

# مقایسه روش‌های پرسپترون چندلایه، تابع اساسی شعاعی و بهینه‌سازی کمینه ترتیبی برای دسته‌بندی داده‌ها

مهدی نصیری<sup>۱</sup>، علی هادیان<sup>۲</sup>، مهسا نجف‌زاده<sup>۳</sup>، بهروز مینایی<sup>۴</sup>

## چکیده

در این مقاله سعی شده است که با در دست داشتن ۵ مجموعه داده‌ای خاص، دقت سه روش پرسپترون چندلایه‌ای، تابع شعاعی اساسی و بهینه‌سازی کمینه ترتیبی برای دسته‌بندی داده‌ها با در نظر گرفتن معیار دقت مورد مقایسه قرار گیرند. تأثیر عواملی مانند اندازه مجموعه داده، تعداد صفات دودویی و عددی آن و درصد داده‌ای که در شرایط استقلال مجموعه داده از مسئله، برای آزمون استفاده شده‌اند، مورد بررسی قرار گرفته است.

برای یک مجموعه داده خاص این روش‌ها در دو حالت بررسی شده‌اند. در یک حالت داده آموزشی و داده آزمایشی یکی در نظر گرفته شده‌اند و در حالت دیگر ۳۰٪ از داده، آزمایشی و ۷۰٪ آن آموزشی می‌باشد.

## کلمات کلیدی

شبکه عصبی، آموزش، پرسپترون چندلایه‌ای، تابع شعاعی اساسی، بهینه‌سازی کمینه ترتیبی، مجموعه داده، داده‌های آزمایشی، داده‌های آموزشی

## ۱. مقدمه

از سال ۱۹۴۰، نوروفیزیولوژیست‌ها سعی می‌کردند سیستم یادگیری و تجزیه و تحلیل مغز را کشف کنند و از سویی ریاضی‌دانان تلاش می‌کردند تا مدل ریاضی بسازند که قابلیت فراگیری و تجزیه و تحلیل عمومی مسائل را دارا باشد. از آن زمان، بارها این اتفاق افتاد که ریاضی‌دانان یافته‌های نوروفیزیولوژیست‌ها را پیاده‌سازی کردند، بدون این که دلیل آن را بدانند. آنها در عمل مشاهده کردند که سیستم پیاده شده کارایی شگفت‌انگیز سیستم طبیعی را دارد.

در سال ۱۹۴۳ مدل اولیه یک نرون پیشنهاد شد. در سال ۱۹۵۸ روز نبلدت یک استفاده عملی از نرون مصنوعی را نشان داد. او یک شبکه عصبی را برای شناسایی یک الگو (یک حرف الفبا) به کار برد. چون نظریه او از بینایی انسان نشأت گرفته است، اسم آن را پرسپترون نامید. اگر چه از همان ابتدا، ریاضی‌دانان توانسته بودند مدل ریاضی یک سلول عصبی یا نرون را بسازند، اما تا سال ۱۹۷۴ هنوز دانش مربوط به نوع اتصال این واحدهای شبه نورونی به یکدیگر تکامل لازم را نیافته بود.

امروزه برنامه‌های کاربردی متعددی در دسترس هستند که با روش‌های عصبی کار می‌کنند. هرچند کاربرد این برنامه‌ها بویژه برای افراد عادی کمی مشکل است، اما محققین روز به روز بیشتر آنها را به کار می‌گیرند. و چون عمل تجزیه و تحلیل و یادگیری در مغز شبکه اتفاق می‌افتد نه در مغز محقق، برای تجزیه و تحلیل یک سیستم پیچیده بوسیله روش شبکه‌های عصبی، نیاز به دانش زیادی درباره سیستم مورد مطالعه نمی‌باشد. اما به هر حال بهره‌گیری از دانش کلی درباره طرز کار این شبکه‌ها برای کاربران آنها ضروری است، چرا که تنظیمات ساده و کلی در این برنامه‌ها وجود دارند که آگاهی از آنها برای ساختن یک مدل موفق ضروری است.

