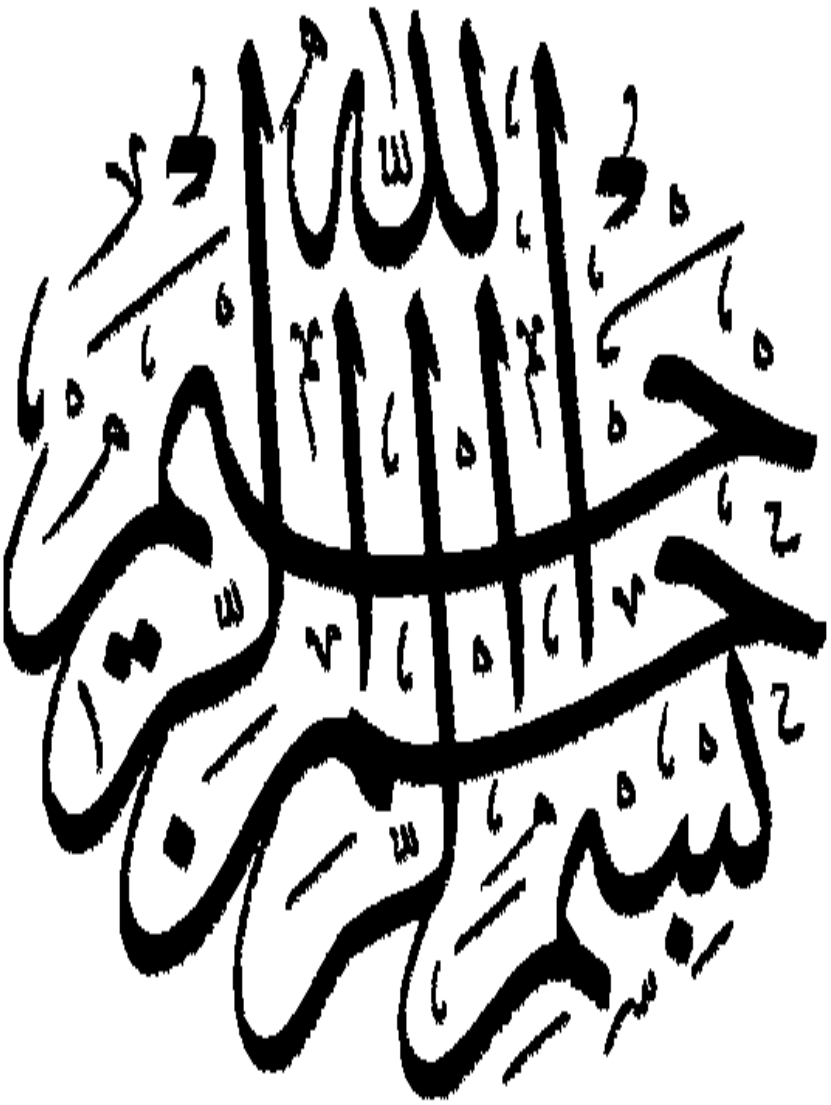

بسم الله الرحمن الرحيم

دانشگاه پیام نور – مرکز مشهد

پایان نامه جهت دریافت درجه کارشناسی ارشد ریاضی کاربردی
گرایش تحقیق در عملیات

حل مسائل بهینه سازی محدب با استفاده از شبکه های عصبی و کاربرد آن در حل مسائل مکمل غیر خطی

استاد راهنما
سرکار خانم دکتر عقیله حیدری

استاد مشاور
جناب آقای دکتر سهراب عفتی

نگارش
نرگس لطیفه‌گو

بهمن ۱۳۹۰

تقدیم به

پدر و مادر عزیزم که همراهان همیشگی من در ادامه تحصیل بوده و هستند. همچنین همسرم که مرا در این امر مهم یاری کردند.

تقدیر و تشکر

سپاس خدایی را که علم و دانش بر بندگان جویای حقیقت ارزانی داشت تا بر راهی که به نور تلاش و همّت روشن می‌شود گام بردارند. به بیان علامه شهید مطهری علم زیبایی عقل است. با علم انسان به سوی خدا تقرب می جوید و هر چه علمش افزون گردد تقربش به سوی خدای تعالی بیشتر می‌شود. اکنون که خداوند مرا به اتمام دوره‌ای دیگر از تحصیلاتم توفیق کرامت فرموده شایسته است از استاد گرانقدرم سرکار خانم دکتر عقیله حیدری نهایت تشکر را داشته باشم همچنین از استاد مشاورم آقای سهراب عفتی قدر دانی کنم و از اینکه از محضر این دو بزرگوار کسب فیض کردم خدا را شاکرم.

باسمه تعالی **مشخصـات پایان نامه تحصیلی دانشجویان دانشگاه پیـام نور**	

عنوان پایان نامه: بهینه سازی با استفاده از تابع فیشربرمیستر در شبکه‌های عصبی	
نام دانشجو: نرگس لطیفه‌گو	**مقطع تحصیلی:** کارشناسی ارشد
دانشکده: علوم پایه	**رشته تحصیلی – گرایش:** ریاضی-کاربردی
نام استاد راهنما: دکتر عقیله حیدری	**نام استاد مشاور:** دکتر سهراب عفتی
تاریخ دفاع: ۱۳۹۰/۱۱/۳۰	**تعداد صفحات:** ۸۲

چکیده پایان نامه: در این پایان‌نامه شبکه‌های عصبی برای بهینه‌سازی استفاده شده است. کاربرد این نوع شبکه‌ها را در مسائل مختلف بهینه‌سازی از جمله مسائل خطی و غیرخطی و مسائل درجه دوّم و مسائل مکمل غیرخطی را بیان کرده‌ایم. ایده اصلی از تقریب‌سازی شبکه‌های عصبی برای مسائل بهینه‌سازی، ساخت یک تابع انرژی و برقرار کردن یک دستگاه دینامیکی برای نشان دادن یک شبکه عصبی مصنوعی است. دستگاه دینامیکی به کاربرده شده از نوع معادلات دیفرانسیلی مرتبه اوّل می‌باشد.

