بسم الله الرحمن الرحيم

دانشگاه صنعتی اصفهان

دانشکده برق و کامپیوتر

تشخیص خطا به روش ماشین بردار پشتیبان

پایان‌نامه کارشناسی ارشد مهندسی برق-کنترل

امیر امینی

اساتید راهنما

دکتر جواد عسگری
دکتر فرید شیخ الاسلام

۱۳۸۸

دانشگاه صنعتی اصفهان

دانشکده برق و کامپیوتر

پایان نامه کارشناسی ارشد رشته برق-کنترل آقای امیر امینی

تحت عنوان

تشخیص خطا به روش ماشین بردار پشتیبان

در تاریخ ۱۳۸۸/۱/۳۰ توسط کمیته تخصصی زیر مورد بررسی و تصویب نهائی قرار گرفت.

۱- استاد راهنمای پایان نامه — دکتر جواد عسگری

۲- استاد راهنمای پایان نامه — دکتر فرید شیخ الاسلام

۲- استاد مشاور پایان نامه — دکتر مازیار پالهنگ

سرپرست تحصیلات تکمیلی دانشکده — دکتر علیمحمد دوست حسینی

کلیه حقوق مادی مترتب بر نتایج مطالعات، ابتکارات و نوآوریهای ناشی از تحقیق موضوع این پایان نامه (رساله) متعلق به دانشگاه صنعتی اصفهان است.

تقدیم به همه رهروان راه علم، همـه دوسـتداران حق و عدالت و همه کسانی کـه بـا مـرگ سـرخ خویش در راه حق گام برداشتند و با دلی مالامال از عشق به پروردگار در خون خود خفتند.

تشکر و قدردانی

لازم می‌دانم ابتدا از پدر و مادر مهربانم که در زندگی تکیه‌گاه من بوده‌اند و اندک موفقیت خویش را از برکت دعای خیر ایشان دارم، تشکر و قدردانی نمایم.

از اساتید بزرگوارم، جناب آقایان دکتر جواد عسگری ودکتر فرید شیخ الاسلام که با رهنمودهایشان نه تنها در طول انجام پایان نامه، بلکه در تمام دوره همراه بنده بوده‌اند قدردانی می‌کنم. همچنین از استاد گرامی جناب آقای دکتر مازیار پالهنگ که در طول این دوره از نعمت مشاوره با ایشان بهره‌مند بوده‌ام, تشکر و قدردانی می‌کنم.

از اساتید بزرگوار آقایان دکتر موسوی و استاد ذاکری که زحمت داوری این پایان‌نامه را تقبل نمودند, سپاسگزارم. همچنین لازم است از جناب آقای دکتر علی‌محمد دوست حسینی، سرپرست تحصیلات تکمیلی دانشکده و سرکار خانم نکوئی به دلیل زحماتی که به خاطر اینجانب متحمل گشته‌اند، سپاسگزاری نمایم.

از دوستانم سعیدی، روحی، حیدری، محمدی، درویش زاده، قلی پور، سلیقه دار، رستمی، شیرازی، پیمانی، جلالی، رفیعی و اسماعیلی که در طول این دوره در کنار من بودند تشکر نموده و توفیق روز افزون آرزو می نمایم.

