

مدل طراحی بهینه معماری برای شبکه‌های عصبی مصنوعی و به‌کارگیری آن در پیش‌بینی مصرف ماهانه نفت‌گاز کل کشور

محمدرضا امین‌ناصری^{۱*}، احمد کوچک‌زاده^۲

۱- دانشیار مهندسی صنایع، دانشگاه تربیت مدرس، تهران، ایران
۲- دانشجوی دکتری تخصصی، مهندسی صنایع، دانشگاه تربیت مدرس، تهران، ایران

پذیرش: ۸۶/۱۲/۷

دریافت: ۸۶/۲/۲۵

چکیده

یکی از گام‌های مهم در توسعه شبکه‌های عصبی مصنوعی طراحی معماری شبکه است که تأثیر زیادی بر عملکرد شبکه دارد. در طراحی معماری شبکه‌های عصبی مصنوعی، عواملی از قبیل تعداد لایه‌های پنهان، تعداد نرون‌ها در هر لایه، توابع تبدیل و الگوریتم آموزش باید تعیین شوند. محققان در طراحی معماری شبکه به‌طور عمده از طریق سعی و خطا عمل می‌کنند و یا اینکه اثر متقابل بین عوامل مختلف در طراحی معماری شبکه را در نظر نمی‌گیرند. در این تحقیق، یک مدل مبتنی بر تکنیک طراحی آزمایشها برای طراحی بهینه معماری شبکه عصبی با یادگیری تحت نظارت، با توجه به اثر متقابل بین عوامل ذکر شده، ارائه می‌شود.

در این تحقیق، از مدل پیشنهادی برای طراحی معماری شبکه عصبی به منظور پیش‌بینی مصرف نفت‌گاز کل کشور استفاده شده است. به‌منظور مقایسه عملکرد مدل پیشنهادی با استفاده از روش سعی و خطا به‌عنوان یکی از روشهای مرسوم در طراحی معماری، یک مدل شبکه عصبی برای پیش‌بینی مصرف نفت‌گاز توسعه داده شده و طی آن برتری مدل پیشنهادی نشان داده شده است. همچنین برای مقایسه عملکرد شبکه عصبی با روشهای آماری، دو مدل با استفاده از رگرسیون و ARIMA طراحی شده‌اند. نتایج به دست آمده در پیش‌بینی مصرف نفت‌گاز در این بخش نیز نشان می‌دهد که پیش‌بینی با شبکه عصبی طراحی شده جوابهای بهتری دارد.

کلید واژه‌ها: معماری شبکه‌های عصبی مصنوعی، طراحی آزمایشها، پیش‌بینی، مصرف نفت‌گاز.



۱- مقدمه

پس از توسعه الگوریتم پس انتشار خطا، برای آموزش شبکه عصبی پیش‌خوراند، شبکه عصبی مصنوعی به صورت فزاینده‌ای در انواع مسائل مهندسی و تجارت مورد توجه و استفاده قرار گرفت [۱:۲، ص ۹۳-۱۰۵]. یکی از عمده‌ترین موارد کاربرد شبکه عصبی مصنوعی، پیش‌بینی است. اگرچه در طی سالهای متمادی استفاده از مدل‌های آماری خطی در پیش‌بینی مرسوم بوده است، اما این مدلها علی‌رغم برخورداری از مزایایی از قبیل سادگی و تعدد نرم‌افزارهای تجاری، توانایی در نظرگرفتن روابط غیرخطی را -که اساس دنیای واقعی را تشکیل می‌دهند- ندارند [۲]. در واقع استفاده روزافزون از شبکه عصبی در پیش‌بینی از این واقعیت ناشی می‌شود که شبکه عصبی، یک مدل مبتنی بر داده با قابلیت الگوبرداری غیرخطی و عاری از بسیاری فرضیه‌های محدودکننده است. در این ارتباط، هورانگ، به مقایسه عملکرد شبکه عصبی و مدل‌های سری زمانی $ARMA(p,q)$ به صورت تجربی پرداخت و در آن نشان داد که عموماً شبکه‌های عصبی مصنوعی عملکرد بهتری نسبت به مدل $ARMA(p,q)$ دارند [۳:۴، صص ۲۷۳-۲۸۹]. Zhang و همکاران به بررسی تجربی عملکرد شبکه عصبی در پیش‌بینی سری‌های زمانی غیرخطی در مقایسه با مدل‌های خطی مرسوم پرداخته و نشان دادند که شبکه‌های عصبی مصنوعی عملکرد بسیار بهتری داشته‌اند [۵، صص ۳۶-۳۸۱]. Marquez با مقایسه رگرسیون، یکی دیگر از ابزارهای متداول در پیش‌بینی، نشان داد که شبکه‌های عصبی به لحاظ قابلیت انطباق بر داده‌ها، مانند مدل رگرسیون عمل می‌کنند [۶].

