبکه‌های عصبی را می‌توان به دوره‌های زیر تقسیم نمود:

۱) **دورهٔ تلاش‌های اولیه:** بدون شک اولین گام در شبیه‌سازی شبکه‌های عصبی، زمانی بود که شخصی به نام سگال در سال ۱۹۱۱ اعلام کرد که مغز از عناصر اصلی ساختاری به نام «نرون » تشکیل یافته است. در این دوره، شبیه‌سازی‌های ابتدایی براساس منطق معمولی انجام گرفت. دیدگاه جدید شبکه‌های عصبی در دهه‌ی چهل (۴۰) قرن بیستم شروع شد، زمانی که وارن‌ـک کلاچ و والتر پیتز نشان دادند که شبکه‌های عصبی در اصل می‌توانند هر تابع حسابی و منطقی را محاسبه نمایند. شبکه آن‌ها براساس نرون‌های ساده‌ای که به صورت دودویی و یک حد آستانه‌ی ثابت بود، بنا شده بود. نتایج مدل آن‌ها نیز توابع ساده منطقی نظیر «و» و «یا» بود. تلاش بعدی، شبیه‌سازی توسط کامپیوتر بود. دو گروه، یکی فرلی وکلارک در سال ۱۹۵۴ و دیگری دودا و هایبیت در سال ۱۹۵۶ در این زمینه فعال بودند. گروه اول که محققان شرکت IBM بودند، با عصب شناسان دانشگاه مک گیل ارتباط نزدیکی داشتند، بنابراین هرزمان که مدلشان کار نمی‌کرد به مشاهده با عصب شناسان می‌پرداختند که باعث ایجاد یک روند میان رشته‌ای شد که تاکنون نیز ادامه دارد.

۲) **دوره‌ی تکنولوژی نوظهور و امیدبخش آینده:** در این دوره علاوه بر عصب شناسان که در توسعه‌ی شبیه‌سازی شبکه‌های عصبی نقش مؤثری داشتند، مهندسان و روانشناسان نیز در این توسعه بسیار مؤثر بودند. فرانک رزنبلات در سال ۱۹۵۸ تحرک قابل ملاحظه‌ای در هنگام طراحی و توسعه‌ی پرسپترون‌ها به وجود آورد. پرسپترون‌ها شامل سه لایه هستند که لایه‌ی میانی را لایه‌ی همراه و یا لایه‌ی مخفی نیز می‌گویند. سیستم دیگر، سیستم شبکه عصبی تطبیقی خطی آدلاین، یک وسیله‌ی الکتریکی آنالوگ ساخته شده از اجزای ساده بود. روشی که برای آموزش در آن استفاده می‌شد با روش استفاده شده در پرسپترون متفاوت بود. این سیستم از روش حداقل میانگین مربعات خطا برای یادگیری و آموزش استفاده می‌کرد.

۳) **دوره‌ی ناکامی:** در سال ۱۹۶۹ مینسکی و پاپرت کتابی در مورد عمومی کردن محدودیت پرسپترون تک لایه به چند لایه نوشتند. آن‌ها درکتاب خود ذکر کردند که: « به طور مشهودی به نظر ما توسعه و بسط پرسپترون‌های چند لایه ناکام و عقیم است.» امّا نتیجه قابل قبول ازکتاب آن‌ها، عدم اطمینان از شبیه‌سازی شبکه‌های عصبی بود.

۴) **دوره‌ی ابداع و نوآوری:** اگرچه در این دوره، علاقه و اعتماد به شبکه‌های عصبی به حداقل خود رسیده بود، امّا بعضی از محققان هنوز بر روی توسعه‌ی این موضوع، به ویژه مسائلی نظیر شناسایی الگوها، کار می‌کردند. در طی این دوره، نمونه‌های متفاوتی ازکارهای جدید ساخته شد که این روند هنوز هم به طور صعودی ادامه دارد.