نقطه تعادل این دستگاه با جواب بهینه مسئله مورد نظر با یک نقطه شروع اوّلیه معادل است. مدل اوّل از شبکه عصبی معرفی شده مستقیماً از شرایط بهینگی برای یک مسئله بهینه‌سازی بدست می‌آید و مدل دوّم از شبکه عصبی معرفی شده از تبدیل یک مسئله مینیمم‌سازی ناقید به یک مسئله مکمل غیرخطی با استفاده از تعمیم تابع فیشربرمیستر بدست می‌آید. همچنین همگرایی در خط سیر این شبکه‌ها را بررسی کرده و بر روی پایداری‌های لیاپانوف و مجانبی و سراسری این شبکه‌های عصبی کار شده است و در نهایت چند مثال عددی برای فهم بیشتر موضوع آورده شده است.

کلمات کلیدی:

مسائل مکمل غیرخطی

شبکه عصبی

همگرایی نمایی

تعمیم تابع فیشربرمیستر

امضای استاد راهنما:

امضای استاد مشاور :

تاریخ:

اینجانب نرگس لطیفه‌گو دانشجوی ورودی سال ۱۳۸۸ مقطع کارشناسی ارشد رشته ریاضی کاربردی گواهی می‌نمایم چنانچه در پایان نامه خود از فکر، ایده و نوشته دیگری بهره گرفته‌ام با نقل قول مستقیم یا غیر مستقیم منبع و مأخذ آن را نیز در جای مناسب ذکر کرده‌ام. بدیهی است مسئولیت تمامی مطالبی که نقل قول دیگران نباشد بر عهده خویش می‌دانم و جوابگوی آن خواهم بود.

دانشجو تأیید می‌نماید که مطالب مندرج در این پایان نامه (رساله) نتیجه تحقیقات خودش می‌باشد و در صورت استفاده از نتایج دیگران مرجع آن را ذکر نموده است.

نام و نام خانوادگی دانشجو
تاریخ و امضاء

اینجانب نرگس لطیفه‌گو دانشجوی ورودی سال ۱۳۸۸ مقطع کارشناسی ارشد رشته ریاضی کاربردی گواهی می‌نمایم چنانچه بر اساس مطالب پایان نامه خود اقدام به انتشار مقاله، کتاب و ... نمایم ضمن مطلع نمودن استاد راهنما با نظر ایشان نسبت به نشر مقاله، کتاب و ... و به صورت مشترک و با ذکر نام استاد راهنما مبادرت نمایم.

نام و نام خانوادگی دانشجو
تاریخ و امضاء

کلیه حقوق مادی مترتب از نتایج مطالعات، آزمایشات و نوآوری ناشی از تحقیق موضوع این پایان نامه متعلق به دانشگاه پیام نور می باشد.

فصل اوّل
آشنایی با شبکه‌های عصبی

مقدمه ۱
۱-۱ تاریخچه ۲
۱-۲ شبکه عصبی زیستی ۴
۱-۳ شبکه عصبی مصنوعی (ANN) ۵
۱-۴ مزیت‌های شبکه‌های عصبی ۶
۱-۵ طبقه بندی الگو ۶
۱-۶ فراگیری و تعمیم ۷
۱- ۷ آشنایی با مدل نورون و معماری شبکه‌های عصبی ۷
۱-۸ توابع انتقال ۸
۱-۹ نورون با یک بردار به عنوان ورودی ۱۱
۱-۱۰ معماری شبکه‌های عصبی ۱۱
۱-۱۱ ورودی‌ها و لایه‌ها ۱۲
۱-۱۲ شبکه‌های چند لایه ۱۳
۱-۱۳ شبکه‌های عصبی پرسپترون ۱۴
۱-۱۴ قاعده یادگیری پرسپترون ۱۵
۱-۱۵ ایجاد قواعد یادگیری ۱۷
۱-۱۶ آموزش پرسپترون‌های چند نورونی ۲۱
۱-۱۷ محدودیت‌های پرسپترون ۲۱
۱-۱۸ شبکه عصبی هاپفیلد ۲۴

فصل دوّم
پایداری در سیستم‌های دینامیکی

مقدمه ۲۷
۲-۱ مفاهیم پایداری ۲۸
۲-۲ پایداری و ناپایداری ۲۸
۲-۳ نظریه لیاپانوف ۲۹
۲-۴ پایداری مجانبی و پایداری نمایی ۳۰
۲-۵ پایداری موضعی و پایداری کلی ۳۱
۲-۶ خطی‌سازی و پایداری موضعی ۳۱
۲-۷ توابع معین مثبت و توابع لیاپانوف ۳۲

۲-۸ قضایای نقطه تعادل ۳۳
۲-۹ روش کراسوفسکی ۳۴
۲-۱۰ روش گرادیان متغیر ۳۵
۲-۱۱ نظریه پیشرفته پایداری ۳۷
۲-۱۲ یکنواختی مفاهیم پایداری ۳۹
۲-۱۳روش مستقیم لیاپانوف در سیستم‌های وابسته به زمان ۳۹

