همایش ملی انرژی
دانشگاه آزاد اسلامی واحد خمینی شهر
آذر ۱۳۹۴

مدل‌سازی میزان فروش انرژی برق بر اساس روش‌های شبکه‌های عصبی مصنوعی و رگرسیون و ارزیابی نتایج

علی ایمان [۱،*]، مریم شریف‌دوست [۲]

* نویسنده مسئول: ali.iman @iaukhsh.ac.ir

چکیده

امروزه با پیشرفت فن‌آوری و رشد جوامع بشری، درخواست برای انرژی، به ویژه انرژی الکتریکی، از سوی مصرف‌کنندگان رو به افزایش گذارده است. این نیاز از جمله نیازهای ضروری به‌شمار می‌آید که پاسخ‌گویی بدان مستلزم ساخت تاسیسات بزرگی مانند سدها، نیروگاه‌ها، آموزش متخصصین و... می‌باشدکه تامین هر یک از موارد ذکر شده بسیار زمان‌بر خواهد بود. درصورتی که قادر به پیش‌بینی دقیقی برای میزان تقاضای این انرژی در سال‌های آینده مقدور باشد، زیرساخت‌های لازم شامل آموزش، ساخت و نصب تجهیزات و... برای رسیدن به این سطح از تولید برای سال‌های آینده فراهم می‌شود. از این رو هر چه پیش‌بینی انجام شده برای تقاضای انرژی در سال‌های آینده دقیق‌تر باشد، منجر به برنامه‌ریزی دقیق‌تر در به‌کارگیری بهتر در زمان و سرمایه می‌شود. برای پیش‌بینی متغیرها، روش‌های مختلفی در دسترس می‌باشد که روش‌های آماری و روش‌های داده‌کاوی بیشتر مورد توجه هستند. در این مقاله از روش‌های آماری موجود، مدل رگرسیون و از روش‌های داده‌کاوی، شبکه‌های عصبی مصنوعی مورد استفاده قرار گرفته‌اند. در این تحقیق سعی بر آن شده است که با دانستن تعداد مشترکین در بخش‌های خانگی، صنعتی، کشاورزی و عمومی، در استان‌های تهران، اصفهان، فارس در یازده سال، به پیش‌بینی فروش انرژی برق بر حسب میلیون کیلو وات ساعت پرداخته شود. به منظور انجام مقایسه بین دو روش یاد شده از معیار MSE استفاده شده است.

واژه‌های کلیدی

داده‌کاوی، رگرسیون، شبکه‌های عصبی مصنوعی، انرژی، میانگین توان دوم خطا

۱-دانشجوی ارشد آمار اجتماعی و اقتصادی، دانشگاه آزاد اسلامی واحد خمینی‌شهر.

۲- استادیار، گروه ریاضی و آمار، دانشگاه آزاد اسلامی واحد خمینی‌شهر.

۱- مقدمه

امروزه با وجود تمامی پیشرفت‌هایی علمی حاصل شده، هم‌چنان ذخیره‌سازی انرژی الکتریکی در مقیاس وسیع مقدور نمی‌باشد، بنابراین مدیریت حاکم بر تولید و عرضه انرژی الکتریکی باید بر مبنای تطبیق عرضه تقاضای انرژی برق، اقدام به ارائه برنامه، سرمایه‌گذاری و بهره‌برداری، بهینه نماید. به‌ویژه با توجه به این‌که فرایند سرمایه‌گذاری در این صنعت یک عملیات زمان‌بر می‌باشد، مهم‌ترین گام در برنامه‌ریزی‌ها انجام شده، در اختیار داشتن اطلاعات کافی و کامل از چگونگی مصرف انرژی الکتریکی و پیش‌بینی روند تکاملی آن با توجه به عوامل مختلف در مصرف انرژی الکتریکی است؛ زیرا هر نوع برنامه‌ریزی در این مورد مشروط به داشتن اطلاعات جامع مربوط به میزان مصرف انرژی در مقاطع مختلف زمانی و مکانی است و یا به عبارت دیگر آگاهی از نیازهای آینده صنعت برق می‌تواند برنامه‌ریزان را در تعیین قدرت تولیدی، نوع نیروگاه‌ها، مشخصات سیستم‌های انتقال و توزیع و پست‌هایی که باید ساخته شوند یاری می‌رساند. این آگاهی بر اساس تحلیل اطلاعات گذشته، بررسی روند رشد بار و یا فرض قوانین تجربی یا مدل ریاضی انجام می‌گیرد.