در سال ۱۹۷۲ تئوکوهنن و جیمز اندرسون ، به طور جداگانه و بدون اطلاع از هم، شبکه‌های عصبی جدیدی را معرفی نمودند که قادر بودند به عنوان عنصر «ذخیره سازی اطلاعات» عمل نمایند. استفان گروسبرگ ، در این دهه روی شبکه‌های عصبی خود سازمانده فعالیت می‌کرد. وربوس در سال ۱۹۷۴ روش آموزش پس انتشار را، اگرچه قبلاً نیز شناخته شده بود، توسعه داد. شبکه‌های پس‌انتشار، مشهورترین و پرکاربردترین شبکه‌های عصبی هستند که امروزه استفاده می‌شوند. در اصل یک شبکه پس‌انتشار، یک پرسپترون چند لایه با تابع آستانه‌ای متفاوت در نرون‌ها است.

۵) **دوره‌ی ظهور مجدد:** پیشرفتی که شبکه‌های عصبی در واخر دهه‌ی ۷۰ و اوایل دهه‌ی ۸۰ داشت، سبب ظهور مجدد این زمینه شد. فاکتورهای مختلفی در این حرکت مؤثر بودند. به عنوان مثال می‌توان از مقالات و کنفرانس‌هایی که در این زمینه برگزار شد و یا مطرح شدن این مبحث به عنوان یک درس پیش‌دانشگاهی در

دانشگاه‌های آمریکا و اروپا نام برد. [۳]

۴.۱ ویژگی‌های شبکه‌های عصبی مصنوعی

قبل از آن که به توضیح و تشریح شبکه‌های عصبی مصنوعی بپردازیم، برخی از ویژگی‌های آن را بیان می‌کنیم.

۱.۴.۱ قابلیت یادگیری[1]

بدون تردید یکی از جالب‌ترین قابلیت‌های شبکه‌های عصبی، قابلیت یادگیری آن‌ها است. یادگیری یعنی توانایی تنظیم پارامترهای شبکه (وزن‌های سیناپسی) به منظور دریافت اطلاعاتی که بتواند در آینده به یاد آورده شود، همچنین با این هدف که اگر شبکه برای یک وضعیت خاص آموزش دید و تغییر کوچکی در شرایط محیطی شبکه (وضعیت خاص) رخ داد، شبکه بتواند باآموزش مختصر، برای شرایط جدید نیز کارآمد باشد. رویه‌های مختلفی برای تغییر وزن‌ها وجود دارد که به آن‌ها قوانین یادگیری می‌گویند. (در ادامه به آن می‌پردازیم)

زمانی که یک شبکه عصبی رفتار خود را تغییر می‌دهد و بهبود می‌بخشد، مقادیر بردارهای ورودی، از درون شبکه به دفعات عبور داده می‌شود و در هر عبور، شبکه وزن‌های خاص خود را به گونه‌ای اصلاح می‌کند که با توجه به ورودی، بتواند خروجی مورد نظر را تولید نماید. هر عبور بردار ورودی و مقدار خروجی مورد نظر از درون شبکه، یک سیکل[2] نامیده می‌شود.

۲.۴.۱ قابلیت تعمیم[3]

داده‌های آموزش (یادگیری)، در اصل گلچینی مناسب از یک گستره‌ی وسیع از مجموعه داده‌ها «ورودی-خروجی» را نمایندگی می‌کند. شبکه در صورت آموزش از روی مجموعه داده‌های یادگیری، قدرت تعمیم به کل مجموعه را خواهد داشت. شبکه‌ای از توانایی تعمیم‌پذیری برخوردار است که رابطه‌ی ورودی-خروجی محاسبه شده توسط شبکه، برای الگوی ورودی-خروجی آزمایشی (الگویی که هیچ گاه در فرآیند یادگیری به شبکه ارائه نشده است) صحیح عمل کند.