امیر امینی

فروردین ۱۳۸۸

فهرست مطالب

عنوان صفحه

۱-فصل اول مقدمه ۲

۱-۱ تاریخچه و کاربردهای ماشین بردار پشتیبان ۲

۱-۲ اهداف انجام پایان نامه ٤

۱-۳ روند ارائه مطالب ٥

۲-فصل دوم شبکه های عصبی ۷

۲-۱ تاریخچه و زمینه شبکه های عصبی ۸

۲-۲ پرسپترون برای دو دسته الگو ۹

۲-۳ الگوریتم های آموزش ۱۱

۲-٤ تابع معیار ۱۳

۳-فصل سوم ماشین بردار پشتیبان ۱٦

۳-۱ تئوری یادگیری آماری ۱٦

۳-۲ بعد VC ۱۹

۳-۳ ماشین بردار پشتیبان دو کلاسه ۲۰

۳-۳-۱ ابر صفحه جدا ساز ۲۰

۳-۳-۲ ابر صفحه جداساز بهینه ۲۰

۳-۳-۳ ابرصفحه جدا ساز بهینه تعمیم یافته ۲٦

۳-۳-٤ تعمیم به فضای دارای ویژگی ابعادی بالا ۲۸

۳-۳-٥ توابع هسته ۲۹

۳-۳-٦ سنجش عملکرد ۳۱

۳-۳-۷ چگونگی انتخاب هسته مناسب ۳۲

۳-۳-۸ مقایسه بین شبکه های عصبی مصنوعی و ماشین بردار پشتیبان ۳۲

۳-٤ ماشین بردار پشتیبان فازی(FSVM) ۳٤

۳-٤-۱ دخالت دادن میزان اهمیت نمونه ها ۳٤

۳-٤-۲ تصمیم گیری نرم با استفاده از منطق فازی ۳٥

٤-فصل چهارم تشخیص امضای صحیح با استفاده از ویژگی های تابعی در حوزه زمان ۳۸

٤-۱ روشهای تشخیص امضای صحیح ۳۸

٤-۱-۱ روشهای پارامتری ۳۹

٤-۱-۲ روشهای تابعی ۳۹

٤-۲ پیش پردازش ٤۰

٤-۲-۱ نمونه برداری دوباره ٤۰

٤-۲-۲ نرمالیزه کردن اندازه امضا ٤۱

٤-٢-٣ هموارسازی ٤١
٤-٢-٤ حذف چرخش ٤٢
٤-٣ استخراج ویژگیها ٤٤
٤-٤ محاسبه شباهت بین امضاها با استفاده از رگرسیون توسعه یافته ٤٥
٤-٤-١ مروری بر رگرسیون و رگرسیون توسعه یافته ٤٥
٤-٤-٢ پیچش زمانی پویاDTW ٤٧
٤-٤-٣ یکسانسازی طول زمانی سیگنالها با استفاده از تطابق همه نقاط آن ها ٥٠
٤-٤-٤ یکسانسازی طول زمانی سیگنالها با استفاده از تطابق نقاط اکسترمم آن ها ٥١
٤-٤-٥ نرمالیزه کردن انرژی سیگنالها و محاسبه شباهت بین امضاها ٥٣
٤-٥ جدا سازی امضاهای همراه با خطا و یا جعلی از امضاهای اصلی ٥٤
٤-٥-١ آموزش سیستم تشخیص امضا ٥٥
٤-٥-٢ تائید یا رد امضای ورودی ٥٥
٤-٦ خلاصه ٥٩
٥-فصل پنجم نتایج و شبیه سازی ها ٦٠
٥-١ بررسی اثرات تغییر پارامترها بر نرخ تشخیص ٦٠
٥-٢ مقایسه میان ماشین های بردار پشتیبان چند کلاسه ٦٥
٥-٣ تشخیص خطا در سیستم سه تانک ٦٨
٥-٤ بررسی روش خلماتوف و ارایه یک روش جدید با چند طبقه کلاس بندی ٧٠
٥-٤-١ انتخاب ویژگی ٧٠
٥-٤-٢ آموزش کلاس بندی ٧٢
٥-٤-٣ روش ارائه شده ٧٣
٦-فصل ششم نتیجه گیری و پیشنهادات ٧٦
٦-١ نتیجه گیری ٧٦
٦-٢ پیشنهادات ٧٧
٧-مراجع ٧٨

فهرست اختصارات

SVM	Support Vector Machine
SVC	Support Vector Classification
SVR	Support Vector Regression
SRM	Structural Risk Minimization
ERM	Emperical Risk Minimization
VC	vapnik chervonenkis
DP	Dynamic Programming
DTW	Dynamic Time warping
EER	Equal Error Rate
FAR	False Acceptance Rate
FRR	False Rejection Rate

چکیده

با افزایش پیچیدگی و پیشرفت سیستم های کنترلی و استفاده از آن ها در محیط ها و کاربردهای حساس، تمایل روزافزونی در زمینه تشخیص خطا ایجاد شده است. در گذشته شبکه های عصبی به عنوان ابزاری برای تشخیص مدل یا خرابی در یک سیستم به کار گرفته شده اند. اما مشکل الگوریتم بهینه سازی آن ها برای انتخاب پارامتر و کم کردن خطا در هر مرحله به جای کم کردن خطای کل مدل باعث شده است تا ماشین بردار پشتیبان جایگزین مناسبی برای آن ها شوند. ماشین بردار پشتیبان بر پایه تئوری یادگیری آماری وپنیک از جمله الگوریتم های یادگیری موفق در زمینه تشخیص و ایزوله نمودن خطا در سیستم های دینامیکی می‌باشد. یکی از مدل های مورد استفاده توسط ماشین بردار پشتیبان، سیستم های تائید امضا می باشد. هدف از تایید امضا ، جداسازی امضاهای جعلی از امضاهای اصلی است. استفاده از رگرسیون توسعه یافته، در مقایسه با فاصله اقلیدسی و DTW، معیار بهتری از میزان شباهت دو امضا بدست می‌دهد، برای این منظور باید طول زمانی سیگنال‌های متناظر دو امضا یکسان شود. استفاده از تطابق همه نقاط برای یکسان‌سازی طول زمانی این سیگنال‌ها سبب کاهش تمایز بین امضاهای اصلی و امضاهای جعلی می‌شود. در این پایان‌نامه برای حفظ تمایز بین امضاهای اصلی و امضاهای جعلی، روشی بر مبنای تطابق نقاط اکسترمم برای یکسان‌سازی طول زمانی سیگنال‌ها ارائه شده است.همچنین به کمک الگوریتم ماشین بردار پشتیبان اقدام به تشخیص خطا در سیستم سه تانک نموده و میزان دقت این الگوریتم در تشخیص خطا را با چند الگوریتم آموزش شبکه عصبی از جمله RBF و بازگشتی مقایسه می نماییم. در روال کار این پایان نامه قصد داریم پس از انجام مطالعات لازم در مورد روابط و معادلات بهینه سازی ماشین بردار پشتیبان در جهت جداسازی کلاس ها از یکدیگر، به شرح جزئیات ماشین بردار پشتیبان در فرم های خطی و غیر خطی پرداخته وبا یک سری شبیه سازی ها اثر پارامترهای مختلف (ازجمله پارامترهای هسته و پارامتر C وبردارهای پشتیبان و..) را در میزان شناسایی و تخمین خطا بررسی کنیم.