یکی از موضوعات تحقیقی مهم در ادبیات شبکه عصبی مصنوعی، طراحی معماری شبکه عصبی است. در ادبیات طراحی معماری شبکه عصبی، عموماً عواملی از قبیل تعداد لایه‌های پنهان، تعداد نرون‌ها در هر لایه و توابع انتقال مطرح می‌شوند. محققان در طراحی معماری شبکه عصبی، از رویکردهای متفاوتی استفاده کرده‌اند: ۱- رویکرد سعی و خطا که در آن برای تعیین معماری شبکه عصبی، ترکیبهای متعددی مورد آزمایش قرار می‌گیرد و با مقایسه عملکرد هر ترکیب با سایر ترکیبها، معماری شبکه تعیین می‌شود. Zhang و همکاران به بررسی تأثیر نرونهای لایه ورودی و نرونهای لایه مخفی در سری‌های زمانی غیرخطی پرداخته و تأثیر هر عامل را مستقل و بدون در نظرگرفتن تأثیر متقابل با عوامل دیگر، بررسی

1. Back propagation algorithm
2. Feed forward Neural network

کرده و در تعیین معماری شبکه، از طریق سعی و خطا عمل کرده‌اند [۵، صص ۳۶-۳۸۱]. روش سعی و خطا، معماری بهینه را تضمین نمی‌کند و به خبرگی فرد، بستگی دارد؛ ۲- تکنیکهای سازنده و یا تکنیک هرس^۱ که بر پایه اضافه و یا حذف یک نرون از معماری اولیه شبکه با توجه به تأثیر آن بر عملکرد شبکه، قرار دارد. در این تکنیک، در زمانی که میانگین مربعات خطای شبکه، از یک مقدار از پیش تعیین شده بزرگتر باشد، یک نرون به شبکه اضافه می‌شود و در حالتی که وزنهای شبکه با اضافه کردن نرون مزبور، در تعداد زیادی تکرار، تغییر نکند و یا خطای شبکه به میزان قابل توجهی بهبود نیابد، از شبکه حذف می‌شود. مشهورترین الگوریتمهای سازنده روش DNC^۲ و روش CC^۳ است که در روش DNC یک شبکه با یک لایه پنهان ساخته می‌شود و در روش CC یک شبکه با چندین لایه پنهان ساخته می‌شود [۷، صص ۳۶۵-۳۷۵؛ ۸]. روش DNC در یادگیری مسائل پیچیده با مشکل روبرو می‌شود [۸]. با افزایش تعداد لایه‌های پنهان و بروز پدیده انطباق بیش از حد بر داده‌ها، قابلیت تعمیم در شبکه حاصل از روش CC در برخی موارد تنزل می‌یابد [۹، صص ۵۱۱-۵۱۶]. Muras و Munrul Islam، روش CNNDA^۴ را برای توسعه یک شبکه با دولایه پنهان ارائه دادند [۹، صص ۵۱۱-۵۱۶]. در روش CNNDA، ابتدا نرون‌ها یکی پس از دیگری به یک شبکه با دولایه پنهان اضافه می‌شوند، آنگاه پس از همگرایی شبکه و دستیابی به نتیجه مطلوب (تا زمانی که جواب شبکه به میزان قابل توجهی تنزل پیدا نکرده) به هرس و حذف نرون‌ها و اتصالات شبکه ادامه می‌دهد [۱۰، صص ۱۲۶۵-۱۲۷۸]. از آنجا که این روشها مبتنی بر روش نزول گرادیان است، نمی‌توان جواب بهینه کلی^۵ را تضمین کرد و ممکن است پاسخ نهایی، یک جواب بهینه محلی باشد؛ ۳- الگوریتم ژنتیک، که در آن با تغییر تعداد لایه‌های پنهان و تعداد نرونهای هر لایه با استفاده از عملگرهای الگوریتم ژنتیک، فضای گزینه‌های مختلف معماری شبکه، براساس تابع هدف معماری شبکه، جستجو می‌شود و بهترین معماری انتخاب می‌شود. میلر و همکاران با رمزگذاری الگوی اتصالات شبکه از الگوریتم ژنتیک در طراحی معماری شبکه عصبی استفاده کردند [۱۱، صص ۳۷۹-۳۸۴]. Schaffer و همکاران نیز از الگوریتم ژنتیک براساس روش رمزگذاری مستقیم برای تعیین دامنه نرخ یادگیری در

-
1. Pruning algorithm
 2. Back propagation algorithm
 3. Feed forward Neural network
 4. Feed forward Neural network
 5. Global Optimal Solution



الگوریتم پس انتشار خطا، تعداد نرونهای لایه پنهان و دامنه وزنهای اولیه در شبکه پیش‌خوراند استفاده کردند [۱۲، صص ۲۴۴-۲۴۸]. Gencay و Arifovic از روش Schaffer برای تعیین دامنه وزنهای اولیه شبکه، ساختار اتصالات شبکه، تعداد نرونهای لایه پنهان و نوع ورودی‌ها در یک شبکه پیش‌خوراند با یک لایه پنهان استفاده کردند [۱۳، صص ۵۷۴-۵۹۴]. Vosniakos و Benardos برای تعیین تعداد لایه‌های پنهان و تعداد نرونهای لایه پنهان به‌صورتی که خطای حاصل از آموزش و همچنین قابلیت تعمیم حداقل شود، از الگوریتم ژنتیک استفاده کردند [۱۴]. استفاده از الگوریتم ژنتیک در طراحی معماری بسیار زمان بر است و پارامترهای الگوریتم ژنتیک، به‌وسیله کاربر تعیین می‌شود [۱۵، صص ۶۳۵-۶۴۵]. همچنین در الگوریتم ژنتیک، عوامل دیگر نظیر الگوریتم آموزش، تابع تبدیل و اثر متقابل بین عوامل مختلف در نظر گرفته نمی‌شود؛ ۴- روش تاگوچی^۱ که بر اساس آن با انجام آزمایشهای هدفمند از بین ترکیبهای متعدد، ترکیب بهینه انتخاب می‌شود. Khaw و همکاران از این روش در تعیین تعداد لایه‌های مخفی، تعداد نرونها در هر لایه و اندازه مجموعه آموزش در یک شبکه پیش‌خوراند، استفاده کردند [۱۶، صص ۲۲۵-۲۴۵]. Lin و Chen و Tseng از این روش برای طراحی معماری شبکه LVQ^۲ استفاده کردند [۱۷، صص ۲۳۶-۱۸:۲۴۱، صص ۳-۱۴]. Teo و Sim، از این روش در طراحی معماری شبکه نقشه خود سازمانده^۳ استفاده کردند [۱۹، صص ۸۵-۹۴]. این روش اثر متقابل بین عوامل مختلف را در نظر نمی‌گیرد و از پشتوانه آماری لازم برخوردار نیست؛ ۵- سایر تکنیک‌ها نظیر تجزیه کانونی^۴، معیار اطلاعاتی شبکه^۵، الگوریتم چندجمله‌ای زمانی^۶ [۲۰، صص ۲۹۱-۳۰۰؛ ۲۱، صص ۸۶۵-۲۲:۸۷۲، صص ۵۳۵-۵۴۵]. هیچ‌کدام از این روشها، جواب بهینه را در همه حالات تضمین نمی‌کنند [۲۳، صص ۳۵-۶۲].