1) Learning ability 2) Epoch 3) Generalization

تعمیم از جمله ویژگی‌های مطرح شبکه‌های عصبی است و فرآیند یادگیری را در صورتی می‌توان موفق نامید که پس از آن، شبکه قادر به تعمیم باشد. تعمیم تحت تأثیر سه عامل است:

۱) اندازه و غنی بودن مجموعه داده‌های آموزشی

۲) ساختار شبکه

۳) پیچیدگی ذاتی مسئله تحت بررسی

چون مسئله ما از قبل مشخص است، پس فقط عوامل ۱ و ۲ باقی می‌ماند. اگر داده‌های آموزش ثابت باشد، برای قدرت تعمیم بیشتر باید به دنبال ساختار مناسب بود. ولی اگر ساختار شبکه داده شده باشد، باید به دنبال این باشیم که تعداد داده‌های آموزش را به آن گونه بدست آوریم که شبکه از تعمیم‌پذیری قابل قبولی برخوردار باشد.

۳.۴.۱ پردازش موازی[1]

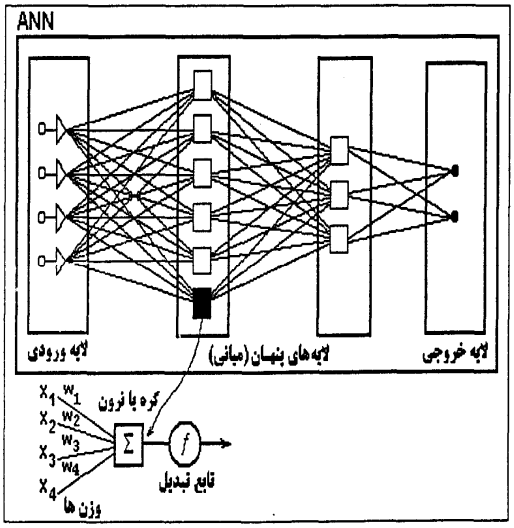
هنگامی که شبکه عصبی مصنوعی در قالب سخت‌افزار پیاده می‌شود، سلول‌هایی که در یک تراز و یک سطح قرار می‌گیرند، می‌توانند به طور همزمان به ورودی‌های آن ترازاز سطح پاسخ دهند. این ویژگی باعث افزایش سرعت پردازش می‌شود. در واقع در چنین سیستمی، وظیفه‌ی کلی پردازش، بین پردازنده‌های کوچکتر، مستقل از یکدیگر توزیع می‌گردد.

۵.۱ ساختار شبکه‌های عصبی مصنوعی

شبکه‌های عصبی ترکیبی از «نرون‌ها» یا «گروه‌ها» یا «واحدهای پردازشگر اطلاعات» هستند، که به صورت سلسله مراتبی در لایه‌ها سازماندهی شده‌اند. اولین لایه در شبکه، «لایه‌ی ورودی » است که اطلاعات را از محیط خارج شبکه، دریافت می‌کند. این لایه در مورد مغز، به پنج حس انسان و در مورد یک شبکه عصبی مصنوعی، به متغیرهای مستقل انتخاب شده توسط محقق، ارتباط دارد. «لایه‌های میانی یا پنهان» ورودی‌ها را به صورت تدر یجی از مبداء دریافت و به خروجی نهایی، تبدیل می‌کنند. آخرین لایه در شبکه « لایه خروجی» نامیده می‌شود. این لایه به پاسخ انسان در

1) Parallel Processing

مقابل تحریک ورودی در مغز و به پاسخ متغیر وابسته در مقابل تحریک متغیرهای مستقل در شبکه عصبی مصنوعی، مربوط می‌شود.

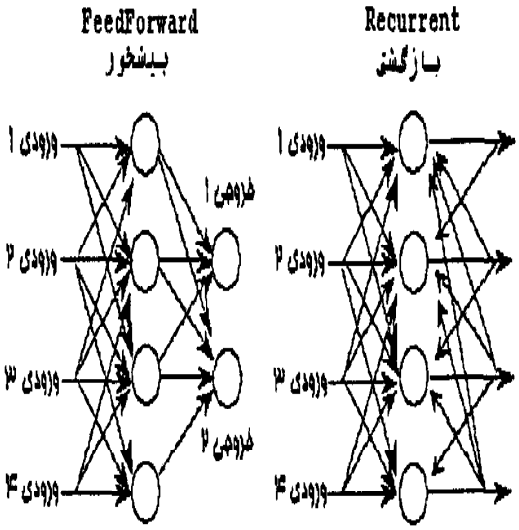


شکل(۱-۲): ساختار یک شبکه‌ی عصبی

برای تشریح «ساختار» یا «معماری » یا «الگوی » شبکه‌های عصبی به توضیح در رابطه با لایه ورودی، لایه خروجی، لایه پنهان، گره‌های ورودی، گره‌های خروجی، گره‌های پنهان، اتصالات یا وزن‌ها، توابع تبدیل و انواع ارتباط در شبکه می‌پردازند.