<sup>۱</sup> دانشجوی کارشناسی ارشد هوش مصنوعی دانشگاه علم و صنعت ایران [nasiri@comp.iust.ac.ir](mailto:nasiri@comp.iust.ac.ir)

<sup>۲</sup> دانشجوی کارشناسی مهندسی نرم‌افزار دانشگاه علم و صنعت ایران [hadian@comp.iust.ac.ir](mailto:hadian@comp.iust.ac.ir)

<sup>۳</sup> دانشجوی کارشناسی ارشد نرم‌افزار دانشگاه علم و صنعت ایران [mnajfzade@iust.ac.ir](mailto:mnajfzade@iust.ac.ir)

<sup>۴</sup> استادیار دانشکده مهندسی کامپیوتر دانشگاه علم و صنعت ایران [b\\_minaei@iust.ac.ir](mailto:b_minaei@iust.ac.ir)

شبکه های عصبی مصنوعی دارای ۳ مفهوم هستند: ۱- سیستم تجزیه و تحلیل داده ها ۲- نرون یا سلول عصبی ۳- شبکه یا قانون کار گروهی نرون ها. در یک تعریف کلاسیک، شبکه های عصبی مجموعه ای از عناصر پردازش گر ساده که به هم متصل اند. عناصر پردازش در شبکه های عصبی نسبت به پردازش گرهای معمولی، بسیار ساده ترند با این تفاوت که تعداد آنها بسیار زیاد است.

هر کدام از نرون ها با تعدادی از نرون های دیگر اتصال مستقیم و مستقل دارد و وزن این اتصالات میزان ارتباط آن ها را تعیین می کند و اطلاعات در این وزن ها قرار می گیرند.

شبکه های عصبی دارای ویژگی های زیر است:

پردازش واحد انجام می دهند.

۲- توازی زیادی دارند.

۳- قسمت مجازی حافظه را ندارد و اطلاعات در یک مجموعه از وزن ها ذخیره می شوند.

۴- از بین رفتن بخشی از شبکه باعث خرابی و از کار افتادن شبکه نمی شود و نسبت به نویز و خرابی سخت افزار مقاوم است.

در مقایسه با شبکه های عصبی مصنوعی، روش ماشین بردار پشتیبانی از روش های نسبتاً جدیدی است که در سال های اخیر کارایی خوبی نسبت به شبکه های عصبی پرسپترون نشان داده است. مبنای کاری دسته بندی کننده ماشین بردار پشتیبانی، دسته بندی خطی داده ها است و در تقسیم خطی داده ها سعی می کند خطی را انتخاب کند که حاشیه اطمینان بیشتری داشته باشد. حل معادله پیدا کردن خط بهینه برای داده ها به وسیله روش های برنامه نویسی درجه دوم که روش های شناخته شده ای در حل مسائل محدودیت دار هستند صورت می گیرد.

ماشین بردار پشتیبانی اصولاً یک ماشین خطی است که ایده اصلی آن ایجاد یک فوق صفحه به عنوان سطح تصمیم گیری می باشد، به طوری که حد تفکیک بین نمونه های مثبت و منفی حداکثر شود. این روش با استفاده از یک شیوه که بر پایه تئوری آموزش آماری بنا نهاده شده، به خصیصه های بهینه فوق دست پیدا می کند. به صورت دقیقتر ماشین بردار پشتیبانی یک پیاده سازی تقریبی از روش "حداقل کردن ریسک ساختاری" است. اما این ماشین ها نمی توانند به آسانی مسئله برنامه نویسی درجه دوم<sup>۱</sup> را در مقیاس بزرگ با تکنیک های استاندارد برنامه نویسی درجه دوم حل کنند. بنابراین الگوریتم جدیدی به نام بهینه سازی کمینه ترتیبی برای این امر به کار گرفته شد که از نظر مفهومی ساده و پیاده سازی آن آسان می باشد و در کل سریعتر از الگوریتم قبلی است.

در این مقاله سعی شده است تا دقت روش های پرسپترون چندلایه و توابع اساسی شعاعی را که جزو روش های شبکه عصبی مصنوعی هستند و نیز روش بهینه سازی کمینه ترتیبی را با به کار بردن آنها برای دسته بندی<sup>۲</sup> ۵ مجموعه داده خاص با یکدیگر مقایسه کنیم. ابتدا به طور مختصر با این سه روش آشنا می شویم و سپس نتایج حاصل را بیان می نماییم.

## ۲. پرسپترون چندلایه<sup>۳</sup>

یکی از انواع شبکه های عصبی، پرسپترون می باشد که بصورت های پرسپترون تک لایه و پرسپترون چند لایه موجود هستند. شبکه های عصبی پرسپترون جزو شبکه های عصبی پیش خور دسته بندی می باشند. پرسپترون تک لایه تنها می تواند مسایل مجزای خطی را دسته بندی کند و برای مسایل پیچیده تر لازم است که از تعداد بیشتری لایه استفاده کنیم. شبکه های پیش خور چند لایه از یک و یا تعداد بیشتری لایه های میانی تشکیل شده اند.