روشهای ذکر شده، اولاً به‌طور عمده نظیر الگوریتم ژنتیک زمان بر و بسیار پیچیده هستند، ثانیاً در طراحی معماری شبکه عصبی، اثر متقابل بین عوامل طراحی را در نظر نمی‌گیرند. به همین دلیل نمی‌توان ساختار حاصل از آنها را بهینه دانست.

-
1. Taguchi method
 2. LVQ Net: Learning Vector Quantization Net
 3. SOM Net: Self Organizing Map Net
 4. Canonical Decomposition Technique
 5. Network Information Criterion
 6. Polynomial TimeAlgorithm

در این تحقیق به منظور در نظر گرفتن آثار متقابل بین عوامل مختلف در طراحی معماری شبکه عصبی، مدلی با استفاده از تکنیک طراحی آزمایشها^۱ برای شبکه‌های با یادگیری تحت نظارت^۲ ارائه شده است. طی استفاده از این تکنیک از عواملی نظیر نوع شبکه، تعداد لایه‌های پنهان، تعداد نرونهای هر لایه، توابع پیش‌پردازش و توابع انتقال استفاده می‌شود. در این تحقیق از مدل طراحی شده برای پیش‌بینی مصرف ماهانه نفت‌گاز به‌عنوان یک مورد استفاده شده است. در ادامه و در بخش دوم، مدل پیشنهادی تبیین می‌شود. در بخش سوم مدل طراحی شده برای پیش‌بینی مصرف ماهانه نفت‌گاز کل کشور اجرا می‌گردد. در بخش چهارم با استفاده از روش مرسوم در طراحی معماری، یک مدل شبکه عصبی برای پیش‌بینی مصرف نفت‌گاز توسعه داده شده و با شبکه عصبی حاصل از مدل پیشنهادی مورد مقایسه قرار می‌گیرد. همچنین برای مقایسه عملکرد مدل پیش‌بینی مبتنی بر شبکه عصبی با روشهای آماری، دو مدل با استفاده از رگرسیون و مدل^۳ ARIMA طراحی و مورد مقایسه قرار گرفته‌اند. در بخش پنجم هم به نتیجه‌گیری پرداخته می‌شود.

۲- مدل پیشنهادی در طراحی معماری

درحقیقت، طراحی معماری شبکه، یک مسأله بهینه‌سازی است که در آن بردار متغیرهای طراحی معماری از قبیل تعداد لایه‌های پنهان و تعداد نرون‌ها در هر لایه، متغیرهای تصمیم‌گیری و شاخصهای اندازه‌گیری خطای شبکه، تابع هدف است.

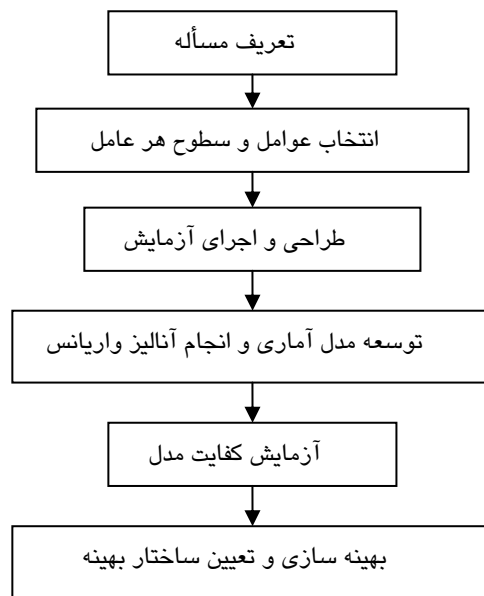
برای حل این‌چنین مسأله بهینه‌سازی از روشهای رایج نظیر برنامه‌ریزی خطی به سبب ماهیت متغیرهای طراحی و پیچیدگی تابع هدف نمی‌توان استفاده کرد. به همین دلیل در این تحقیق برای بهینه‌سازی مسأله طراحی معماری شبکه عصبی، از مدلی برپایه تکنیک طراحی آزمایشها استفاده می‌شود.