واژه های کلیدی:۱-ماشین بردار پشتیبان۲-تئوری یادگیری آماری۳-پیچش زمانی پویا٤-تطابق نقاط اکسترمم

فصل اول
مقدمه

در ابتدای سال های ۱۹۷۰ با افزایش پیچیدگی و پیشرفت سیستم های کنترلی و استفاده از آن ها در محیط ها و کاربردهای حساس، تمایل روزافزونی در زمینه تشخیص خطا[۱] ایجاد شده است. عدم تشخیص به موقع خطا به وسیله کاربر در سیستم های حساس، منجر به صدمه دیدن و از بین رفتن مقادیر قابل توجهی از امکانات و اطلاعات خواهد شد. در حال حاضر با توجه به پیشرفت فناوری، ضرورت هوشمندی سیستم کنترلی در تشخیص خطا بیشتر احساس می‌شود و لذا روشی مناسب که یک طراحی پایدار را ارائه کرده و قابلیت اعتماد سیستم کنترلی را افزایش دهد، مورد نیاز است [۱].

۱-۱ تاریخچه و کاربردهای ماشین بردار پشتیبان[۲]

یادگیری ماشین به درستی یکی از زیر مجموعه های چند منظوره[۳] می‌باشد و مفاهیم متعددی همچون هوش مصنوعی، پیچیدگی محاسباتی، تئوری کنترل، علم ادراک[٤]، فلسفه، روان شناسی و آمار را می‌توان از آن جمله تصور نمود. به طور خود کار ماشین یادگیری به ساخت یک مدل از یک فرآیند یا سیستم در میان آنالیزهای موجود

---

[۱] Fault Detection
[۲] Support Vector Machine
[۳] Multi Disciplinary
[٤] Cognitive Science

از الگو، می‌پردازد و یا قادر است مدل جدیدی بر اساس داده هایی که بدست می‌آورد طراحی کند تا خروجی یک سیستم مشاهده پذیر را پیش بینی کند. این توانایی یادگیری در پیش بینی درست، بدون مشاهده ورودی را توانایی تعمیم[۱] می‌نامند و همه روش های یادگیری ماشین در پی بهبود توانایی تعمیم خود می‌باشند.

از جمله مشکلات اولیه که یادگیری ماشین با آن مواجه است عبارتند از [۲] : ۱- اندازه دسته داده ها محدود و گاهی اوقات کوچک است، ۲- نمونه برداری به صورت غیر یکنواخت صورت می‌گیرد،۳- بعد داده ها به طور طبیعی زیاد می‌باشد. علاوه بر این می‌توان به مشکلات خطای اندازه گیری، نویز و مقادیر اشتباه در داده ها نیز اشاره نمود. به دنبال این موضوع الگوریتم های بهینه سازی برای انتخاب پارامتر و اندازه گیری آماری در انتخاب بهترین مدل به کار برده شدند. در گذشته مدل سازی تجربی بر اساس مینیمم خطای تجربی[۲] صورت می‌گرفت که بر اساس مینیمم سازی خطای اطلاعات آموزشی در هر مرحله عمل می‌کرد. این خود از جمله مشکلات بهینه سازی می‌باشد که از اصل مینیمم سازی محلی[۳] در اکثر الگوریتم های آموزش شبکه های عصبی استفاده می‌شود.

بر این اساس در اواخر دهه ۸۰ و اوایل دهه ۹۰ نیاز شدیدی در پیدا کردن یک ماشین یادگیری کارآمد و قدرتمند احساس شد که بر اساس مینیمم خطای تجربی نباشد. بدین منظور به یکی از روش های مناسب برای حل مشکلات فوق به نام مینیمم خطای ساختاری[۴] دست پیدا کردند که اساس آن بر محدود کردن باند بالایی خطا می باشد، که یک مینیمم سراسری[۵] را می یابد. این روش زیر بنایی برای بدست آوردن ابزاری قدرتمند در جهت تشخیص خطا به نام ماشین بردار پشتیبان شد.

ماشین بردار پشتیبان بر پایه تئوری یادگیری آماری وپنیک [۳]از جمله الگوریتم های یادگیری موفق در بسیاری از مسائل و کاربردها شد. از جمله اینکه ماشین بردار پشتیبان در آموزش مشکل بهینه های محلی را نداشت، دسته بندی کننده را با حداکثر تعمیم بنا می کرد، ساختار و توپولوژی خود را بصورت بهینه تعیین می نمود، و توابع تمایزغیرخطی را به راحتی و با محاسبات کم، با استفاده از هسته های[۶] غیر خطی و مفهوم حاصلضرب داخلی در فضای هیلبرت، تشکیل می‌داد. البته ماشین بردار پشتیبان دارای دو شاخه اصلی به نام های کلاس بندی و رگرسیون می باشد که در این پایا ن نامه در زمینه کلاس بندی کار شده است.