برخی از محققین نیز از اصطلاح «توپولوژی» شبکه استفاده کرده‌اند.

شکل ۱ـ۳ دو نمونه از معماری‌های شناخته شده شبکه عصبی را نشان می‌دهد. نحوه‌ی ترکیب اجزاء ذکر شده در بالا، یک معماری خاص را ایجاد می‌کند.



شکل(۱-۳): نمونه ای از معماری شبکه های عصبی مصنوعی

در شبکه پیشخور[1]، حرکت همواره از ورودی به سمت خروجی شبکه است. ولی در شبکه بازگشتی، حداقل یک مسیر پسخور در شبکه وجود دارد. به عبارت دیگر در شبکه پیشخور، حافظه وجود ندارد و خروجی نرون فقط وابسته به نرون‌های لایه‌های قبل است. ولی در شبکه پسخور[2]، خروجی یک نرون، تابع خروجی لحظهٔ قبل خود یا سایر نرون‌ها است.

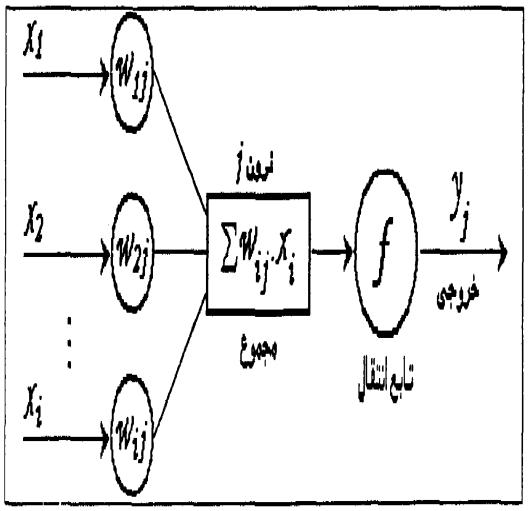
برای نشان دادن معماری یک شبکه از عبارت ساده‌ی $N^{n_1-n_2-n_3}$ استفاده می‌شود. به طور نمونه $N^{۳-۴-۲}$، یک شبکه‌ی عصبی با یک لایه پنهان را نشان می‌دهد که دارای سه نرون در لایه‌ی ورودی و چهار نرون در لایه‌ی پنهان و دو نرون در لایه‌ی خروجی است.

به عنوان یک اصل کلی در تمام مسائل مدل سازی شبکه‌ی عصبی، سعی براین است که ساده‌ترین مدل شبکه‌ی عصبی، که به قدر کفایت داده‌های یادگیری را نمایندگی نماید، مورد استفاده قرار گیرد. این اصل، «اصل امساک»[3] نامیده می‌شود. اصل امساک، هم در مود تعداد لایه‌های پنهان و هم در مورد نرون‌های درون این لایه‌ها مصداق دارد. با ارائه‌ی این توضیحات اولیه، در ادامه به توضیح هر یک از اجزای ساختار یک شبکه پرداخته می‌شود. [۳]

1) Feed-Forward 2) Feed-Back 3) Parsimony

۱.۵.۱ عناصر پردازش (نرون یا گره)

یک شبکه عصبی از نرون‌های مصنوعی تشکیل شده است که عناصر پردازش هستند. هر یک از نرون‌ها، ورودی‌ها را دریافت می‌کند، این ورودی‌ها را پردازش می‌کند و یک سیگنال خروجی را تحویل می‌دهد. می‌توان گفت یک نرون در شبکه به عنوان مرکز پردازش و توزیع اطلاعات عمل می‌کند و ورودی‌ها و خروجی‌های خود را دارد. این پردازش در شکل (۱ـ۴) نشان داده شده است.