در مدل پیشنهادی، پس از تعریف مسأله موردنظر، عوامل و سطوح هر یک، در طراحی معماری تعیین می‌شود. آنگاه تعدادی آزمایش بر اساس همه ترکیبهای متفاوت از عوامل و سطوح انجام می‌شود. سپس با انجام آزمون آماری آنالیز واریانس، وجود و یا عدم وجود

1. DOE: Design of Experiment
2. Supervised Learning
3. Taguchi method



تأثیر هر یک از عوامل و اثر متقابل بین آنها تعیین می‌شود. پس از آزمایش کفایت مدل^۱، بر اساس نتایج آنالیز واریانس و مشخص شدن آثار معنادار^۲، یک رویه تحت عنوان رویه پاسخ^۳ بر داده‌ها انطباق داده می‌شود. رویه پاسخ، مبنای پیش‌بینی و محاسبه ترکیب بهینه از بین ترکیبهای متعدد بین عوامل منتخب و سطوح آنها است. در شکل ۱ مدل گام به گام پیشنهادی در طراحی معماری شبکه نشان داده شده است.



شکل ۱ فرایند طراحی آزمایشها

در این تحقیق از مدل پیشنهادی طراحی بهینه معماری شبکه عصبی برای پیش‌بینی مصرف ماهانه نفت‌گاز در کل کشور استفاده می‌شود. به سبب آنکه نفت‌گاز در بخشهای متنوع اقتصادی کشور استفاده می‌شود، تأمین نابهنگام آن تبعات نامطلوبی دارد. بنابراین باید

1. Model Adequacy
2. Significant Effects
3. Response Surface

به منظور برنامه ریزی در تامین فراورده مزبور، مصرف آن به صورت دقیق پیش بینی شود.

۲-۱- انتخاب عوامل و سطوح هر عامل

بر طبق شکل ۱ پس از تعریف مسأله و مشخص شدن نوع کاربرد شبکه عصبی، لازم است تا عوامل و سطوح هر یک از آنها در طراحی معماری شبکه عصبی تعیین شود. در این تحقیق، متغیرهای طراحی نوع شبکه، تعداد لایه های پنهان، تعداد نرونهای هر لایه، تابع تبدیل و تابع پیش پردازش در نظر گرفته شده اند.

۲-۱-۱- نوع شبکه

شبکه پیش خوراند، متداولترین شبکه ای است که در پیش بینی استفاده می شود [۲۴؛ صص ۳۰۳-۳۱۰؛ ۲۵]. علاوه بر شبکه های پیش خوراند، انواع دیگری از شبکه های عصبی مصنوعی نیز در پیش بینی استفاده شده است. برای مثال، شبکه عصبی بازگشتی^۱ در زمینه الگوهای غیرخطی پویا، جانشین مناسبی برای شبکه پیش خوراند در پیش بینی سری های زمانی است [۲۶؛ صص ۱۲۵-۱۳۳؛ ۲۷؛ صص ۳۶-۳۷؛ ۲۸؛ صص ۷۱-۸۳؛ ۲۹؛ صص ۲۳۷-۲۴۰؛ ۳۰؛ صص ۷۶۵-۷۸۶؛ ۳۱؛ صص ۲۲۳-۲۳۵؛ ۳۲؛ صص ۱۲۶-۱۳۲]. در شبکه عصبی بازگشتی، خروجیها می توانند به صورت مستقیم، به ورودیها بازخورد بدهند که نتیجه آن بازخورد پویا روی خطای حاصل از الگوهای قبلی است. براین اساس درست مانند مزیتی که مدل های ARMA بر AR در پیش بینی دارند، پویایی بیشتری نسبت به شبکه های پیش خوراند دارند. با وجود این، تحقیقات چندانی در بررسی مزیت نسبی شبکه های بازگشتی نسبت به شبکه های پیش خوراند در ادبیات پیش بینی صورت نگرفته است. در این تحقیق از دو نوع شبکه عصبی پیش خوراند و بازگشتی برای توسعه مدل پیش بینی استفاده شده است.

شایان ذکر است در تمام تحقیقات، نوع شبکه در طراحی معماری شبکه عصبی در نظر گرفته نمی شود و روش پیشنهادی منحصر به یک نوع شبکه است، اما در این تحقیق با توجه به انعطاف پذیری روش پیشنهادی می توان این عامل را در زمره متغیرهای طراحی در نظر گرفت.



۲-۱-۱- لایه پنهان

نرونهای لایه پنهان شبکه را در کشف الگوی نهفته در داده‌ها و روابط غیرخطی حاکم بر ورودیها و خروجیها، توانمند می‌سازند.

در ادبیات موضوع، برخی از محققان تنها از یک لایه پنهان استفاده می‌کنند [۲۳، صص ۳۵-۶۲] و برخی دیگر به استفاده از دو لایه پنهان توصیه می‌نمایند [۳۳، صص ۳۳-۳۴؛ ۳۵؛ ۳۴، صص ۲۲۷-۲۳۴]. Zhang نشان داده است که استفاده از دو لایه پنهان در مسائل پیش‌بینی دقت و کیفیت پیش‌بینی را ارتقا می‌دهد [۳۵، صص ۱۵۱-۱۷۰]. چنانچه در این تحقیق نشان داده می‌شود، استفاده از یک یا دو لایه پنهان به عوامل دیگری بستگی دارد و چنانچه Zhang بررسی کرده مستقل از سایر عوامل نمی‌تواند بررسی شود.

در این تحقیق از شبکه‌ای با دو لایه پنهان استفاده شده است. برای تعیین تعداد نرونهای لایه پنهان اول، بر مبنای قاعده Hecht و Nielsen از حداکثر $2n+1$ نرون که در آن n تعداد نرونهای لایه ورودی است، عمل شده است. [۳۶].