ماشین بردار پشتیبان یک شبکه پیش رونده عمومی است، یا به طور خاص تر، یک ماشین یادگیری خطی می باشد که قادر به پیش بینی خروجی در بسیاری از کاربردهای متفاوت می باشد. اولین بار از ماشین بردار پشتیبان در کلاس بندی الگوهای دوتایی استفاده شد [٤,٥,٦] . به دنبال افزایش توجه از این ابزار یادگیری، به توسعه و گسترش آن پرداختند که می توان از جمله کاربردهای آن به موارد زیر اشاره کرد: توانایی کار با دسته داده های

---

[۱] Generalization
[۲] Empirical Risk Minimisation
[۳] Local
[۴] Structural Risk Minimization
[۵] Global
[۶] Kernel

غیر خطی از طریق لم هسته [۷]، استفاده از کار یادگیری ماشین در مسائل رگرسیون [۸,۹,۱۰]، خوشه بندی[۱] [۱۱]، تشخیص الگو [۱۲]، یادگیری نیمه خودکار [۱۳]، تغییر ماشین بردار پشتیبان به فرم خطی و نرم یک [۱٤,۱٥,۱٦]. این پیشرفت کار با ماشین بردار پشتیبان همراه با گسترش الگوریتم های سریع همانند بهینه سازی مینیمم ترتیبی شده بود [۱۷]، در اواخر دهه ۹۰ تا کنون روش های نقطه یابی درونی برای دسته داده های بزرگ [۱۸]، و الگوریتم های شبه پرواز [۱۹] و تجزیه بر پایه دستاوردها [۲۰] برای ماشین بردار پشتیبان خطی جای خود را به عنوان به روزترین کاربردها از این ماشین پیدا کرده بودند.

ماشین بردار پشتیبان به طور موفق در بسیاری از کاربردها همچون دارویی، زیست شناسی، دارایی، روباتیک، بینایی ماشین، پردازش تصویر، شناسایی متون دست نویس و پیش پردازش متن ها به کار رفته است و به عنوان یکی از برجسته ترین تکنیک های یادگیری ماشین در زمان حال مطرح است.

۱-۲ اهداف انجام پایان نامه

در این پایان نامه به کمک الگوریتم ماشین بردار پشتیبان اقدام به تشخیص خطا در سیستم سه تانک و سیستم های تشخیص امضا نموده و میزان دقت این الگوریتم در تشخیص خطا را تعیین کرده و نتایج حاصله از تشخیص خرابی را با چند الگوریتم آموزش شبکه عصبی مقایسه می نماییم. در روال کار این پایان نامه قصد داریم پس از انجام مطالعات لازم در مورد روابط و معادلات بهینه سازی ماشین بردار پشتیبان در جهت جداسازی کلاس ها از یکدیگر، به شرح جزئیات ماشین بردار پشتیبان در فرم های خطی و غیر خطی پرداخته وبا یک سری شبیه سازی ها اثر پارامترهای مختلف (ازجمله پارامترهای هسته و پارامترC وبردارهای پشتیبان و..) را در میزان شناسایی و تخمین خطا بررسی کنیم.

با گسترش ماشین بردار پشتیبان از شکل دو کلاسه به حالت چند کلاسه می توانیم وجود خطا و نوع آن را در یک سیستم چند حالته (منظور وجود خطاهای مختلف در سیستم می باشد) تشخیص دهیم، برای این منظور روی دو نوع معروف از ماشین بردار پشتیبان های چند کلاسه با نام های زوج- زوج[۲] و یکی در مقابل بقیه[۳] کار خواهیم کرد وپس از شبیه سازی های مربوطه، این دو را از نقطه نظر دقت تشخیص خطا و زمان پاسخ دهی با هم مقایسه می نماییم. به علت وجود مناطق غیر قابل کلاس بندی در ماشین بردار پشتیبان های چند کلاسه به ماشین بردار پشتیبان فازی روی خواهیم آورد و با تعیین پارامترهایsi وmi برای آموزش و آزمون داده ها می توانیم به پاسخ مطلوب تری دست یابیم.

---

[۱] Clustering

[۲] Pairwise

[۳] One Against All

در تمامی شبیه سازی های مذکور سعی بر استفاده از داده های معتبر از پایگاه های اطلاعاتی خواهیم نمود تا جواب دهی ماشین بردار پشتیبان با واقعیت قابل فهم تر باشد. اما علاوه براین، اقدام به پیاده سازی یک مدل سه تانک به عنوان یک سیستم کنترلی ساده و غیر خطی و پرکاربرد می نماییم و با استخراج داده ها از شبیه سازی های صورت گرفته، ویژگی های مورد نیاز جهت آموزش ماشین بردار پشتیبان را بدست می آوریم و به کمک آن حالت نرمال رااز چهار حالت خطای ایجاد کرده در سیستم، تشخیص می دهیم.