فصل سوّم
بهینه‌سازی با استفاده از شبکه‌های عصبی

مقدمه ۴۲
۳-۱ پیشنیازها ۴۳
۳-۲ سطوح کارایی و نقاط بهینه ۴۵
۳-۳ مشتقات سویی ۴۷
۳-۴ مینیمم گیری ۴۷
۳-۵ شرایط لازم برای بهینگی ۴۸
۳-۶ توابع درجه دوّم ۴۹
۳-۷ بهینه‌سازی کارایی ۵۰
۳-۸ روش سریعترین کاهش شیب ۵۱
۳-۹ سرعت یادگیری پایا ۵۲
۳-۱۰ روش نیوتن ۵۳
۳-۱۱ گرادیان الحاقی ۵۳
۳-۱۲ یک مدل شبکه عصبی برای حل مسئله بهینه سازی مقیّد ۵۶
۳-۱۳ یک مدل شبکه عصبی بر اساس تعمیمی از تابع فیشربرمیستر برای مسائل مکمل غیر خطی ۵۸
۳-۱۴ ویژگی‌های مسائل مکمل غیرخطی (NCP) ۶۰
۳-۱۵ مدل شبکه عصبی ۶۲
۳-۱۶ همگرایی و پایداری از خط سیر ۶۴
۳-۱۷ نتایج عددی ۶۵
پیوستA. کد برنامه‌های اجرا شده در نرم افزار Matlab ۷۶
واژه نامه فارسی به انگلیسی ۸۱
مراجع ۸۳

فصل اوّل
آشنایی با شبکه‌های عصبی

مقدمه

یکی از روش‌های کارآمد در حل مسائل پیچیده، شکستن آن به زیر مسئله‌های ساده‌تر است که هر کدام از این زیربخش‌ها به نحو ساده‌تری قابل درک و توصیف باشند. در حقیقت یک شبکه، مجموعه‌ای از این ساختارهای ساده است که در کنار یکدیگر سیستم پیچیده نهایی را توصیف می‌کنند. شبکه‌ها انواع مختلفی دارند امّا همگی آن‌ها از دو مؤلفه تشکیل می‌شوند .

مجموعه‌گره‌ها. هر گره در حقیقت واحد محاسباتی شبکه است که ورودی‌ها را گرفته و بر روی آن پردازش انجام می‌دهد تا خروجی بدست آید. پردازش انجام شده توسط گره می‌تواند از ساده‌ترین نوع پردازش‌ها نظیر جمع کردن ورودی‌ها تا پیچیده-ترین محاسبات را شامل شود. در حالت خاص، یک گره می‌تواند خود شامل یک شبکه دیگر باشد.

مجموعه یال‌ها. این یال‌ها نحوه گذر اطلاعات بین گره‌ها را مشخص می‌کند در حالت کلی یال‌ها می‌توانند تک سویه[۱] یا دوسویه[۲] باشند. تعامل بین گره‌ها از طریق این یال‌ها سبب بروز یک رفتار کلی از سوی شبکه می‌گردد که چنین رفتاری به تنهایی در هیچ یک از مؤلفه‌های شبکه دیده نمی‌شود. جامع بودن این رفتار کلی بر عملکرد موجود در هر گره سبب تبدیل شبکه به یک ابزار توانمند می‌شود به عبارت دیگر، مجموعه ساده‌ای از مؤلفه‌ها وقتی در قالب یک شبکه باشند می‌توانند رفتاری از خود بروز دهند که هیچ یک از آن مؤلفه‌ها به تنهایی قادر به بروز چنین مشخصه‌ای نیست.

---

۱. Unidirection
۲. Bidiirectiona

۱-۱ تاریخچه

اوّلین کارهای مربوط به شبکه‌های عصبی به سال ۱۹۴۳ برمی‌گردد، زمانی که یک فیزیولوژیست اعصاب و یک ریاضیدان از دانشگاه MIT رساله‌شان را درباره نحوه عملکرد احتمالی نرون‌ها منتشر نمودند. در این رساله آن‌ها نشان دادند که می‌توان یک شبکه عصبی را فقط با استفاده از ریاضیات و الگوریتم پیاده‌سازی نمود. آن‌ها برای توضیح حدسشان از نحوه کار نرون‌ها در مغز، مدل ساده‌ای از شبکه عصبی با مدارهای الکتریکی ساختند.

تفسیر آن‌ها از شبکه عصبی آن بود که از اتصال مجموعه‌ای از واحدهای تصمیم‌گیری دودویی، می‌توان شبکه‌ای با قابلیت حل هر مسئله محاسباتی ایجاد نمود. در سال ۱۹۴۹ دونالد هب[۱] کتابی با عنوان "سازماندهی رفتار" منتشر نمود و آموزش را در شبکه‌های عصبی (فقط از نظر روانشناسی) معرفی نمود. او در این کتاب به مفاهیم نحوه یادگیری در انسان پرداخته و اشاره نموده است که وقتی نورونی توسط نورون دیگری که به آن متصل است تحریک می‌شود، اتصالشان قویتر می‌شود و در نتیجه احتمال اینکه این دو نورون دوباره تحریک شوند بیشتر می‌شود. کتاب هب به وسیله روانشناسان بسیار مورد استفاده قرار گرفت امّا متأسفانه مهندسین علاقه‌ای از خود نشان ندادند.

در سال ۱۹۵۹ برنارد ویدرو[۲] و مارکین هوب[۳] از استنفورد، دو مدل عصبی به نام‌های مدلاین[۴] و آدالاین[۵] ساختند. نام این دو مدل از استفاده آن‌ها از عناصر خطی تطبیقی چندگانه گرفته شده است برای تشخیص الگوهای دودویی ساخته شد به گونه‌ای که بتواند بعد از خواندن رشته‌های بیت جاری در یک خط تلفن، بیت بعدی را پیش‌بینی کند. مدلاین اوّلین شبکه عصبی بود که در دنیای واقعی مورد استفاده قرار گرفت. این شبکه یک فیلتر تطبیقی بود که پژواک‌ها را در خطوط تلفن حذف می‌کرد. این شبکه همچنان استفاده تجاری دارد. در سال ۱۹۶۲ قانونی برای به روز کردن وزن‌ها (آموزش) در شبکه عصبی ارائه شد.