۲-۱-۲- پیش پردازش

پیش‌پردازش داده‌ها قبل از آموزش شبکه موجب تسهیل در یادگیری شبکه، فراهم آوردن ملزومات الگوریتم آموزش و اجتناب از پیچیدگیهای محاسباتی می‌شود [۳۷، صص ۴۴۲-۴۵۶؛ ۳۸، صص ۴۹۱-۴۹۴]. در ادبیات موضوع به تابع پیش‌پردازش خاصی اشاره نشده است. لذا در این تحقیق از دو تابع متداول پیش‌پردازش MinMax و انحراف معیار (StDev) بر اساس رابطه‌های (۱) بهره‌گیری شده است.

$$X_n = \frac{X - X_{Min}}{X_{Max} - X_{Min}} \quad (1)$$

$$X_n = \frac{X - X_{Mean}}{StdDev} \quad (2)$$

در تابع اول مقادیر ورودی در بازه [۱/۰] به‌طور یکنواخت بر گردانده می‌شود و در تابع دوم، میانگین داده‌ها صفر و انحراف معیار آنها برابر یک می‌شود.

۳-۱-۲- تابع تبدیل

تابع تبدیل، ارتباط بین ورودی و خروجی یک نرون و شبکه را تعیین می‌کند. تابع تبدیل انواع محدودی دارد، اما در پیش‌بینی در اکثر تحقیقات از یک نوع تابع تبدیل در هر لایه و عموماً از توابع لجستیک و هایپربولیک در لایه پنهان و تابع همانی و یا خطی در لایه خروجی استفاده می‌شود [۲۳، صص ۳۵-۶۲]. در این تحقیق از دو نوع تابع تبدیل لجستیک و هایپربولیک در لایه پنهان و از تابع خطی در لایه خروجی استفاده شده است. در جدول ۱، عوامل منتخب در طراحی معماری شبکه عصبی و سطوح هریک به‌طور خلاصه عنوان شده است.

جدول ۱ عوامل و سطوح عوامل مدل طراحی آزمایشها

عوامل	نوع شبکه (NT)		تعداد نرونهای لایه پنهان اول (N _۱)	تعداد نرونهای لایه پنهان دوم (N _۲)	تابع پیش‌پردازش (PF)		تابع تبدیل (TF)	
	سطوح	پیش خوراند	بازگشتی	۱-۱۵	۰،۱،۲	STDV	MinMax	LogSig

۲-۲- طراحی و اجرای آموزش

در این مرحله، ماتریس آزمایشات شامل تخصیص مناسب عوامل منتخب و سطوح هریک از این عوامل تعیین شده، آنگاه بر اساس ترکیبهای متفاوت در ماتریس آزمایشات، شبکه طراحی و آموزش داده می‌شود. در این تحقیق براساس جدول ۱، دو سطح برای عامل نوع شبکه، ۱۵ سطح برای تعداد نرونهای لایه پنهان اول، ۳ سطح برای تعداد نرونهای لایه پنهان دوم، ۲ سطح برای تابع پیش‌پردازش و ۲ سطح برای تابع تبدیل در نظر گرفته شده است و لذا ماتریس آزمایشات شامل $360 = (2 \times 15 \times 3 \times 2 \times 2)$ ترکیب می‌شود. براساس روش پیشنهادی باید با توجه به تک تک این ترکیبها، شبکه‌های عصبی طراحی شده و با تعدادی تکرار، آموزش داده شوند. در این تحقیق، شبکه عصبی به ازای هر ترکیب، پنج مرتبه آموزش داده



شده است. در شبکه عصبی از الگوریتمهای آموزش متفاوتی استفاده می شود که هر کدام مزایا و معایبی دارند. در این تحقیق، به توصیه Zhang، از الگوریتم لونیبرگ_ مارکوورت^۱ که از نظر سرعت و اشغال حافظه کمتر، مطلوبیت بیشتری دارد، استفاده شده است [۲۳، صص ۳۵-۶۲].

در گام اول در توسعه مدل آماری باید برای ارزیابی عملکرد، متغیر پاسخ^۲ تعیین شود. محققان از شاخصهای متفاوتی برای ارزیابی عملکرد شبکه عصبی استفاده کرده اند. از جمله این شاخصها می توان به مجموع مربعات خطا، میانگین مربعات خطا^۳، میانگین درصد خطا^۴، میانگین انحرافات مطلق^۵ اشاره کرد [۳۹، صص ۵۹۱-۵۹۷؛ ۴۰]. در این تحقیق از میانگین درصد خطا به عنوان شاخص ارزیابی عملکرد شبکه عصبی و همچنین متغیر پاسخ در طراحی آزمایشها استفاده شده است.

۲-۳- توسعه مدل آماری و انجام آنالیز واریانس

در این مرحله پس از آموزش شبکه های عصبی طراحی شده در مرحله قبل، مدل آماری بسط و توسعه داده می شود. برای مثال، مدل آماری یک طرح سه عاملی به شرح رابطه^۳ است.