یکی از کاربردهای قابل بحث در تحقیقات این پایان نامه کار برروی سیستم های تشخیص امضا برخط[۱] به کمک ماشین بردار پشتیبان خواهد بود تا بتواند امضا اصلی را از امضاهای همراه با خطا و یا جعلی تشخیص دهد. خطا در امضا ممکن است در اثر امضا کردن و یا دستگاه نمونه بردار امضا رخ دهد که عواملی همچون کثیفی دستگاه و یا قلم و همچنین عدم دقت در نمونه برداری از امضا از جمله این خطاها می توانند باشند.

همچنین داده هایی که از امضاهای جعلی بدست می آیند بسیار به امضاهای اصلی نزدیک می باشند زیرا امضاهای جعلی به صورت حرفه ای جعل شده اند و باید بهترین ویژگی ها را انتخاب کرده و با آن ماشین بردار پشتیبان را آموزش دهیم. در ابتدا یک سری بهینه سازی ها از جمله نرمالیزه کردن داده ها(جهت هم اندازه کردن امضاها)، هموار سازی[۲] داده ها(در مواقعی که امضا همراه با لرزش دست به صورت جزئی صورت می گیرد)، همسان سازی طیف سیگنال داده ها از روش پیچش زمانی پویا[۳] (که یک روش جستجوگر پیش رونده خواهد بود و سعی بر پیدا کردن کوتاهترین مسیر می نماید) و ایجاد تمایز بیشتر میان امضاهای اصلی و امضاهای همراه با خطا ویا جعلی از روش تطابق نقاط اکسترمم (که می توان از آن به عنوان یکی از کارهای نو در این زمینه نام برد) استفاده کرده تا بهترین داده ها را در اختیار داشته باشیم و سپس اقدام به آمورش ماشین بردار پشتیبان و آزمون آن خواهیم نمود. اولین روش برای تشخیص امضا اصلی از امضاهای همراه با خطا و یا جعلی روش رگرسیون توسعه یافته یا ماکزیمم شباهت می باشد و پاسخ دارای ۱۰٪ خطا می باشد. در سال ۲۰۰٤ شخصی به نام خلماتوف[٤] توانست با تعیین ویژگی های موثرو کارا در تعیین امضا یک شخص، خطا را به ۵.۵٪ رساند در این پایان نامه قصد داریم با تلفیق دو روش مذکور میزان خطا را به کمک ماشین بردار پشتیبان کاهش دهیم.

۱-۳ روند ارائه مطالب

همان‌طور که درقسمت‌های قبل بیان شد، هدف از انجام این تحقیق، تشخیص خطا به کمک ماشین بردار پشتیبان می باشد اما باید برتری این ابزار نسبت به شبکه های عصبی که از به روزترین ابزار جهت کلاس بندی

---

[۱] Online
[۲] Smoothing
[۳] Dynamic Time Warping
[٤] Kholmatov

می باشد نشان داده شود. بنابراین لازم است ابتدا آشنایی مختصری با شبکه های عصبی داشته باشیم. از این‌رو در فصل دوم این پایان‌نامه، به معرفی شبکه عصبی، پرسپترون و الگوریتم های آموزش می پردازیم.

در فصل سوم، پس از انجام مطالعات لازم در مورد روابط و معادلات بهینه سازی ماشین بردار پشتیبان در جهت جداسازی کلاس ها از یکدیگر،به شرح جزئیات ماشین بردار پشتیبان در فرم های خطی و غیر خطی پرداخته و با یک سری شبیه سازی ها اثر پارامترهای مختلف (ازجمله پارامترهای هسته و پارامترC وبردارهای پشتیبان و..) را در میزان شناسایی و تخمین خطا بررسی می کنیم. همچنین به گسترش ماشین بردار پشتسبان در فرم های چند کلاسه وفازی می پردازیم. در فصل چهارم، به توضیح تشخیص امضای صحیح با استفاده از ویژگی های تابعی در حوزه زمان پرداخته و با یک سری بهینه سازی ها از جمله نرمالیزه کردن داده ها، هموار سازی داده ها، همسان سازی طیف سیگنال داده ها از روش پیچش زمانی پویا و ایجاد تمایز بیشتر میان امضاهای اصلی و امضاهای همراه با خطا و یا جعلی از روش تطابق نقاط اکسترمم استفاده کرده تا بهترین داده ها را در اختیار داشته باشیم و سپس اقدام به آمورش ماشین بردار پشتیبان و آزمون آن خواهیم نمود.

در فصل پنجم، به شبیه سازی و بررسی نتایج حاصله پرداخته و با به کار بردن جداول مقایسه ای دلایل برتری هر مرحله نسبت به مراحل قبل را مشخص می نماییم. در فصل ششم، نتیجه تحقیقات و پیشنهاداتی برای ادامه کارهای این تحقیق، ارائه خواهیم کرد.