شکل(۱-۴): پردازش اطلاعات در یک نرون مصنوعی

در مورد کار نرون‌ها، چند نکته مهم وجود دارد:

۱) نرون‌ها تنها به اطلاعات محلی نیاز دارند. تمام اطلاعات مورد نیاز یک نرون برای تولید یک مقدار خروجی، در ورودی نرون و خود نرون، موجود است. هیچ اطلاعاتی در مورد سایر مقادیر شبکه، مورد نیاز نرون نیست. در حقیقت هیچ نیازی به دانستن حالت سایر نرون‌ها، وقتی هیچ ارتباط روشنی بین آن‌ها نیست، وجود ندارد.

۲) نرون‌ها، تنها یک مقدار خروجی تولید می‌کنند. این مقدار خروجی از طریق ارتباطات، به عنوان ورودی به نرون‌های دیگر وارد می‌شود یا این که به عنوان خروجی شبکه در نظر گرفته می‌شود.

۳) هر نرون به صورت مستقل از سایر نرون‌ها، کار می‌کند. یعنی خروجی هر نرون، تنها وابسته به ورودی‌های آن نرون است.

۲.۵.۱ لایه‌ی ورودی

لایه ورودی شبکه عصبی، اولین لایه در شبکه است و اطلاعات ورودی به شبکه، از طریق این لایه، در اختیار شبکه قرار داده می‌شود. تعداد نرون‌ها در لایه ورودی که اصطلاحاً «تعداد گره‌های ورودی» نامیده می‌شود، به تعداد پارامترهایی که قرار است به شبکه وارد شود، بستگی دارد. هر ورودی به یک ویژگی منفرد، مرتبط است. برای مثال، در صورتی که مسئله، تصمیم‌گیری بر روی تصویب یا عدم تصویب یک وام است، بعضی ویژگی‌ها می‌توانند سطح درآمد متقاضی، سن و مالکیت منزل باشد. مقدار عددی یا بازنمایی یک ویژگی، ورودی به شبکه است. انواع مختلف داده‌ها نظیر متن، تصویر و صدا می‌تواند به عنوان ورودی استفاده شود. بنابراین ممکن است حتی نیاز به پیش پردازش هم وجود داشته باشد. محاسبات عصبی تنها می‌تواند اعداد را پردازش نماید. اگر یک مسئله با ویژگی‌های کیفی یا تصاویر سروکار داشته باشد، باید قبل ازاین که توسط شبکه عصبی به آن پرداخته شود، این ویژگی‌ها به معادل عددی آن‌ها تبدیل شود.

۳.۵.۱ لایه‌های پنهان

لایه‌های پنهان یا میانی لایه‌هایی هستند که در شبکه بین لایه‌ی ورودی و لایه‌ی خروجی قرار می‌گیرند. این لایه‌ها، شبکه را برای «تعمیم» که یکی از قابلیت‌های اساسی شبکه عصبی است، قدرتمند می‌سازند. به صورت تئوری، یک شبکه، تنها با یک لایه‌ی پنهان و تعداد کافی نرون در آن، قابلیت تقریب‌زنی هر تابع پیوسته‌ای را خواهد داشت. در عمل شبکه‌هایی با یک یا دو لایه‌ی پنهان، به صورت گسترده مورد استفاده قرار می‌گیرند و خوب کار می‌کنند. افزایش تعداد لایه‌های پنهان، سبب افزایش زمان محاسبات و زمان آموزش شبکه می‌شود. همچنین در مدل‌های پیش‌بینی، سبب مشکل «انطباق بیش از حد» خواهد شد. در بخش‌های بعد در مورد این مشکل، توضیحات بیشتر ارائه می‌شود.

کاسترا و بوید[1]، معتقدند قبل از افزودن سومین لایه‌ی پنهان، لازم است آزمایش‌هایی، با ورودی‌های مختلف به شبکه و تعداد کافی نرون در لایه‌های پنهان آن انجام گردد و در صورت نیاز لایه‌ی سوم، بعد از انجام این آزمایش‌ها اضافه گردد. همچنین بیان می‌کنند که همه‌ی کارهای تجربی که تا این زمان انجام شده، نشان می‌دهد که شبکه‌ای با بیش از چهار لایه، نتایج عملکرد شبکه را بهبود نخواهد داد.