ایده این قانون آن بود که وقتی یک پرسپترون (که در بخش ۱-۴ به آن اشاره خواهد شد) فعال شده دارای خطای بزرگی است می‌توان مقادیر وزن‌ها را طوری تنظیم کرد که خطا در شبکه یا حداقل در پرسپترون‌های مجاور توزیع شود. در همین سال، یک زیست‌شناس اعصاب به نام فرانک رزنبلانت[۶] از دانشگاه کُرنل، شروع به کار بر روی پرسپترون نمود. او به کار بر روی سیستم بینایی مگس علاقه‌مند بود و عقیده داشت که بیشتر پردازش‌هایی که مگس برای فرار انجام می‌دهد در چشمانش اتفاق می‌افتد. پرسپترونی که از تحقیقات رزنبلانت بیرون آمد، به صورت سخت افزاری ساخته شد و قدیمی‌ترین شبکه عصبی محسوب می‌شود که هنوز مورد استفاده قرار می‌گیرد. پرسپترون تک‌لایه به عنوان ابزاری مفید در دسته‌بندی مجموعه‌های داده به دو کلاس معرفی گردید. رزنبلانت همچنین برای قانون آموزش پرسپترون پایداری را اثبات نمود. توانایی شبکه‌های عصبی در تشخیص الگوها باور نکردنی بود. نتایج اخیر، موجی از علاقه‌مندی در دانشمندان ایجاد نمود ولی این علاقه دیری نپاید [۱۷].

---

۱. Donald Hebb
۲. Bernard Withrow
۳. Markin Hebb
۴. Madaline
۵. Adaline
۶. Frank Rosenblant

در سال ۱۹۶۹، ماروین مینسکی[۱] و سیمور پیرت[۲] در رساله شان و کتابی به نام پرسپترون‌ها نشان دادند که شبکه عصبی (تک‌لایه) درجداسازی مجموعه داده‌هایی که به صورت غیرخطی جداپذیرند ضعیف عمل می‌کند. مینسکی و پیرت ضعف‌های دیگری از شبکه‌های عصبی را نیز نشان دادند. به ویژه آنکه اگر تعداد ورودی‌های شبکه عصبی افزایش یابد، زمان آموزش شبکه به صورت نمایی افزایش می‌یابد و در نتیجه بر روی کارایی شبکه محدودیت ایجاد می‌کند.

آن‌ها همچنین اظهار داشتند که چندلایه کردن شبکه عصبی، تأثیری در حل محدودیت‌های گفته شده ندارد. پرسپترون نشان داده است که با وجود (یا حتی به دلیل) محدودیت‌های شدیدش ارزش پژوهش دارد. پرسپترون ویژگی‌های جذّاب زیادی دارد خطی بودن، تئوری جذّاب آموزش و سادگی مدل که یک نوع پردازش موازی است. دلیلی وجود ندارد که فرض کنیم هر کدام از این ویژگی‌ها به حالت چندلایه آن هم انتقال می‌یابد. با این وجود، عقیده داریم که روشن کردن (یا رد کردن حکم) شهودیمان مبنی بر بی‌نتیجه بودن بسط به سیستم چندلایه، یک مسئله مهم پژوهشی می‌باشد. (مینسکی و پیرت، ۱۹۶۹) و بسیاری از پژوهشگران، از تابع آموزشی استفاده می‌کردند که اساساً دارای مشکل بود چون در سراسر خط مشتقپذیر نبود. در نتیجه این عوامل، پژوهش و سرمایه‌گذاری در زمینه شبکه عصبی به شدت کاهش یافت. کُندی پژوهش‌ها در زمینه شبکه عصبی تا سال ۱۹۸۱ ادامه داشت. تنها اتفاق مهم و قابل ذکر در این سال‌ها معرفی الگوریتم آموزش پس انتشار خطا توسط پل وربوس[۳] در سال ۱۹۷۴ بود که کار با مسائل بزرگتر را ممکن می‌ساخت در این الگوریتم به یک نورون اجازه داده می‌شد که خطایش را به لایه‌های تشکیل دهنده شبکه پس انتشار کند.

این روش بعدها و در سال ۱۹۸۶ توسط دیوید روملهارت[۴]، جئوفری هینتون[۵] و رونالد ویلیامز[۶] به صورت کاملتری در آموزش شبکه‌های چند لایه و برای جداسازی داده‌هایی که به صورت خطی جداپذیر نیستند مورد استفاده قرار گرفت. البته این روش در کارهای دیوید پارکر ۱۹۸۲, ۱۹۸۵ و یان لیکان ۱۹۸۶ نیز آورده شده است. ولی در حقیقت افتخار ابداع این روش به پل وربوس می‌رسد که این روش را در رساله دکترایش در دانشگاه هاروارد سال ۱۹۷۴ ارائه نموده است. در سال ۱۹۸۲، اتفاقات زیادی موجب علاقه به شبکه عصبی شد. جان هپفیلد [۷] مقاله‌ای در آکادمی ملی علوم ارائه نمود.