فصل دوم
شبکه های عصبی

در سالیان اخیر شاهد حرکتی مستمر، از تحقیقات صرفاً تئوری به تحقیقات کاربردی بخصوص در زمینه پردازش اطلاعات، برای مسائلی که برای آن ها راه حلی موجود نیست و یا به راحتی قابل حل نیستند بوده ایم. با عنایت به این امر، علاقه فزاینده ای در توسعه تئوریک سیستم های دینامیکی هوشمند مدل- آزاد[۱]- که مبتنی بر داده های تجربی هستند ایجاد شده است. "شبکه عصبی مصنوعی[۲] "جزء این دسته از سیستم های دینامیکی قرار دارند، که با پردازش روی داده های تجربی، دانش یا قانون نهفته در ورای داده ها را به ساختار شبکه منتقل می کنند. به همین خاطر به این سیستم ها هوشمند[۳] گویند، چرا که براساس محاسبات روی داده های عددی یا مثال ها، قوانین کلی را فرا می گیرند. این سیستم ها در مدل سازی ساختار نروسیناپتیکی[۴] مغز بشر می کوشند[۲۱].

با اذعان به این مسائل، در این بخش قصد داریم به معنای شبکه های عصبی مصنوعی، حدود انتظارات ما از این شبکه ها و شباهت های آن ها به شبکه های واقعی بپردازیم. با امید به آن که در این فصل نمایی کلی از شبکه های عصبی در ذهن ما نقش بندد. در فصل سوم این پایان نامه به بیان نوعی جدید از ماشین های آموزشی با نام ماشین بردار پشتیبان می پردازیم و سپس مقایسه ای بین این دو مدل(شبکه های عصبی و ماشین بردار پشتیبان

---

[۱] Model Free
[۲] Artificial Neural Networks(ANN)
[۳] Intelligent
[۴] Neuro-synaptic

) انجام می دهیم. بنابراین در این فصل سعی بر بیان مسائل مقدماتی و کلی شبکه های عصبی داریم و از بیان جزئیات مربوط به شبکه های عصبی کناره گیری می کنیم.

۲-۱ تاریخچه و زمینه شبکه های عصبی

دسته بندی کننده بیز برای جامعه های گوسی به طور کامل با بردار میانگین وماتریس کوواریانس هر دسته مشخص می شود. الگوهایی که برای تخمین این عوامل استفاده شدند، معمولاً الگوهای آموزش خوانده می شوند و مجموعه چنین الگوهایی از هر یک از دسته ها، مجموعه آموزشی خوانده می شود. فرآیندی که در آن با استفاده از مجموعه آموزشی، توابع تصمیم بدست می آیند، یادگیری یا آموزش خوانده می شود.

الگوهای آموزش هر دسته به روش ساده ای برای محاسبه عوامل تابع تصمیم متناظر با آن دسته به کار می روند. بعد از این که عوامل مورد نظر تخمین زده شدند، ساختار دسته بندی کننده تعیین می شود. کارآیی نهایی تابع تصمیم به میزان برآورده شدن شرایط آماری که در به دست آوردن روش دسته بندی فرض شده اند بستگی دارد[۲۲].

خواص آماری دسته های الگو دریک مسئله غالباً نامعلوم هستند یا نمی توان آنها را تخمین زد. درعمل بهترین راه پرداختن به چنین مسایل تصمیم- نظریه ای، روش هایی هستند که از طریق آموزش مستقیماً توابع تصمیم موردنیاز را به دست می آورند. با این شرایط انجام فرض هایی راجع به شکل کلی توابع چگالی احتمال، یا سایر اطلاعات احتمالی درباره دسته های الگوی مورد بررسی، لازم نیست.

اساس مطلبی که در ادامه می آید، استفاده از تعداد زیادی عناصر محاسباتی غیر خطی (که نورون خوانده می شوند) می باشد. این عناصر به صورت شبکه هایی که مجسم کننده روش اتصال سلول های عصبی در مغز هستند سازمان یافته اند. مدل های حاصل با نام های گوناگونی شامل شبکه های عصبی، رایانه های عصبی، مدل های پردازش متوازی توزیع شده، سامانه های نورونی، شبکه های وفقی لایه ای و مدل های اتصالی مورد استفاده قرار می گیرند. در این جا از نام شبکه های عصبی استفاده می کنیم.

نخستین کاربرد علمی شبکه های عصبی در اواخر دهه ۵۰ قرن بیستم مطرح شد، زمانی که فرانک روزنبلات[۱] در سال ۱۹۵۸ شبکه پرسپترون را معرفی نمود. روزنبلات و همکارانش شبکه ای ساختند که قادر بود الگوها را از هم شناسایی نماید. در همین زمان بود که برنارد ویدرو[۲] در سال ۱۹۶۰ شبکه عصبی تطبیقی خطی آدلاین[۳] را با قانون یادگیری جدید مطرح نمود که از لحاظ ساختار، شبیه شبکه پرسپترون بود[۲۱] .

---

[۱] Frank Rosenbalt

[۲] Bernurd Widrow

[۳] Adaptive Linear Element(ADALINE)

وقتی پرسپترون ها با مجموعه های آموزشی جداناپذیر و خطی آموزش داده شوند، ظرف چند مرحله تکرار محدود به جواب همگرا می شوند. این جواب به شکل ضرایب ابر صفحه هایی بود که دسته های حاصل از الگوهای مجموعه آموزشی را به درستی از هم جدا می کردند. هر دوی این شبکه ها، پرسپترون و آدالاین، دارای این محدودیت بودند که توانایی طبقه بندی الگوهایی را داشتند، که به طور خطی از هم متمایز می شدند. ویدرو و روزنبلات هر دو از این امر آگاه بودند، چون آن ها قانون یادگیری را برای شبکه های عصبی تک لایه مطرح نموده بودند که توانایی محدودی جهت تخمین توابع داشتند. هر چند آن ها توانستند شبکه های چند لایه را مطرح نمایند، لکن نتوانستند الگوریتم های یادگیری شبکه های تک لایه را بهبود بخشند.