یک پرسپترون چند لایه شامل اجزاء سازنده لایه ها و وزن ها می باشد. رفتار شبکه نیز وابسته به ارتباط بین اعضاء است. در حالت کلی در پرسپترون چند لایه ۳ نوع لایه نرونی وجود دارد:

۱. لایه ورودی: دریافت اطلاعات خامی که به شبکه تغذیه شده است.

۲. لایه های پنهان: عملکرد لایه های پنهان به وسیله ورودی ها و وزن ارتباط بین آنها و لایه های پنهان تعیین می شود. وزن های بین واحدهای ورودی و پنهان تعیین می کند چه وقت یک واحد پنهان باید فعال شود.

۳. لایه خروجی: عملکرد واحد خروجی به فعالیت واحد پنهان و وزن ارتباط بین واحد پنهان و خروجی بستگی دارد [۵]، [۲].

پرسپترون چند لایه یک شبکه کاملاً بهم مرتبط می باشد زیرا هر نرون در یک لایه به تمامی نرونهای لایه بعدی مرتبط می باشد. اگر بعضی از این ارتباطات وجود نداشته باشد شبکه، یک شبکه مرتبط ناقص است. زمانی که می گوئیم شبکه از  $n$  لایه تشکیل شده است، ما تنها لایه های میانی و لایه خارجی را شمرده ایم و لایه ورودی شمارش نمی شود چرا که این نرون ها محاسبه ای را انجام نمی دهند. بنابراین شبکه تک لایه شبکه ای با تنها یک لایه خارجی می باشد [۵].

## ۳. توابع اساسی شعاعی<sup>۴</sup>

توابع اساسی شعاعی نوعی از شبکه‌های عصبی مصنوعی هستند که در دهه‌ی ۸۰ پدید آمده‌اند. تابع اساسی شبکه در یک شبکه عصبی دو لایه‌ای قرار گرفته است که در آن هر واحد پنهان یک تابع فعال شعاعی را پیاده سازی می‌کند و واحدهای خروجی مجموع وزن‌داری از خروجی های واحد پنهان را پیاده سازی می‌نمایند. ورودی یک شبکه عصبی تابع اساسی شعاعی غیر خطی و خروجی این شبکه خطی است [۷]. شبکه‌های تابع اساسی شبکه به خاطر خصوصیت تقریبی غیرخطی خود برای مدل‌سازی نگاشت‌های پیچیده مناسب‌اند این در حالی است که شبکه‌های عصبی پرسپترون این نگاشت‌های پیچیده را به وسیله‌ی چندین لایه‌ی واسطه مدل می‌کنند [۶]. برای استفاده از یک شبکه‌ی تابع اساسی شعاعی، باید تابع فعالیت واحد پنهان، تعداد واحدهای پردازش، معیار مدل‌سازی و الگوریتم آموزش برای پیدا کردن پارامترهای شبکه مشخص شوند. پیدا کردن وزن‌های تابع اساسی شبکه، آموزش شبکه نامیده می‌شود و اگر مجموعه‌ای از جفت‌های ورودی- خروجی داشته باشیم تحت عنوان مجموعه‌ی آموزش به کار گرفته می‌شوند. برای این‌که خروجی‌های شبکه متناسب با ورودی‌ها باشند پارامترهای شبکه را بهینه‌سازی می‌کنیم این درخورسازی به وسیله یک تابع هزینه ارزیابی می‌شود. بعد از آموزش، شبکه‌ی تابع اساسی شبکه می‌تواند با داده‌ی آماری مشابه با مجموعه آموزش به کار گرفته شود [۸].