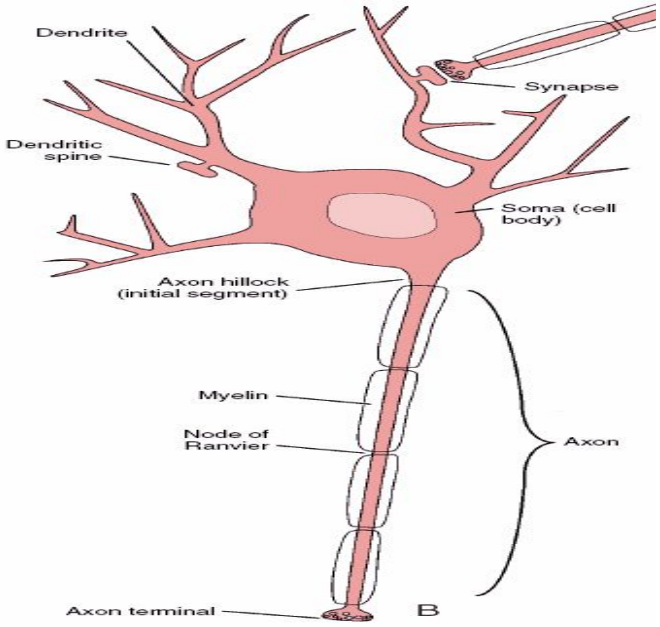
هدف هپفیلد ارائه مدل ساده‌ای از مغز نبود بلکه ایجاد ابزاری با کارایی بیشتر با استفاده از خطوط دوطرفه بین نرون‌ها بود . تا قبل آن، نرون‌ها فقط در یک مسیر به هم اتصال داشتند. در همان زمان، ریلی و کوپر از یک" شبکه هایبرید "چندلایه استفاده کردند که در آن هر لایه استراتژی حل مسئله مختلفی داشت. همچنین درسال ۱۹۸۲، اتفاق دیگری نیز روی داد. کنفرانسی آمریکایی - ژاپنی با عنوان شبکه‌های عصبی همیاری/ رقابتی در کیوتوی ژاپن برگزار شد. ژاپن اعلام کرد که نسل پنجم محاسبات وارد دوره جدید مبنی بر تلاش بر روی شبکه‌های عصبی شده است. با انتشار این خبرمجلات آمریکایی این ترس را ایجاد کردند که ممکن است آمریکا در این زمینه عقب بماند و طولی نکشید که سرمایه گذاری و در نتیجه پژوهش-های بیشتر دراین زمینه دوباره آغاز شد. (نسل اوّل شامل استفاده از سوئیچ‌ها و سیم‌ها بود، نسل دوّم از ترانزیستور بهره می‌برد، نسل سوّم از تکنولوژی حالت-جامد مانند مدارهای مجتمع و زبان‌های برنامه نویسی سطح بالاتر استفاد می‌کرد و نسل چهارم تولید کننده‌های کد بودندو پنجمین نسل محاسبات شامل هوش مصنوعی می‌شود).

---

۱ . Marvin Minsky
۲ . Seymour Papert
۳ . paul Vrbvs
٤ . David Rumelhart
٥ .Hyvfry Hinton
٦ .John Ronald Williams
۷ . John Hopfield

۱-۲ شبکه عصبی زیستی

شبکه‌های عصبی مصنوعی رابطه با نقطه مقابل طبیعی خود رابطه دارد. مغز انسان شامل بیش از $۱۰^{۱۱}$ نورون عصبی می‌باشد که از طریق حدود $۱۰^{۴}$ اتصال به ازای هر نورون به هم متصل شده‌اند. به طور کلی می‌توان سه قسمت اصلی برای نورون‌ها در نظر گرفت. دندریت، بدنه سلول وآکسون. **دندریت‌ها** دریافت کننده‌های درخت شکل از جنس فیبرهای عصبی هستند که سیگنال‌های الکتریکی را به بدنه سلول منتقل می‌کند. **بدنه سلول** این سیگنال‌ها را جمع کرده و یک حد آستانه بر روی آن‌ها اعمال می‌کند. در نهایت **آکسون** یک فیبر عصبی بلند است که سیگنال‌ها را از بدنه سلول عصبی به نورون‌های دیگر منتقل می‌کند به نقطه اتصال بین آکسون یک سلول عصبی با دندریت سلول‌های عصبی دیگر **سیناپس** گفته می‌شود. نحوه چینش نورون‌ها و قدرت سیناپس‌ها با توجه به فرایندهای پیچیده شیمیایی می‌گردد که آن‌ها در واقع بنا کننده تابع شبکه‌های عصبی می‌باشند. شکل زیر یک طرح ساده از نورون عصبی را نشان می‌دهد.



شکل ۱- ۱.

برخی از ساختارهای عصبی در هنگام تولّد تعریف می‌شوند و برخی دیگر در طول زندگی و از طریق فرایند یادگیری توسعه می‌یابند. در واقع **یادگیری** عبارت از ایجاد یال‌های جدید و یا قطع کردن برخی یال‌های عصبی قدیمی است. توسعه این روند در مراحل اوّلیه زندگی بسیار قابل توجه است. روند تغییرات ساختارهای عصبی در طول زندگی ادامه می‌یابد. تغییرات بعدی که در طول زندگی ایجاد می‌شود عمدتاً متأثر از قدرت و ضعف یال‌های سیناپسی هستند. به عنوان مثال ثابت شده که حافظه انسان با توجه به قدرت سیناپتیک شکل می‌گیرد. بنابراین فرایند یادگیری چهره یک دوست جدید با تغییر در قدرت چندین سیناپس ممکن می‌شود. شبکه عصبی مصنوعی فعلی هرگز به پیچیدگی مغز انسان نیستند امّا به هر حال دو شباهت اساسی بین شبکه‌های عصبی زنده و مصنوعی وجود دارد. شباهت اوّل در این است که ساختارها هر دو از یک ابزار محاسباتی با به هم پیوستگی بسیار بالا تشکیل شده‌اند (البته شبکه‌های عصبی خیلی ساده‌تر از گونه زنده آن می‌باشند). شباهت دوّم این است که در هر دو مورد یال‌های بین نورون‌ها تعیین کننده تابع شبکه می‌باشند[۱۸].