در این فصل مدل ساده ای از یک پرسپترون ارائه خواهد شد و سپس توضیح خواهیم داد که چگونه می توان الگوریتم آموزش را روی آن پیاده سازی کرد.

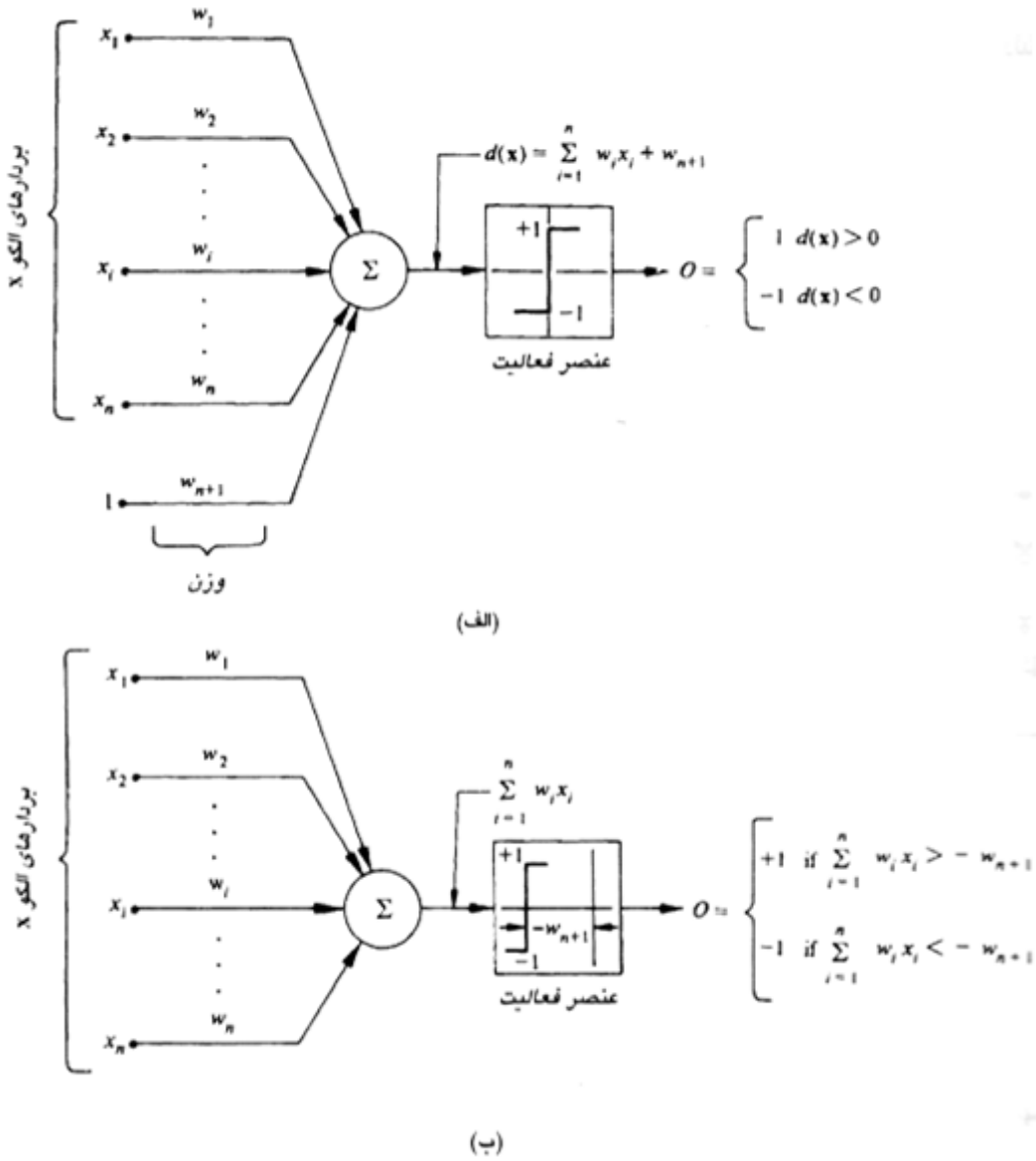
۲-۲ پرسپترون برای دو دسته الگو

پرسپترون یک تابع تصمیم خطی، که دو مجموعه آموزشی جداپذیر و خطی را از هم جدا می کند، از ابتدایی ترین شبکه های آموزشی عصبی می باشد. شکل۲-۱-الف مدل پرسپترون برای دو دسته الگو را نشان می دهد. پاسخ این الگوریتم بر مبنای جمع وزن دار ورودی هایش است یعنی،

$$d(x) = \sum_{j=۱} w_i x_i + w_{n+۱} \qquad (۱-۲)$$

که یک تابع تصمیم خطی بر حسب مولفه های بردارهای الگو است.