روش‌های گوناگونی برای براورد و تخمین بارالکتریکی وجود دارند که اکثرا برای حالت بلند مدت و میان مدت طرح‌ریزی شده‌اند. برخی از این روش‌ها برپایه روش‌های آماری و برخی دیگر از جمله روش‌های داده‌کاوی[1] می‌باشند. از جمله‌ی این روش‌ها می‌توان به رگرسیون[2]، سری‌های‌زمانی[3]، شبکه‌های عصبی مصنوعی[4]، الگوریتم ژنتیک[5]، الگوریتم کلونی مورچگان[6] و... اشاره کرد. روش‌های دیگری که در این زمینه کاربرد دارند، مدل‌های تلفیقی هستند که از ترکیب دو یا چند روش در براورد بار استفاده می‌کنند.

در این مقاله سعی بر آن شده است که با دانستن تعداد مشترکین در سال‌های (۱۳۸۳-۱۳۹۳) به طور جداگانه، در بخش‌های خانگی، صنعتی، کشاورزی و عمومی، در استان‌های تهران، اصفهان و فارس ، به پیش‌بینی فروش انرژی برق که رابطه‌ی تنگاتنگی با میزان بار مصرفی دارد، پرداخته شود.

این پیش‌بینی با استفاده از دو روش رگرسیون و شبکه‌های عصبی مصنوعی که از جمله معروف‌ترین روش‌های آماری و داده‌کاوی می‌باشد، صورت گرفته و در آخر به مقایسه‌ی دو روش یاد شده در زمینه‌ی پیش‌بینی متغیر پاسخ پرداخته شده است.

در سال‌های اخیر، شاهد حرکتی گسترده تحقیقات صرفا تئوری به تحقیقات کاربردی برای حل مسائلی هستیم که برای آن‌ها مدل خاص ریاضی موجود نبوده و یا مدل ریاضی آنها پیچیده و کاملا غیرخطی است، به طوری که به‌دست آوردن آن مدل از روش‌های کلاسیک ریاضی بسیار سخت و همراه با خطای زیادی می‌باشد (البته منظور از خطا، خطای ناشی از گرفتن فرضیات خاص برای ساده شدن سیستم می‌باشد). با عنایت به این امر، علاقه فزاینده‌ای در توسعه تئوریک سیستم‌های دینامیکی هوشمند مدل–آزاد که مبتنی بر داده‌های تجربی هستند، ایجاد شده است. شبکه‌های عصبی مصنوعی یا ANN[7] جزء این دسته از سیستم‌های دینامیکی قرار دارند که با پردازش‌های لازم روی داده‌های تجربی، قوانین نهفته در بین داده‌ها را پیدا می‌نمایند.

در زمینه‌ی پیش‌بینی بار الکتریکی و میزان فروش انرژی برق نیز به‌نظر می‌رسدکه هر مدل ریاضی که برای بررسی بار یک منطقه یا یک کشور در نظر گرفته شود، نمی‌تواند یک مدل کاملا دقیق باشد که بتواند تمام پارامترهای موثر در بار آن منطقه را در خود نشان دهد. در ضمن اگر برای دقت بیشتر پارامترهای موثر در بار زیادتر شود، سیستم غیر خطی‌تر شده و به‌دست آوردن مدل ریاضی آن مشکل‌تر و هزینه‌بر می‌شود. لذا ایده‌ی به‌کارگیری شبکه‌های عصبی

[1] Data Minning
[2] Regression
[3] Times series
[4] Artificial Neural Network
[5] Genetic Algorithm
[6] Ant Colony Algorithm
[7] Artificial Neural Network

مصنوعی به عنوان یکی از مدل‌های استفاده شده در این مقاله به‌نظر منطقی می‌رسد.