۱-۳ شبکه عصبی مصنوعی (ANN)[۱]

آن‌چنان که بیان شد انواع مختلفی از شبکه‌ها وجود دارد. در این بین شبکه‌ای وجود دارد که نورون مصنوعی را به عنوان یک گره درنظر می‌گیرد. در اصطلاح، این چنین شبکه‌هایی را شبکه عصبی مصنوعی یا به اختصار ANN می‌نامند. یک شبکه عصبی مصنوعی ایده‌ای است برای پردازش اطلاعات که از سیستم عصبی زیستی الهام گرفته شده و مانند مغز به پردازش اطلاعات می‌پردازد. عنصر کلیدی این ایده، ساختار جدید سیستم پردازش اطلاعات است. این سیستم از شمار زیادی عناصر پردازشی فوق العاده به هم پیوسته تشکیل شده که برای حل یک مسئله به صورت هماهنگ عمل می‌کند. ANNها، نظیر انسان‌ها، با مثال یاد می‌گیرند.

یک ANN برای انجام وظیفه‌ای مشخص، مانند شناسایی الگوها و دسته‌بندی اطلاعات، در طول یک فرایند یادگیری، تنظیم می‌شود. در سیستم‌های زیستی یادگیری با تنظیماتی در یال‌های سیناپسی که بین اعصاب قرار دارد همراه است. این روش ANNها هم می‌باشد. شبکه‌های عصبی را می‌توان با اغماض زیاد، مدل‌های الکترونیکی از ساختار عصبی مغز انسان نامید. مکانیسم فراگیری و آموزش مغز اساساً بر تجربه استوار است. مدل‌های الکترونیکی شبکه‌های عصبی طبیعی نیز بر اساس همین الگو بنا شده‌اند و روش برخورد چنین مدل‌هایی با مسائل، با روش‌های محاسباتی که به‌طور معمول توسط سیستم‌های کامپیوتری در پیش گرفته شده‌اند، تفاوت دارد. می‌دانیم که حتی ساده‌ترین مغزهای جانوری هم قادر به حل مسائلی هستند که اگر نگوییم که کامپیوترهای امروزی از حل آن‌ها عاجز هستند، حداقل در حل آن‌ها دچار مشکل می‌شوند.

به عنوان مثال، مسائل مختلف شناسایی الگو، نمونه‌ای از مواردی هستند که روش‌های معمول محاسباتی برای حل آن‌ها به نتیجه مطلوب نمی‌رسند. درحالی که مغز ساده‌ترین جانوران به‌راحتی از عهده چنین مسائلی برمی‌آید. تصوّر عموم کارشناسان IT بر آن است که مدل‌های جدید محاسباتی که بر اساس شبکه‌های عصبی بنا می‌شوند، جهش بعدی صنعت IT را شکل می‌دهند. تحقیقات در این زمینه نشان داده است که مغز، اطلاعات را همانند الگوها ذخیره می‌کند. فرآیند ذخیره‌سازی اطلاعات به‌صورت الگو و تجزیه و تحلیل آن اساس روش نوین محاسباتی را تشکیل می‌دهند. این حوزه از دانش محاسباتی به هیچ وجه از روش‌های برنامه‌نویسی سنتی استفاده نمی‌کند و به جای آن از شبکه‌های بزرگی که به صورت موازی آرایش شده‌اند و تعلیم یافته‌اند، بهره می‌جوید.

شبکه‌های عصبی مصنوعی در واقع مثلثی هستند که ۳ ضلع مفهومی دارند:

- سیستم تجزیه و تحلیل داده‌ها
- نورون یا سلول عصبی
- شبکه یا قانون کار گروهی نورون‌ها

در یک تعریف کلاسیک هایکین[۲] می‌گوید" **شبکه عصبی** عبارت است از مجموعه‌ای عظیم از پردازشگرهای موازی که استعداد ذاتی برای ذخیره اطلاعات تجربی و بکارگیری آن دارند و این شبکه دست‌کم از دو بابت شبیه مغز است :

- مرحله‌ای موسوم به یادگیری دارد.
- وزن‌های سیناپسی جهت ذخیره دانش به کار می‌روند."

---

۱. Artificial Neural Network

۲ . Haykyn

۱-۴ مزیت‌های شبکه‌های عصبی

شبکه‌های عصبی، با قابلیت قابل توجه در استنتاج معانی از داده‌های پیچیده یا مبهم، برای استخراج الگوها و شناسایی روش-هایی که آگاهی از آن‌ها برای انسان و دیگر تکنیک‌های کامپیوتری بسیار پیچیده و دشوار است به کار گرفته می‌شوند. یک شبکه عصبی تربیت یافته می‌تواند به عنوان یک متخصص در مقوله اطلاعاتی که برای تجزیه و تحلیل به آن داده شده به حساب آید. از این متخصص می‌توان برای برآورد وضعیت‌های دلخواه جدید و جواب سؤال‌های" چه می‌شد اگر " استفاده کرد. مزیت‌های دیگر آن شامل موارد زیر می‌شود.