ابتدا ضرایب $I = ۱,۲,..,n,n+,w_i$ که وزن ها خوانده می شوند، در ورودی ها ضرب و سپس حاصل جمع وارد عنصر آستانه گیر می شوند. از این نظر، وزن ها مشابه سیناپس های سامانه عصبی انسان هستند. تابعی که خروجی جمع کننده را به خروجی نهایی می نگارند، گاهی اوقات تابع فعالیت خوانده می شود.



شکل۲-۱- دو نمایش معادل از مدل پرسپترون برای دسته های دو الگویی

وقتی $d(x) > ۰$ عنصر آستانه گیر باعث می شود که خروجی پرسپترون ۱+ باشد که نشان دهنده این است که الگوی x به عنوان عضو دسته $w_۱$ تشخیص داده شده است و وقتی $d(x) < ۰$ باشد خروجی پرسپترون ۱- می شود می شود که الگوی x به عنوان عضو دسته W۲ به حساب می آید.

وقتی $d(x) = ۰$، x روی سطح تصمیمی که دو دسته الگو را جدا می کند قرار دارد و یک وضعیت میان را می دهد مرز تصمیم پرسپترون با برابر صفر قرار دادن رابطه ی (۲-۱) بدست می آید

$$d(x) = \sum_{j=۱} w_i x_i + w_{n+۱} = ۰ \qquad (۲\text{-}۲)$$

$$w_۱ x_۱ + w_۲ x_۳ + \cdots + w_n x_n + w_{n+۱} = ۰ \qquad (۳\text{-}۲)$$

که معادله یک ابر صفحه در فضای الگوی n بعدی است. از نظر هندسی،n ضریب اول جهت ابر صفحه را تعیین می کنند و آخرین ضریب متناسب با فاصله عمودی تا ابر صفحه است. یعنی اگر $w_{n+1} = 0$،ابر صفحه از روی داده های الگو می گذرد، به طور مشابه اگر $w_j = 0$، ابر صفحه موازی محور x_j است.
خروجی عنصر آستانه گیر در شکل ۲-۱-الف به علامت $d(x)$ بستگی دارد. به جای آزمودن کل تابع برای تعیین مثبت یا منفی بودن آن، می توانیم بخش مجموع رابطه ی ۲-۱ را با w_{n+1} مقایسه کنیم، که در این مورد خروجی سامانه عبارت است از

$$O = \begin{cases} +1 & if \quad \sum_{i+1}^{n} w_i x_i > -w_{n+1} \\ -1 & if \quad \sum_{i+1}^{n} w_i x_i < -w_{n+1} \end{cases} \qquad (۲-۴)$$

این پیاده سازی معادل شکل ۲-۱-الف است که در شکل ۲-۱-ب نیز نشان داده شده است. تنها تفاوت این است که تابع آستانه ای به میزان $-w_{n+1}$ جابجا می شود و ورودی ثابت یک نیز دیگر وجود ندارد. رابطه دیگری که معمولاً در عمل به کار می رود، این است که بردارهای الگو را با افزودن عنصر (n+1)ام که بدون توجه به دسته دسته عضویت الگو، همواره معادل است، افزایش دهیم. یعنی، بردار الگوی افزایش یافته y از بردار الگوی x با قرار دادن $y_i = x_i$، برای i=۱,۲,…,n و افزودن عنصر اضافی $y_{n+1} = 1$ به دست می آید. آنگاه رابطه ی۲-۱ به صورت

$$d(y) = \sum_{i=1}^{n+1} w_i y_i = w^T y \qquad (۲-۵)$$

می شود که اکنون $y = (y_1, y_2, \dots, y_n, 1)^T$ بردار الگوی افزایش یافته است، و $w = (w_1, w_2, \dots, w_n, w_{n+1})^T$ بردار وزن خوانده می شود. این طرز بیان از نظر نوشتن مناسب تر است. با این وجود بدون توجه به رابطه مورد استفاده، مشکل اصلی، پیدا کردن w با استفاده از یک مجموعه آموزشی داده شده از هر دو دسته می باشد.

۲-۳ الگوریتم های آموزش

الگوریتم هایی که در زیر مطرح می شوند نماینده روش های متعددی هستند که در طول سال ها برای آموزش پرسپترون ها پیشنهاد شده اند. دسته های جداپذیر خطی، یک الگوریتم تکراری ساده را برای به دست آوردن بردار وزن نهایی در مجموعه آموزشی جداپذیر خطی پیاده سازی می کنند. ابتدا برای دو مجموعه آموزشی از بردارهای الگوی افزایش یافته که به ترتیب به دسته های الگوی متفاوت تعلق دارند، فرض کنید$w(1)$ بیانگر بردار وزن ابتدایی باشد، که می توان آن را به دلخواه انتخاب کرد. آنگاه درk امین مرحله تکرار، اگر $y(k) \in w_1$، $w^{\mathrm{T}}(\mathrm{k})\mathrm{y}(\mathrm{k}) \leq 0$ و $w(k)$ را با

$$w(k+1) = w(k) + cy(k) \qquad (۲-۶)$$