تحلیل رگرسیون نیز که از آن به عنوان یکی از روش‌های آماری نام برده شد، برای تحلیل داده‌های چند عاملی به‌کار گرفته می‌شود که حوزه کاربرد آن بیش‌ترین وسعت را دارد نتایج پرجاذبه آن از نظر مفهومی، فرایند ساده به کارگیری یک معادله است که ارتباط بین مجموعه‌ای از متغیرها را بیان می‌کند. تحلیل رگرسیون هم‌چنین از جهت نظری به لحاظ ظرافت و زیبایی ذاتی ریاضیات، دارای جذابیت می‌باشد. موفقیت در تحلیل رگرسیونی نیاز به درک و تیزبینی در دو مقوله تئوری و مسائل عملی دارد و زمانی بروز می‌کند و خود را نشان می‌دهد که تکنیک در استخدام داده‌ها و اطلاعات عمومی قرار گیرد.

در ادامه، مقاله به صورت زیر پیش خواهد رفت: در بخش۲ به طور اجمالی به معرفی مدل رگرسیون و در بخش ۳ به معرفی شبکه‌های عصبی مصنوعی پرداخته شده است. نتایج حاصل از محاسبات و برازش مدل‌ها در بخش ۴ ارائه شده است و سرانجام در بخش ۵ نتیجه‌گیری‌های حاصل از تحقیق بیان گردیده است.

۲- رگرسیون

رگرسیون در لغت، به معنی بازگشت بوده و در اصطلاح کاربردی، به معنای پی‌بردن به رفتار یک متغیر به کمک رفتار سایر متغیرها می‌باشد. در آمار، رگرسیون به معنای یک نوع رابطه ریاضی می‌باشد که بین متغیر وابسته از یک سو و متغیرهای مستقل از سوی دیگر برقرار می‌باشد. دریافت روابط بین عوامل تاثیرگذار و تاثیرپذیر، حصول اطمینان از وجود همبستگی معنی‌دار بین دو متغیر و هم‌چنین هدف‌گذاری برای براورد یک متغیر بر حسب متغیر دیگر است. در حقیقت تحلیل رگرسیونی فن و تکنیکی آماری برای بررسی و به مدل در آوردن ارتباط بین متغیرها می‌باشد؛ هرچند مدل رگرسیون یک ارتباط علت و معلولی بین متغیرها را نتیجه نمی‌دهد. از جمله کاربردهای دیگر رگرسیون می‌توان به توصیف داده‌ها، براورد پارامترها، پیش‌گویی متغیر پاسخ و کنترل اشاره کرد[۱].

به مرور زمان محققین با مواجه شدن با انواع مختلف مسائل کاربردی، بر آن شدند تا به بسط و گسترش این تکنیک بپردازند. نتایج سال‌ها پژوهش و تحقیق این محققین منجر به پدید آمدن مدل‌های مختلفی برای این شیوه‌ی آماری گردید که از جمله‌ی آنها می‌توان به رگرسیون خطی‌ساده[۱]، رگرسیون خطی چندگانه[۲]، لجستیک[۳]، پروبیت[۴]، ترتیبی[۵]، رگرسیون غیرخطی[۶]، کاکس[۷] و... اشاره کرد.

یکی از مهم‌ترین شرایط ایده‌آل در بررسی‌های رگرسیونی عدم وجود ارتباط خطی بین متغیرهای مستقل است. اگر هیچ رابطه‌ی خطی بین متغیرهای مستقل رگرسیونی وجود نداشته باشد، گفته می‌شود که متغیرها متعامد هستند. در این صورت، استنباط‌هایی مانند مشخص کردن اثرات نسبی متغیرهای رگرسیونی، پیش‌بینی و یا براورد و انتخاب یک مجموعه مناسب از متغیرها برای برازش مدل، نسبتا به‌سادگی صورت خواهد پذیرفت. اما متاسفانه در بیش‌تر کاربردهای رگرسیونی، متغیرهای رگرسیونی متعامد نیستند. گاهی اوقات متعامد نبودن مسئله‌ی مهمی نمی‌باشد، در حالی که در بعضی وضعیت‌ها متغیرهای رگرسیونی تقریبا به طور کامل ارتباط خطی دارند و در چنین مواردی استنباط‌های بر اساس این مدل رگرسیونی می‌تواند به‌شدت گمراه کننده و غلط باشد. هنگامی که ارتباط نزدیکی بین متغیرهای رگرسیونی وجود دارد، گفته می‌شود مسئله‌ی هم‌خطی‌چندگانه رخ داده است. روش‌های متعددی برای برخورد با مشکلات ناشی از هم‌خطی‌چندگانه پیشنهاد شده است. از جمله‌ی این روش‌ها می‌توان به جمع‌آوری داده‌های اضافی، تخصیص مدل مجدد و