- یادگیری انطباق پذیر. قابلیت یادگیری نحوه انجام وظایف بر پایه اطلاعات داده شده برای تمرین و تجربه‌های مقدماتی.
- سازماندهی توسط خود . یک ANN می‌تواند خودش سازماندهی یا ارائه‌اش را، برای اطلاعاتی که در طول دوره یادگیری در یافت می‌کند، ایجاد کند.
- عملکرد بهنگام. محاسبات ANN می‌تواند به‌صورت موازی انجام شود و سخت افزارهای مخصوصی طراحی و ساخته شده است که می‌تواند از این قابلیت استفاده کند.
- تحمل اشتباه بدون ایجاد وقفه در هنگام کد گذاری اطلاعات. خرابی جزئی یک شبکه منجر به تنزل کارایی آن می-شود اگر چه تعدادی از قابلیت‌های شبکه ممکن است حتی با خسارت بزرگی هم باقی بماند.

۱-۵ طبقه بندی الگو

در حالی که مشغول خواندن این جمله هستید، مغز شما باید سیگنال‌هایی را که از چشمتان دریافت می‌کند از هم جدا کند، به طوری که بتواند حروف روی صفحه را تشخیص دهد و آن‌ها را به صورت کلمات، جملات، پاراگراف‌ها و غیره به هم بچسباند. عمل تشخیص حروف مجزا، تشخیص الگو است. نمادهای روی صفحه الگوهایی هستند که باید تشخیص داده شوند. چون این جملات چاپ شده است، تمام حروف، فرم چاپی مشخصی دارند. حال اگر بخواهیم ماشینی بسازیم که بتواند حروف دست نویس را بخواند چه اتفاقی می‌افتد؟ این مسئله به دلیل تنوع بسیار زیادی که حتی بین نمونه‌هایی از یک حروف وجود دارد خیلی مشکل است.

لذا برای هر حرف باید مجموعه‌ای از حروف نوعی ساخته شود. هر مجموعه از حروف یک **طبقه** نامیده می‌شود. اکنون باید ماشینی بسازیم، به نام تشخیص دهنده الگو، که تشخیص دهد یک الگو ورودی به کدام طبقه تعلق دارد. بنابراین، اگر مسئله‌ای را بتوان ذاتاً به صورت طبقه بندی الگو درآورد، می‌توان برای حل یا پیاده‌سازی آن از یک شبکه عصبی استفاده کرد. در موارد بسیاری این کار امکان‌پذیر است. عمل شبکه عصبی این است که الگوهای ورودی را دریافت کند و الگویی را که برای آن طبقه صحیح باشد در خروجی‌اش تولید کند. می‌توان گفت یک شبکه عصبی یک طبقه بندی کننده الگو است.

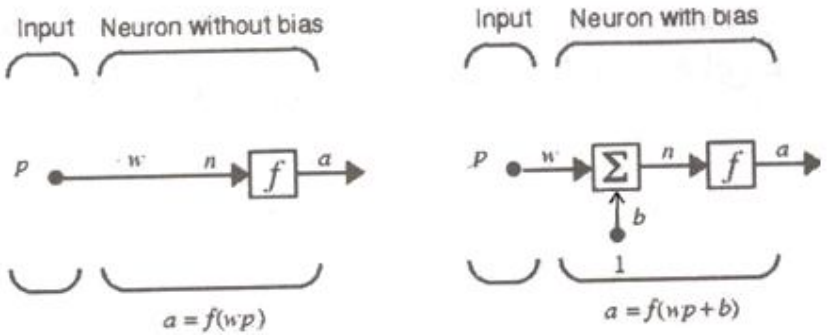
۱-۶ فراگیری و تعمیم

شبکه‌های عصبی طبقه بندی کننده الگو هستند، امّا همه طبقه‌بندی کننده‌های الگو شبکه‌های عصبی نیستند. بنابراین ویژگی‌هایی را معرفی می‌کنیم که شبکه‌های عصبی را از انواع دیگر طبقه‌بندی کننده‌ها متمایز می‌سازد. دو خاصیت از مهمترین خواص شبکه‌های عصبی، احتمالاً فراگیری و تعمیم آن‌هاست. به طور خلاصه یک شبکه عصبی باید بتواند :

- الگوها را طبقه بندی کند.
- به قدر کافی کوچک باشد تا به لحاظ فیزیکی قابل پیاده‌سازی باشد.
- با آموزش برنامه‌ریزی شود، بنابراین بایستی قابلیت فراگیری داشته باشد.
- با استفاده از مثال‌های نشان داده شده در طی آموزش تعمیم یابد.

۱- ۷ آشنایی با مدل نورون[۱] و معماری شبکه‌های عصبی

در شکل ۱-۲ مدل یک نورون با یک ورودی نمایش داده است. این نورون ساده از دو عنصر وزن w و تابع انتقال f تشکیل شده است که n ورودی تابع انتقال می‌باشد.



شکل ۱- ۲. مدل یک نورون

ورودی p به نورون اعمال شده از طریق ضرب در وزن w وزن‌دار می‌شود و حاصل به تابع انتقال[۲] f به عنوان ورودی اعمال شده و خروجی نهایی حاصل می‌گردد. با اضافه کردن بایاس[۳] (اوّلین وزن شبکه عصبی با ضریب یک) به ساختار نورون شکل قبل نورون بایاس‌دار به صورت نورون سمت راست ایجاد می‌شود. ورودی بایاس یک مقدار ثابت ۱ است. مقدار بایاس با حاصل wp جمع شده و در واقع تابع را به سمت چپ شیفت می‌دهد. منظور از f تابع انتقال می‌باشد. این تابع معمولاً یک تابع پلّه‌ای[۴] و یا یک تابع پیچیده[۵] می‌باشد که آرگومان n را دریافت کرده و خروجی a را تولید می‌کند. که به صورت زیر محاسبه می‌شود.