## ۴. بهینه‌سازی کمینه تریبی<sup>۵</sup>

ماشین بردار پشتیبانی یک الگوریتم کامپیوتری است که به وسیله‌ی مثال می‌آموزد که برچسب‌ها را به اشیاء اختصاص دهد. برای مثال، یک ماشین بردار پشتیبانی می‌تواند از طریق آموزش کلاهدرداری‌هایی را که از طریق کارت اعتباری روی می‌دهد به وسیله‌ی امتحان کردن صدها یا هزاران گزارش کلاهدرداری انجام شده با کارت اعتباری شناسایی کند [۹]. همان‌طور که می‌دانیم شبکه‌های عصبی مصنوعی یک مسیر اکتشافی را دنبال می‌کنند که در آن کاربردها و آزمایشات گسترده قبل از تئوری می‌آیند، ولی توسعه ماشین بردار پشتیبانی اول تئوری، سپس پیاده سازی و آزمایشات را دربرمی‌گیرد. ماشین بردار پشتیبانی در مقایسه با شبکه‌های عصبی مصنوعی دو امتیاز بیشتر یکی ساده بودن تفسیر هندسی آن و دیگری ارائه یک راه حل پراکنده دارد و برخلاف شبکه‌های عصبی مصنوعی پیچیدگی محاسباتی ماشین بردار پشتیبانی، وابسته به کیفیت فضای ورودی نیست. در گذشته ماشین پشتیبانی بردار به خاطر کارایی تعمیم عالی آن در زمینه‌ی یادگیری ماشین برجسته شده بود. امروزه استفاده از این ماشین فقط به تعداد کمی از گروه‌های تحقیقاتی محدود شده است به این دلیل که الگوریتم آموزش برای ماشین پشتیبانی بردار برای حل مسائل بزرگ، کند و پیچیده می‌باشد و پیاده‌سازی آن برای برنامه نویسان دشوار است. همچنین این ماشین‌ها نمی‌توانستند به آسانی مسئله برنامه نویسی درجه دوم را در مقیاس بزرگ با تکنیک‌های استاندارد برنامه نویسی درجه دوم حل کنند [۱۱]. بنابراین الگوریتم جدیدی به نام بهینه سازی کاهینه ترتیبی برای این امر به کار گرفته شد که از نظر مفهومی ساده و پیاده سازی آن آسان می‌باشد و در کل سریعتر از الگوریتم قبلی است و شرایط مقیاس‌گذاری بهتری برای حل مسائل ماشین بردار پشتیبانی در مقایسه با الگوریتم استاندارد آموزش ماشین بردار پشتیبانی ارائه می‌کند [۹]، [۱۲].

بهینه‌سازی کمینه ترتیبی دو ضریب لاگرانژ را که می‌توانند نسبت به یکدیگر بهینه شوند پیدا می‌کند و به طور تحلیلی مرحله‌ی بهینه را برای آن دو ضریب محاسبه می‌کند. وقتی که هیچ یک از این دو ضریب لاگرانژ نتوانند بهینه شوند، مسئله‌ی برنامه‌نویسی درجه دوم اصلی حل شده است. بهینه‌سازی کمینه ترتیبی دقیقاً "شامل دو مرحله است: ۱) مجموعه‌ای از اکتشافات برای انتخاب مؤثر جفت ضرایب لاگرانژ که قرار است روی آنها کار کنیم ۲) راه حل تحلیلی برای مسئله‌ی برنامه‌نویسی درجه دوم با سایز دو [۱۰].

## ۵. نتیجه‌گیری

در این تحقیق سه روش دسته‌بندی با استفاده از نرم افزار weka و ۵ نمونه داده‌ی خاص بر مبنای دقت مورد بررسی قرار گرفتند. در جدول ۱ مشخصات مجموعه داده‌ها با توجه به نوع صفات مشخص شده است. در جدول ۲ نتایج سه روش پرسپترون چندلایه، تابع اساسی شعاعی و بهینه‌سازی کمینه ترتیبی با توجه به ۵ مجموعه داده آموزشی مشخص شده است. در این جدول مجموعه داده‌های آموزشی و آزمایشی یکی هستند. در جدول شماره ۳ از ۷۰٪ مجموعه داده‌های جدول ۱ بعنوان داده‌های آموزشی و ۳۰٪ بعنوان داده آزمایشی انتخاب شده است. با توجه به نتایج به دست آمده می‌توان گفت:

اعتماد به دقت روش پرسپترون چندلایه در داده‌هایی که نوع صفات آن فقط اسمی است بیشتر از دو روش دیگر است.

اعتماد به دقت داده‌هایی که نوع صفات آن فقط عددی است در پرسپترون چندلایه بیشتر از دو روش دیگر است.

در داده‌هایی که تعداد نوع صفات اسمی آنها با عددی برابر است دقت دو روش بهینه‌سازی کمینه ترتیبی و تابع اساسی شعاعی برابر است و از روش پرسپترون چندلایه بهتر می‌باشد.