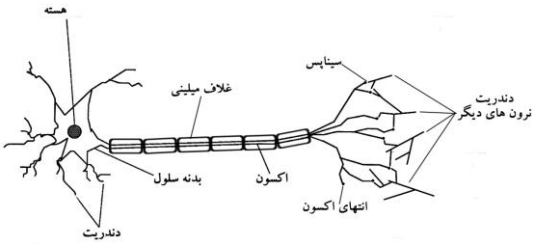
[1] Simple Linear Regression
[2] Multiple Linear Regression
[3] Logistic Regression
[4] Probit Regression
[5] Ordinal Regression
[6] Nonlinear Regression
[7] Cox Regression

به‌کارگیری روش‌های براوردی غیر از روش کمینه توان‌های دوم، اشاره کرد[۱].

۳- شبکه‌های عصبی مصنوعی

مغز یک سیستم پردازش اطلاعات با ساختار کاملا پیچیده است که در حالی‌که تنها دو درصد از وزن بدن را تشکیل می‌دهد، با این وجود بیش از بیست درصد از کل اکسیژن بدن را مصرف می‌کند. نرون‌ها ساده‌ترین واحد ساختاری سیستم‌های عصبی مغز هستند که وظیفه‌ی تحلیل و پردازش اطلاعات را به عهده دارند. از گردهمایی این نرون‌ها بافت‌هایی که عصب نامیده می‌شوند، ایجاد می‌شوند. بافت‌های عصبی، اطلاعات و پیام‌ها را از یک قسمت بدن به قسمت دیگر منتقل می‌کنند.

مغز به عنوان یک سیستم پردازش اطلاعات با ساختار موازی، از 10^{11} نرون به هم مرتبط با تعداد 10^{16} ارتباط تشکیل شده است . بیشتر نرون‌ها شامل اجزای زیر می‌باشند: هسته[۱]، بدنه سلول[۲]، دندریت[۳]، اکسون[۴] و سیناپس[۵]. در شکل (۱) شمایی از یک سلول عصبی بیولوژیک ترسیم شده است.

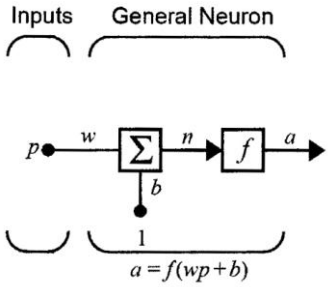


شکل(۱): شمایی از یک سلول عصبی بیولوژیک.

یکی از خصوصیات مهم شبکه‌های عصبی بیولوژیکی، توان تحمل ضایعه در شبکه عصبی است. انسان‌ها با شبکه عصبی مشتمل بر یکصد میلیارد نرون عصبی متولد می‌گردند و اکثر این نرون‌ها پس از مرگ‌شان جایگزین نمی‌گردند. اما علی‌رغم از دست دادن متوالی نرون‌ها، ما به طور دائم آموزش می‌بینیم و سایر نرون‌ها می‌توانند در بسیاری از موارد آموزش دیده و به جای نرون‌های از بین رفته به فعالیت بپردازند.

قابل ذکر است علیرغم این‌که نرون‌های بیولوژیکی در مقایسه با مدارهای الکتریکی بسیار کند عمل می‌کنند(10^{-3} در مقابل 10^{-9} ثانیه)، مغز قادر است بسیاری از کارها را بسیار سریع‌تر از رایانه‌ها انجام دهد. این امر به علت وجود ساختار موازی در شبکه‌های عصبی بیولوژیکی است[۲].