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \varepsilon_{ijkl} \quad \left\{ \begin{array}{l} i = 1, 2, \dots, a \\ j = 1, 2, \dots, b \\ k = 1, 2, \dots, c \\ l = 1, 2, \dots, n \end{array} \right. \quad (3)$$

که در آن، y_{ijkl} مقدار متغیر پاسخ (در این تحقیق میانگین درصد خطا) حاصل از مشاهده

$ijkl$ ام، μ میانگین کل، τ_i تأثیر عامل اول در سطح i ، β_j تأثیر عامل دوم در سطح j ، γ_k تأثیر عامل سوم در سطح k ، $\beta\tau_{ij}$ تأثیر متقابل اول و دوم، $\gamma\tau_{ik}$ تأثیر متقابل اول و سوم، $\beta\gamma_{jk}$ تأثیر متقابل دوم و سوم و $\beta\gamma\tau_{ijk}$ تأثیر متقابل هر سه عامل، b تعداد

-
1. Levenberg – Marquardt Algorithm
 2. Response variable
 3. MSE :Mean Square Error
 4. MAPE :Mean Absolute Percentage Error
 5. MdAPE :Median Absolute Percentage Error
 6. MAD :Mean Absolute Deviation

سطوح عامل دوم، تعداد سطوح عامل سوم، n تعداد تکرار آزمایش^۱ و ε_{ijkl} جمله خطا است. با انجام آنالیز واریانس بر اساس مدل آماری و استفاده از آمار فیشر معنادار بودن هر یک از آثار اصلی و آثار متقابل تعیین می شود [۴۱].

۲-۳- آزمایش کفایت مدل^۲

در این مرحله، برای احراز اطمینان از نتایج حاصل از مدل آماری، دو فرض نرمال بودن باقیمانده‌ها^۳ و برابری واریانس‌ها^۴ بررسی می‌شود. برای تعیین فرض نرمال بودن باقیمانده‌ها، به‌طور معمول از آزمون Kolmogorov-Smirnov استفاده می‌شود [۴۱].
برای آزمون فرضیه برابری واریانس‌ها، از آزمون Bartlett استفاده می‌شود. آماره تست Bartlett به شرح رابطه ۴ است.

$$B = \frac{(\sum v_i) \ln(\sum v_i s_i^2 / \sum v_i) - \sum v_i \ln s_i^2}{1 + \{ \sum (1/v_i) - 1 / \sum v_i \} / \{ 3(k-1) \}} \quad (1)$$

که در آن K تعداد نمونه $V_i = ni - 1$ ، n_i تعداد مشاهدات نمونه i ، $S_i^2 = \sum_{j=1}^{n_j} (x_{ij} - \bar{x}) / (n_j - 1)$

است. آماره B دارای توزیع χ^2 با $K - 1$ درجه آزادی است [۴۱]. $V_i = ni - 1$

۲-۳- بهینه‌سازی و تعیین ساختار بهینه

آخرین گام در مدل پیشنهادی، تعیین ساختار بهینه شبکه عصبی مصنوعی است. در این تحقیق، ساختار بهینه از بین ۳۶۰ ترکیب در جدول ۱ انتخاب می‌شود. در طراحی آزمایشها برای تعیین ترکیب بهینه از تکنیک رویه پاسخ استفاده می‌شود. رویه پاسخ، یک رابطه رگرسیونی است که در آن متغیر پاسخ، متغیر وابسته و عوامل طراحی، متغیرهای مستقل هستند. در اکثر مطالعات، رویه پاسخ به صورت یک عبارت ریاضی درجه ۲ کامل^۵ به شرح رابطه ۵ در نظر گرفته می‌شود [۴۱].

-
1. Replication
 2. Model Adequacy
 3. Normality of Residuals
 4. Equality of Variances
 5. Full Quadratic Equation



$$\hat{Y} = \hat{\beta}_0 + \dots + \hat{\beta}_i x_i + \dots + \hat{\beta}_{ii} x_i^2 + \dots \quad (2)$$

$$+ \hat{\beta}_{ij} x_i x_j + \dots + \hat{\beta}_{ijj} x_i^2 x_j + \dots + \hat{\beta}_{ijij} x_i^2 x_j^2 + \dots + \hat{\beta}_{ijijj} x_i^2 x_j^2 x_j + \dots$$

که در آن x_i سطوح مربوط به تأمین عامل طراحی، $\hat{\beta}$ ضرایب مجهول و \hat{Y} مقدار مشاهده شده شاخص خطا به ازای هر ترکیب است. در رابطه ۵ ضرایب مجهول براساس روش کمترین مربعات به دست می آید. پس از تعیین ضرایب مجهول و انطباق یک رویه بر نتایج حاصل از آزمایش، ترکیب بهینه از حل یک مسأله بهینه سازی بدون محدودیت به دست می آید.