جدول ۱- مشخصات مجموعه داده‌ها

تعداد صفت عددی ۶	تعداد صفت اسمی ۷	مجموعه داده
۰	۳۸	۱) (soybean)
۴۲	۵	۲) (segment)
۴	۰	۳) (iris)
۳	۷	۴) (labor)
۲	۲	۵) (contact-lenses)

جدول ۲- مقایسه با داده‌های آموزشی و آزمایشی یکسان

بهبودسازی کمینه ترتیبی	تابع اساسی شعاعی	پرسپترون چندلایه	مجموعه داده
۹۷.۸۰	۹۷.۲۲	۹۹.۸۵	۱
۹۳.۹۸	۶۶.۷۵	۹۹.۹۶	۲
۹۶.۶۶	۹۷.۳۳	۹۷.۶۶	۳
۹۴.۷۷	۱۰۰	۱۰۰	۴
۱۰۰	۱۰۰	۹۵.۷۱	۵

جدول ۳- مقایسه در حالت: ۷۰٪ داده‌ها به صورت آموزشی و ۳۰٪ به عنوان داده آزمایشی

بهبودسازی کمینه ترتیبی	تابع اساسی شعاعی	پرسپترون چندلایه	مجموعه داده
۹۴.۶۳	۹۱.۲۲	۹۲.۶۸	۱
۸۴.۴۴	۸۸.۱۴	۸۷.۴۱	۲
۹۴.۲۸	۶۲.۹۱	۸۷.۲۳	۳
۶۵.۳۸	۵۹.۶۱	۶۵.۳۸	۴
۳۰.۷۶	۳۰.۷۶	۳۰.۷۶	۵

## ۶. مراجع

- [۱] Poggio,T.,Girosi,F; "Network for approximation and learning" proc. IEEE,vol. IEEE,vol.۷۸,no.۹,pp.۱۴۸۱-۱۴۹۷, ۱۹۹۰.
- [۲] Ripley, B. D.; "Neural networks and related methods for classification", Journal of the Royal Statistical Society, Series B (Methodological), ۵۶(۳), ۴۰۹-۴۵۶, ۱۹۹۴.
- [۳] Haykin, "Neural Networks: A Comprehensive Foundation", Second Edition, Prentice-Hall Inc., ۱۹۹۹.
- [۴] L. Xu, M-Y. Chow, X. Z. Gao; "Comparisons of Logistic Regression and Artificial Neural Network on Power Distribution Systems Fault Cause Identification"; IEEE Mid-Summer Workshop on Soft Computing in Industrial Applications, Finland, June ۲۸-۳۰, ۲۰۰۵.
- [۵] Gary William Flake; "Square Unit Augmented, Radially Extended, Multilayer Perceptrons, Neural Networks: Tricks of the Trade", this book is an outgrowth of a ۱۹۹۶ NIPS workshop, p.۱۴۵-۱۶۳, January ۱۹۹۸.
- [۶] Prak,J., Sandberg,J.W.; "Universal approximation using radial basis functions network" Neural Computation,vol.۳,pp.۲۴۶-۲۵۷, ۱۹۹۱.
- [۷] M.L.J. Orr; "Introduction to Radial Basis Function Networks", Centre for. Cognitive Science, University of Edinburgh, April ۱۹۹۶.
- [۸] Bors,A.G.,Gabbouj,G.; "Minimal topology for a radial basis function neural network for pattern classification", Digital Signal Processing: a review journal,vol.۴,no.۳,pp.۱۷۳-۱۸۸, ۱۹۹۴.
- [۹] S.R. Amendolia, G. Cossu, M.L. Ganadu, B. Golosio, G.L. Masala, G.M. Mura; "A comparative study of K-Nearest Neighbour, Support Vector Machine and Multi-Layer Perceptron for Thalassemia screening"; Chemometrics and Intelligent Laboratory Systems, pp. ۱۳- ۲۰, ۲۰۰۳.

- [10] Platt, J.; “*Sequential minimal optimization: A fast algorithm for training support. vector machines*”, Microsoft Research Technical Report MSR-TR-98-14, 1998.
- [11] G. W. Flake and S. Lawrence, “*Efficient SVM regression training with. SMO*” Mach. Learn., vol. 48, pp. 271–290, 2002.
- [12] Friess, T., Cristianini, N., & Campbell, C.; “*The Kernel-adatron: A fast and simple learning procedure for support vector machines*”. In J. Shavlik (Ed.), Proceedings of the Fifteenth International Conference on Machine Learning, pp. 188–196, 1998.
- [13] Joachims, T.; “*Making large-scale support vector machine learning practical*”, In B. Schölkopf, C. Burges, & A. Smola (Eds.), Advances in kernel methods—Support vector learning (pp. 169–184). MIT Press, 1999.

---

<sup>1</sup> Quadratic problem

<sup>2</sup> Classification

<sup>3</sup> MultiLayer Perceptron (MLP)

<sup>4</sup> Radial Basis Functions (RBF)

<sup>5</sup> Sequential Minimum Optimization (MLP)

<sup>6</sup> Numeric

<sup>7</sup> Nominal