شبکه‌های عصبی مصنوعی، الگوهایی برای پردازش اطلاعات هستند که با الهام از شبکه‌ی عصبی مغز انسان ساخته شده‌اند و جزء سیستم‌هایی قرار دارند که با پردازش داده‌های تجربی، دانش یا قانون نهفته در ورای داده‌ها را به ساختار شبکه منتقل می‌کنند و به همین جهت به این گونه سیستم‌ها، هوشمند گفته می‌شود. زیرا بر اساس محاسبات بر روی داده‌های عددی و یا مثال‌ها، قوانین کلی را فرا می‌گیرند. شبکه‌های عصبی مصنوعی، با وجود این‌که با سیستم‌های عصبی طبیعی قابل مقایسه نیستند، ولی دارای ویژگی‌هایی هستندکه در هر کجا که نیاز به یادگیری یک نگاشت خطی یا غیرخطی باشد، توانمند می‌گرداند. برخی از این ویژگی‌هایی عبارت‌اند از: قابلیت یادگیری، قابلیت تعمیم، سرعت پاسخ‌گویی بالا حین کاربرد، پراکندگی اطلاعات و مقاوم بودن (تحمل پذیری خطا). در شکل (۲) نمایی از یک نرون تک ورودی نشان داده شده است[۳].



شکل(۲): نمایی از یک نرون تک ورودی.

[1] Nucleus
[2] Cell Body
[3] Dendrite
[4] Axon
[5] Synapse

شبکه‌های عصبی را می‌توان بر اساس نحوه یادگیری آن‌ها به سه دسته تقسیم‌بندی کرد: یادگیری با سرپرستی[1]، یادگیری تقویتی[2] و یادگیری بدون سرپرستی[3].

در این مقاله روش یادگیری به‌کار گرفته شده از نوع با سرپرستی می‌باشد. در روش یادگیری با سرپرستی به منظور دست‌یابی به یک آموزش دقیق باید الگوهای ورودی و خروجی مطلوب متناظر آن‌ها به عنوان زوج آموزشی[4] (x,Y) در اختیار شبکه قرار گیرد و سپس به شبکه اجازه داده شود که بر روی اطلاعات موجود پردازش لازم را به عمل آورده، خروجی خود را عرضه نماید. خروجی شبکه($\hat{Y}$) با خروجی مطلوب(Y) مقایسه شده و سپس با در نظر گرفتن خطای یادگیری، از آن‌ها جهت تنظیم پارامترهای شبکه استفاده شود.

در این‌گونه یادگیری‌ها، مسئله آموزش شبکه عصبی منجر به حل یک مسئله بهینه‌سازی می‌شودکه در نهایت باید معیاری از خطا را کمینه گرداند. قانون یادگیری با سرپرستی که در این مقاله مورد توجه قرار گرفته، الگوریتم پس انتشار خطا است.

در نهایت، پارامترهای شبکه (وزن‌ها و بایاس‌ها) بایستی به نحوی تنظیم شوند که عملکرد شبکه بهینه شود و منظور از بهینه شدن عملکرد شبکه، کمینه شدن خطایی است که بین خروجی شبکه($\hat{Y}$) و خروجی مطلوب(Y)وجود دارد. برای این‌که بهینه‌سازی انجام شود، باید دو مرحله کار صورت گیرد. اولین مرحله تعریف مفهوم عملکرد به صورت کمی می‌باشد که این عدد شاخص عملکرد[5] نامیده می‌شود و هر چه مقدار آن کمتر باشد، عملکرد شبکه بهتر است. در بسیاری از موارد از میانگین توان دوم خطاها[6] به عنوان شاخص عملکرد استفاده می‌شود. دومین مرحله، تنظیم پارامترهای شبکه برای کاهش مقدار شاخص عملکرد است. برای این‌کار باید از الگوریتم‌های بهینه‌سازی که نقاط کمینه را در سطح عملکرد می‌یابند، استفاده کرد.

الگوریتم‌های بهینه‌سازی، اکثرا تکرارشونده[7] هستند. یعنی یک مقدار اولیه x_0 حدس زده می‌شود و این مقدار طبق معادله‌ی (۱) در هر تکرار به‌روز[8] می‌شود:

$$x_{k+1} = x_k + \alpha_k p_k, \qquad (۱)$$

که در این رابطه بردار $\underset{\sim}{p}_k$ ، مسیر جستجو[9] و اسکالر α_k، نرخ یادگیری[10] می‌باشند. نرخ یادگیری طول قدم‌ها، زیاد یا کم بودن تغییرات در پارامترهای شبکه در هر تعداد، را معین می‌کند. الگوریتم‌های مختلف در نحوه انتخاب مسیر جستجو و نرخ یادگیری تفاوت دارندکه از جمله‌ی آن‌ها روش بیشترین شیب کاهشی[11] ، روش نیوتن[12] ، روش گرادیان مزدوج[13] و Rprop[14] می‌باشند. در این مقاله از روش Rprop استفاده شده است[۲].