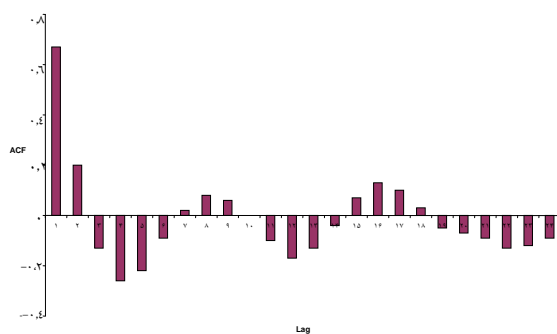
۳- اجرای مدل پیشنهادی

تأمین و توزیع به موقع نفت گان، اصلی ترین منبع تأمین کننده انرژی در بخشهای متنوع اقتصادی کشور از قبیل نیروگاهها، حمل و نقل، صنایع، کشاورزی، کشتیرانی و خانگی اهمیت ویژه ای دارد. بروز هرگونه نارسایی در تأمین به موقع نفت گان تبعات اقتصادی، سیاسی و اجتماعی دارد و به همین دلیل باید مصرف آن به صورت دقیق پیش بینی شود. در پیش بینی کوتاه مدت و بلندمدت تقاضای منابع انرژی از جمله مصرف فرآورده های نفتی در دهه های اخیر تحقیقات زیادی انجام شده و در ذیل آن از مدل های متنوعی استفاده شده است. Edigar و Akar از روش های ARMA و ARIMA برای پیش بینی تقاضای حامل های انرژی در ترکیه طی دوره زمانی ۲۰۰۵ تا ۲۰۲۰، استفاده کردند [۴۲]. Ghosh با استفاده از روش همگرایی و مدل تصحیح خطا برای پیش بینی تقاضای فرآورده های نفتی هند طی دوره ۲۰۰۱-۲۰۱۲، استفاده کرد. [۴۳؛ صص ۲۰۳۲-۲۰۳۷]. Pokharel از سیستم معادلات همزمان برای پیش بینی مصرف حامل های انرژی در نپال استفاده کرد [۴۴]. Nasr و همکاران، از یک شبکه عصبی بازگشتی برای پیش بینی مصرف بنزین استفاده کردند [۴۵؛ صص ۸۹۳-۹۰۵]. Iniyar و Jebaraj مرور جامعی بر مدل های انرژی داشتند [۴۶؛ صص ۵۶-۶۹]. در این بین از میان انبوه مدل های توسعه داده شده، شبکه عصبی نسبت به سایر تکنیکها نظیر اقتصادسنجی، مدل های Box و Jenkins و حتی تلفیق این تکنیکها، نوعاً از عملکرد بهتری برخوردار بوده است [۴۶؛ صص ۵۶-۶۹؛ ۴۷]. در این تحقیق به دلیل مزایای شبکه عصبی مصنوعی برای پیش بینی از شبکه عصبی مصنوعی و برای طراحی معماری شبکه عصبی مربوط به آن از مدل پیشنهادی در طراحی معماری استفاده شد.

داده‌های مسأله از مرکز آمار و اطلاعات شرکت ملی پخش فراورده‌های نفتی به دست آمده و مربوط به مصرف ماهانه نفت‌گاز طی سالهای ۱۳۷۸ تا سه ماهه سوم ۱۳۸۴ است.

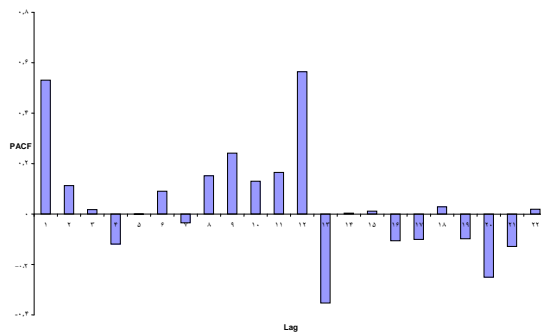
۳-۳- ورودیهای شبکه عصبی

تعیین مجموعه‌ای مناسب از متغیرهای ورودی نقش ویژه‌ای در پیش‌بینی موفق دارد. در پیش‌بینی ماهانه مصرف نفت‌گاز از دو نوع متغیر ورودی بنیادی شامل متغیر سال به‌عنوان نمایه‌ای از رشد سالیانه، متغیر فصل به‌عنوان نمایه‌ای از تغییرات فصلی، متغیر ماه خورشیدی به‌عنوان نمایه‌ای از مقتضیات رسمی کشور و حداکثر و حداقل دمای هوا در تمامی ایستگاههای سینوپتیک کشور و همچنین مقادیر تأخیر یافته (LAG) استفاده شده‌است. در این تحقیق برای تعیین تعداد مناسب متغیرهای تأخیر یافته از تابع خود همبستگی جزئی (PACF)^۱ در مدل‌های باکس و جنکینز استفاده شد [صص ۳۷۴-۳۸۵]. در شکل‌های ۲ و ۳ نمودار ACF و PACF مربوط به داده‌های مسأله مورد نظر آورده است.



شکل ۲ نمودار ACF برای سری زمانی مصرف ماهانه نفت‌گاز

1. PACF: Partial Auto Correlation Function



شکل ۳ نمودار PACF برای سری زمانی مصرف ماهانه نفت گاز

بر اساس شکل‌های ۲ و ۳ ACF یک موج سینوسی را نشان می‌دهد و PACF بعد از تأخیر ۱ قطع می‌شود. در این صورت می‌توان آن را $AR(1)$ دانست که تعداد مقادیر تأخیر یافته مناسب برای مدل برابر عدد ۱ می‌شود.

۳-۳- افزایش داده‌ها

در مسأله پیش‌بینی به‌طور معمول داده‌ها به دو مجموعه آموزش و تست تقسیم بندی می‌شوند. از مجموعه داده‌های آموزش برای تعیین مقادیر اوزان شبکه عصبی و از مجموعه داده‌های تست برای ارزیابی توانایی شبکه عصبی و تعیین قابلیت تعمیم آن، استفاده می‌شود. در این تحقیق، در تقسیم‌بندی داده‌ها به دو مجموعه آموزش و تست از نسبت ۸۰ به ۲۰ استفاده شده است [صص ۲۳۳-۲۳۸].

۳-۳- طراحی بهینه معماری شبکه عصبی

در مسأله مورد نظر، هدف تعیین معماری بهینه شبکه عصبی به منظور پیش‌بینی مصرف ماهانه نفت گاز است.

۳-۵-۱- تعیین عوامل و سطوح هر یک از عوامل

در این بخش، عوامل و سطوح هریک در طراحی معماری شبکه عصبی تعیین می‌شود. در این تحقیق عوامل طراحی نوع شبکه، الگوریتم آموزش، تعداد لایه‌های پنهان، تعداد نرونهای هر لایه،

تابع تبدیل و تابع پیش-پردازش در نظر گرفته شده‌اند. عوامل ذکر شده و سطوح هریک از این عوامل در جدول ۱ عنوان شده است.