افریم و آرونسون[2]، معتقدند استفاده از بیش از سه لایه‌ی پنهان، به ندرت در اغلب سیستم‌های تجاری دیده شده

1) Koastra and Boyd, 1996 2) Efraim and Aronsons, 1908

است. زیرا مقدار محاسبات به ازای هر لایه به صورت نمایی افزایش می‌یابد.

ژانگ و همکاران[1]، معتقدند اغلب محققین تنها یک لایه‌ی پنهان را جهت مقاصد پیش‌بینی به کار می‌برند.

ولی چون یک لایه‌ی پنهان، ممکن است به تعداد زیادی نرون احتیاج داشته باشد که سبب زیاد شدن زمان آموزش و بدتر شدن قابلیت تعمیم شبکه می‌شود، شبکه با دو لایه‌ی پنهان بهتر است. در ادامه به بحث در مورد تعداد نرون‌ها یا گره‌های پنهان- نرون‌هایی که در درون لایه‌های پنهان یا میانی قرارگرفته‌اند- پرداخته می‌شود.

ژانگ وهو[2]، معتقدند در طراحی یک شبکه‌ی عصبی، دو پارامتر بحرانی وجود دارد. این دو پارامتر به ترتیب اهمیت، تعداد گره‌های ورودی و پس از آن تعداد گره‌ها یا نرون‌های پنهان، معرفی شده‌اند. نرون‌های پنهان به شبکه‌ی عصبی اجازه می‌دهند، الگوهای غیرخطی درون داده‌ها را، کشف نماید. به صورت تئوری محدودیتی در تعداد نرون‌های درون لایه‌های پنهان وجود ندارد، ولی به ندرت دیده می‌شود که این تعداد، از دو برابر تعداد گره‌های ورودی تجاوز نماید.

ژانگ در کار تحقیقی دیگری که به همراه همکاران خود انجام داده، روش عمومی مورد استفاده در تعیین تعداد صحیح نرون‌های پنهان را، روش سعی و خطا معرفی می‌کند. همچنین چند راهنمایی عملی در این زمینه را بیان نموده است و معتقد است برخی از محققین برای تعداد نرون‌های پنهان، از رابطهٔ $۲n+۱$، برخی از $۲n$، برخی از n و برخی دیگر از $\frac{n}{۲}$ استفاده نموده‌اند، که n تعداد گره‌های ورودی است.

گاسترا و بوید، معتقدند هیچ فرمول «رویایی» برای محاسبه‌ی تعداد نرون‌های پنهان وجود ندارد، و استفاده از «قاعده‌ی هرم هندسی» ارائه شده توسط مسترز[3] را، یکی از راه‌های پیش‌رو می‌دانند. رابطه‌ی این قاعده به صورت زیر است:

$$n_{Hidden} = \sqrt{n_{Input}.n_{Output}}$$

در این رابطه n به ترتیب تعداد نرون‌های لایه‌ی پنهان، لایه‌ی ورودی و لایه‌ی خروجی را نشان می‌دهد. این محقق تعداد واقعی نرون‌های پنهان را بین نیم تا دو برابر تعداد بدست آمده از این رابطه می‌داند. کاسترا و بوید استفاده از یکی از سه رویکرد «ثابت»، «سازنده» و «مخرب» را برای تعیین تعداد بهینه‌ی تعداد نرون‌های پنهان، توصیه می‌کنند. در رویکرد ثابت، گروهی از شبکه‌ها با تعداد نرون‌های پنهان متفاوت، آموزش داده می‌شوند و سپس توسط ارزیابی یک متغیر نظیر MSE بر روی مجموعه داده‌های آزمایش، شبکه با بهترین تعداد نرون پنهان (کمترین MSE) انتخاب می‌شوند.

1) Zhang etal., 1998 2) Zhang and Hu, 1998 3) Masters