٤- برازش

برای پیش‌بینی میزان فروش انرژی برق بر اساس متغیرهای مستقل که در ادامه به طور مفصل بیان شده‌اند، از دو روش رگرسیون و شبکه‌های عصبی مصنوعی استفاده گردیده است. داده‌های مورد استفاده در این مقاله از سایت مرکز آمار ایران و از نشریه‌ی، "آمار تفصیلی صنعت برق ایران - روند ۱۰ ساله صنعت برق"، جمع آوری گردیده‌اند. در ابتدا مجموعه‌ی کل داده‌ها را به صورت تصادفی به دو بخش (۷۵٪ و ۲۵٪) تقسیم‌بندی نموده؛ از بخش اول (آموزش) برای برازش مدل رگرسیونی به داده‌ها و آموزش شبکه‌های عصبی مصنوعی، استفاده شده است و از بخش دوم (آزمون) که شامل ۲۵٪ کل داده‌ها می‌شود، برای انجام آزمون و مقایسه عملکرد بین دو روش، استفاده شده است.

[1] Supervised Learning
[2] Reinforcement Learning
[3] Unsupervised Learning
[4] Training Pair
[5] Performance Index
[6] Mean of Square Errors (MSE)
[7] Iterative
[8] Update
[9] Search Direction
[10] Learning Rate
[11] Steepest Descent Method
[12] Newton's Method
[13] Conjugate Gradient Method
[14] Resilint backpropagation

متغیرهای مستقل به‌کار گرفته شده در تحلیل‌ها به قرار زیر می‌باشند: X_1: تعداد کل مشترکین خانگی، X_2: تعداد کل مشترکین کشاورزی، X_3: تعداد کل مشترکین صنعتی، X_4: تعداد کل مشترکین عمومی (تعداد مشترکین برای تمام متغیرها بر حسب هزار نفر می‌باشد)، X_5: استان تحت بررسی که شامل استان‌های تهران، اصفهان و فارس می‌باشد. چهار متغیر X_1، X_2، X_3 و X_4 از نوع پیوسته و متغیر X_5 از نوع طبقه‌بندی شده با سه سطح در نظر گرفته شده است که برای انجام تحلیل بر اساس مباحث ویژه رگرسیون به دو متغیر نشانگر C_1 و C_2 تبدیل می‌شود. نحوه‌ی کدبندی متغیرهای نشانگر برای هر سطح از متغیر طبقه‌بندی در جدول (۱) نشان داده شده است.

جدول(۱): نحوه کد بندی متغیرهای نشانگر.

	C_2	C_1
استان تهران	۰	۰
استان اصفهان	۰	۱
استان فارس	۱	۰

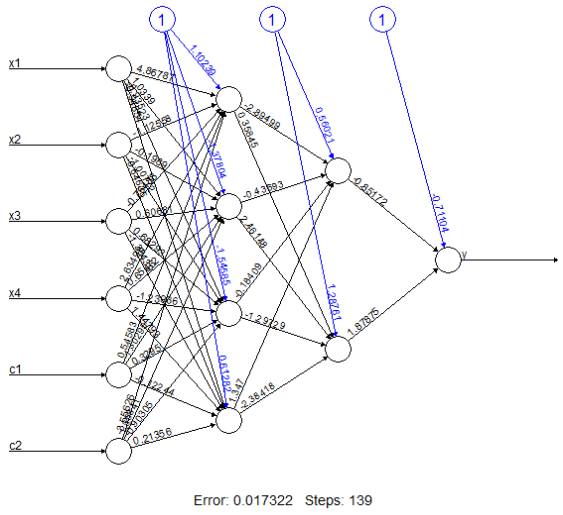
در پایان متغیر پاسخ Y ، میزان فروش انرژی برق بر حسب میلیون کیلو وات ساعت در نظر گرفته شده است. داده‌های تحت بررسی از سال ۱۳۸۳ تا سال ۱۳۹۳ می‌باشند. در شکل(۳) بخشی از داده‌ها نشان داده شده است.