۱ . Neuron
۲. Transfer Function
۳ . bias
٤ .Step Function
٥ .Sigmoid Function

$$a = f(wp + b) , n = wp + b$$

به عنوان مثال در مورد $p = ۲ , w = ۳ , b = -۱.۵$ داریم :

$$a = f(۳(۲) - ۱.۵) = f(۴.۵)$$

b و w دو پارامتر تنظیم شونده در نورون‌ها می‌باشند.

با مقایسه مدل ارائه شده با مدل نورون‌های زنده می‌توان اندازه وزن w را متناظر با قدرت هر یال سیناپسی دانست همچنین سیگما با بدنه سلول متناظر است و در نهایت تابع انتقال و خروجی a متناظر با سیگنال خروجی نورون در آکسون می‌باشند. ایده اصلی شبکه‌های عصبی این است که با تغییر مقادیر w وb شبکه یک رفتار یا تعمیم را اتخاذ کند. توجه داشته باشید که بایاس ورودی نیست بلکه یک پارامتر قابل تنظیم نورون‌ها است[۱۸] .

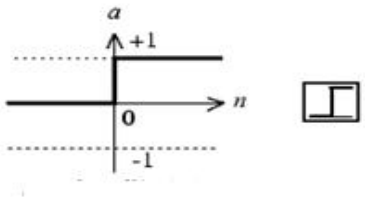
۱-۸ توابع انتقال

تابع انتقال یک تابع خطی و یا غیرخطی از n می‌باشد. از تابع انتقال برای تعیین خصوصیات نورون در راستای حل مسائل مختلف استفاده می‌شود.

تابع انتقال رایج در شبکه‌های عصبی عبارتند از :

- تابع انتقال Hard limit

$$a = hardlim(n) = \begin{cases} ۰ & n < ۰ \\ ۱ & n \geq ۰ \end{cases}$$

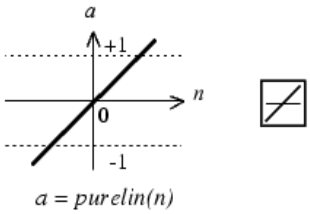


از این نوع تابع انتقال برای دسته‌بندی کردن ورودی‌ها به دو کلاس محدود استفاده می‌شود. از این تابع عمدتاً در نورون‌های سازنده شبکه‌های پرسپترون استفاده می‌شود.

- تابع انتقال خطی[۱]

این تابع انتقال در شکل زیر نشان داده شده که در آن $purelin(n) = n$ می‌باشد

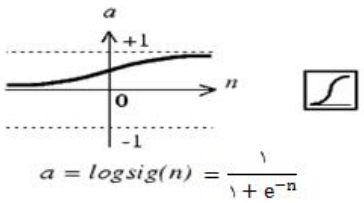
۱. Purelin



نورون‌هایی که از این تابع انتقال استفاده می‌کنند برای تقریب خطی در فیلترهای خطی به کار می‌روند.

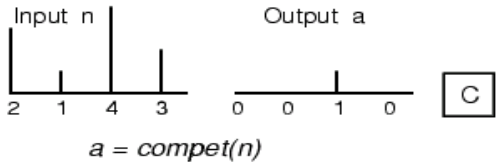
- تابع انتقال log-sigmoid

از این تابع انتقال در شبکه‌های پس انتشار که از جمله شبکه‌های چند لایه هستند استفاده می‌شود (در بخش ۱-۱۳ به شبکه-های چند لایه اشاره خواهد شد). تابع انتقال مقادیر ورودی را در محدود $-\infty$ تا $+\infty$ دریافت کرده و بر مبنای فرمول زیر یک مقدار خروجی بین ۰ و ۱ تولید می‌نماید.



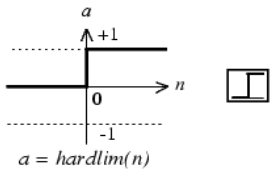
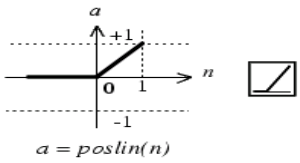
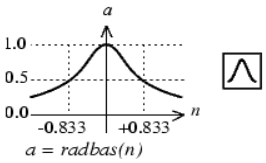
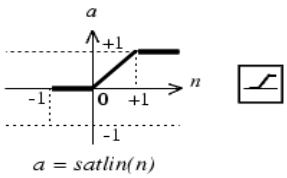
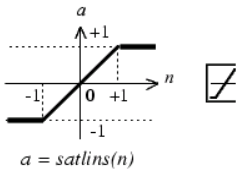
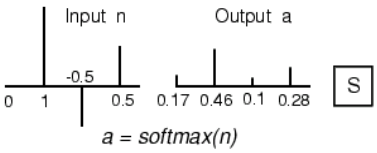
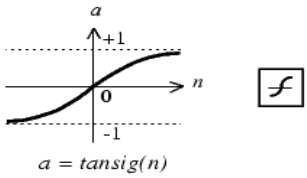
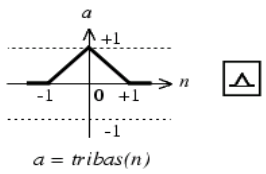
- تابع انتقال رقابتی[۱]

این تابع انتقال در شبکه‌های رقابتی استفاده می‌شود خروجی متناظر با بیشترین مقدار ورودی یک شده و بقیه صفر می‌شوند.



تعدادی دیگری از این توابع در جدول ۱-۲ زیر معرفی شده است :

۱ . Compet

تابع انتقال Hard limit متقارن	
تابع انتقال خطی مثبت	
تابع انتقال Radial Basis	
تابع انتقال Satlin	
تابع انتقال Satlins	
تابع انتقال Softmax	
تابع انتقال tan-sigmoid	
تابع انتقال Triagular Bas	

جدول ۱– ۱.