Data Display

Row	year	x1	x2	x3	x4	c1	c2	y
1	1393	4787	10	38	464	0	0	31423
2	1392	4577	10	34	435	0	0	28749
3	1391	3994	8	31	384	0	0	25504
4	1390	3820	8	29	358	0	0	24690
				⋮				
15	1390	1620	34	23	58	1	0	18706
16	1389	1530	30	23	54	1	0	18384
17	1388	1454	26	22	51	1	0	17162
18	1387	1383	20	25	47	1	0	16923
				⋮				
30	1386	993	26	9	37	0	1	8727
31	1385	952	25	9	31	0	1	7880
32	1384	915	25	8	27	0	1	7357
33	1383	878	24	7	25	0	1	6707

شکل(۳) بخشی از داده‌ها .

برای انجام تحلیل‌ها از نرم افزار R استفاده شده است. روش شبکه عصبی مصنوعی مورد تحلیل در برنامه، با استفاده از پکیج neuralnet و با دو لایه پنهان که در لایه اول تعداد ۴ نرون و در لایه دوم تعداد ۲ نرون در نظر گرفته شده، صورت گرفته است. در شکل(۴) نمایی از مدل شبکه‌ی عصبی به‌کار رفته در برنامه ترسیم شده است.



شکل(٤): شبکه عصبی مصنوعی با دو لایه پنهان.

هم‌چنین انجام رگرسیون بر اساس روش کمترین توان‌های دوم خطا می‌باشد. هدف از مدل رگرسیون خطی چندگانه، برآورد ضرایب ($\beta_i \quad i=1,2,...,6$) در معادله(۲) می‌باشد.

$$Y=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3 \qquad (۲)$$
$$+\beta_4X_4+\beta_5C_1+\beta_6C_2,$$

در نهایت مدل برازش شده‌ی نهایی توسط رگرسیون به‌صورت معادله‌ی(۳) به‌دست آمده است.

$$Y=2186.8+7.12X_1+72.46X_2+186.52X_3 \qquad (۳)$$
$$-29.09X_4-130.19C_1-3746.68C_2,$$

از جمله مزایای استفاده از روش رگرسیون نسبت به شبکه‌های عصبی مصنوعی این است که، مدل‌های رگرسیونی قادر به ارائه‌ی یک فرمول برازش شده به داده‌ها می‌باشند؛ در صورتی که روش شبکه‌های عصبی مصنوعی قادر به برازش چنین فرمول نهایی نمی‌باشد.

در جدول (۲) مقادیر براورد شده برای متغیر Y، با استفاده از دو روش رگرسیون و شبکه‌ی عصبی مربوط به داده‌های ۲۵٪آزمون، نشان داده شده است.

جدول(۲) مقادیر برآورد شده برای متغیر Y.

	سطرانتخابی	سال	Y	$\hat{Y}.reg$	$\hat{Y}.ann$
۱	۸	۱۳۸۶	۲۹۱۵۶	۲۸۰۹۹	۲۷۷۹۲
۲	۱۱	۱۳۸۳	۲۵۶۱۸	۲۵۳۳۸	۲۴۹۰۷
۳	۱۲	۱۳۹۳	۲۱۳۴۳	۲۰۸۳۵	۲۰۸۸۹
۴	۱۳	۱۳۹۲	۲۰۳۵۰	۲۰۲۲۲	۲۰۳۲۰
۵	۱۹	۱۳۸۶	۱۵۶۸۱	۱۵۹۹۶	۱۶۱۴۵
۶	۲۳	۱۳۹۳	۱۲۴۴۵	۱۲۵۵۵	۱۲۴۰۲
۷	۲۶	۱۳۹۰	۱۰۵۳۱	۱۰۵۱۴	۱۰۲۲۸
۸	۲۹	۱۳۸۷	۹۶۷۳	۸۶۸۰	۸۹۰۱

با توجه به مقادیر متغیرهای $\hat{Y}.reg$ و $\hat{Y}.ann$ به نظر می‌رسد که رگرسیون در پیش‌بینی متغیر Y از دقت بیشتری برخوردار است. برای اطمینان بیشتر از درستی این ادعا، از آماره MSE به صورت معادله(۳)، به منظور مقایسه دقت پیش‌بینی دو روش یاد شده استفاده شده است. نتایج حاصل از این آماره در جدول (۳) بیان شده است.

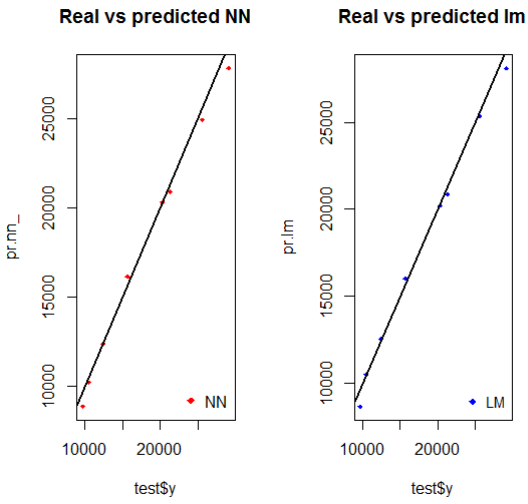
$$MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-k}, \qquad (۳)$$

جدول(۳): نتایج حاصل از آماره MSE .

	Regression	ANN
MSE	320776/4	434185/7

مقدار آماره‌ی MSE مربوط به رگرسیون با توجه به جدول(۳) از مقدار مشابه آن برای روش شبکه‌های عصبی مصنوعی کمتر بوده و درنتیجه می‌توان بیان نمود که مقادیر پیش‌بینی شده برای متغیر Y از دقت بیشتری نسبت به روش شبکه‌های عصبی مصنوعی برخوردار است.

برای درک بهتر نتایج، نمودارهایی رسم شده است که مقادیر پیش‌بینی شده متغیر Y $(\hat{Y})$ را در برابر مقادیر حقیقی آن ترسیم می‌کنند. شکل (۵) نمودار حاصل از برازش مقادیر اصلی Y در برابر مقادیر پیش‌بینی شده بر حسب روش رگرسیون و شبکه‌های عصبی نشان می‌دهد.



شکل(۵): برازش مقادیر Y در برابر مقادیر پیش‌بینی شده (LM : رگرسیون، NN : شبکه‌ی عصبی مصنوعی)

با توجه به شکل(۵)، مجموعه نقاط مربوط به مدل رگرسیون بیش از مجموعه نقاط شبکه‌ی عصبی مصنوعی به خط برازش شده، نزدیک‌تر بوده است که این مطلب تاییدی بر صحت نتایج بیان شده در بالا ، دال بر پیش‌بینی بهتر متغیر وابسته توسط روش رگرسیون می‌باشد.

نتایج به‌دست آمده با نتایج حاصل از شبیه‌سازی که بر اساس همین تعداد متغیر با حجم نمونه مشخص شده است، تطبیق دارد، این مطلب بیانگر آن است که در حجم‌های کم نمونه‌ای با متغیرهای مستقل که از انواع پیوسته و طبقه‌بندی

شده می‌باشد، استفاده از رگرسیون نسبت به شبکه‌های عصبی مصنوعی از ارجحیت برخوردار است.

۵-نتایج

بر اساس نتایج حاصل از آماره‌ی MSE موجود در جدول(۳) به دلیل کمتر بودن مقدار این آماره برای روش رگرسیون نسبت به روش شبکه‌های عصبی مصنوعی، می‌توان اذعان نمود که استفاده از روش رگرسیون نسبت به روش یاد شده در برازش مدل پیش‌بینی فروش انرژی برق، از ارجحیت برخوردار می‌باشد. از دیگر مزایایی که می‌توان برای روش رگرسیون نام برد، وجود معادله‌ی برازش یافته به داده‌ها برای فهم بهتر و سرعت بالای پردازش اطلاعات، می‌باشد.

٦-منابع

[۱]رضوی پاریزی، ابراهیم، مقدمه‌ای بر تحلیل رگرسیون خطی، انتشارات دانشگاه شهید باهنر کرمان، ۱۳۹۰

[۲]فروزان، محمدرضا، روش‌های نوین بهینه‌سازی، جهاددانشگاهی واحد صنعتی اصفهان، ۱۳۸۸

[۳]منهاج، م.ب، مبانی شبکه‌های عصبی، مرکز نشر دانشگاه صنعتی، ۱۳۸۱