۳-۵-۲- طراحی و اجرای آزمایش

در این مرحله، معماری شبکه‌های عصبی با توجه به تک تک ترکیب‌ها در جدول طراحی و شبکه حاصل، آموزش داده می‌شود. در این تحقیق، فرایند آموزش به ازای هر ترکیب، ۵ مرتبه اجرا شده است. بنابراین در مجموع $(5 \times 360 = 1800)$ شبکه عصبی ایجاد و آموزش داده شده است. برای آموزش شبکه عصبی از جعبه ابزار شبکه عصبی نرم‌افزار MATLAB و برنامه‌نویسی در محیط MATLAB استفاده شده است. در این تحقیق چنانچه بیان شد از میانگین درصد خطا برای ارزیابی عملکرد شبکه عصبی و همچنین متغیر پاسخ در طراحی آزمایشها استفاده شده است.

۳-۵-۳- توسعه مدل آماری و آنالیز واریانس

در این مرحله پس از آموزش شبکه‌های عصبی و ثبت مشاهدات، مدل آماری توسعه داده شده و آنالیز واریانس انجام می‌شود که نتایج آن در جدول ۲ آورده شده است.

جدول ۲ نتایج آنالیز واریانس

Source	DF	Seq SS	MS	F	P
NT	۱	۵/۷۰E+۲۴	۵/۷۰E+۲۴	۷/۴۶	۰/۰۰۶
N۲	۲	۳/۴۳E+۲۵	۱/۷۱E+۲۵	۲۲/۴۶	۰/۰۰۰
PF	۱	۳/۹۸E+۲۳	۳/۹۸E+۲۳	۰/۵۲	۰/۴۷۱
TF	۱	۱/۷۷E+۲۴	۱/۷۷E+۲۴	۲/۳۱	۰/۰۲۹
N۱	۹	۲/۲۲E+۲۴	۲/۴۶E+۲۳	۰/۳۲	۰/۹۶۸
NT*N۲	۲	۱/۱۹E+۲۴	۵/۹۳E+۲۳	۰/۷۸	۰/۴۶۰
NT*PF	۱	۲/۱۸E+۲۳	۲/۱۸E+۲۳	۰/۲۹	۰/۵۹۳
NT*TF	۱	۱/۱۶E+۲۳	۱/۱۶E+۲۳	۰/۱۵	۰/۶۹۶
NT*N۱	۹	۵/۹۵E+۲۴	۶/۶۱E+۲۳	۰/۸۷	۰/۵۵۵
N۲*PF	۲	۲/۳۲E+۲۳	۱/۱۶E+۲۳	۰/۱۵	۰/۸۵۹
N۲*TF	۲	۸/۲۰E+۲۴	۴/۱۰E+۲۴	۵/۳۷	۰/۰۰۵



ادامه جدول ۲

Source	DF	Seq SS	MS	F	P
N ₂ *N ₁	۱۸	۲/۰۴E+۲۵	۱/۱۳E+۲۴	۱/۴۸	۰/۰۰۷
PF*TF	۱	۵/۴۴E+۲۳	۵/۴۴E+۲۳	۰/۷۱	۰/۳۹۹
PF*N ₁	۹	۵/۲۱E+۲۴	۵/۷۹E+۲۳	۰/۷۶	۰/۵۵
TF*N ₁	۹	۳/۲۹E+۲۴	۳/۶۶E+۲۳	۰/۴۸	۰/۰۰۹
NT*N ₂ *PF	۲	۱/۷۵E+۲۴	۸/۷۳E+۲۳	۱/۱۴	۰/۳۱۹
NT*N ₂ *TF	۲	۱/۲۱E+۲۴	۶/۰۴E+۲۳	۰/۷۹	۰/۴۵۴
NT*N ₂ *N ₁	۱۸	۱/۲۸E+۲۵	۷/۱۰E+۲۳	۰/۹۳	۰/۵۴۱
NT*PF*TF	۱	۸/۵۵E+۲۲	۸/۵۵E+۲۲	۰/۱۱	۰/۷۳۸
NT*PF*N ₁	۹	۷/۲۰E+۲۴	۸/۰۰E+۲۳	۱/۰۵	۰/۴۰۰
NT*TF*N ₁	۹	۲/۸۳E+۲۴	۳/۱۵E+۲۳	۰/۴۱	۰/۹۲۹
N ₂ *PF*TF	۲	۹/۴۶E+۲۳	۴/۷۳E+۲۳	۰/۶۲	۰/۵۳۹
N ₂ *PF*N ₁	۱۸	۱/۲۰E+۲۵	۷/۷۸E+۲۳	۱/۰۲	۰/۴۳۵
N ₂ *TF*N ₁	۱۸	۱/۱۲E+۲۵	۶/۲۲E+۲۳	۰/۸۱	۰/۶۸۵
PF*TF*N ₁	۹	۵/۹۳E+۲۴	۶/۵۹E+۲۳	۰/۸۶	۰/۵۵۸
NT*N ₂ *PF*TF	۲	۸/۴۶E+۱۹	۴/۲۳E+۱۹	۰/۷۶	۰/۵۵
NT*N ₂ *PF*N ₁	۱۸	۹/۵۴E+۲۴	۵/۳۰E+۲۳	۰/۶۹	۰/۸۲۰
NT*N ₂ *TF*N ₁	۱۸	۹/۹۴E+۲۴	۵/۵۲E+۲۳	۰/۷۲	۰/۷۹۰
NT*PF*TF*N ₁	۹	۶/۴۹E+۲۴	۷/۲۲E+۲۳	۰/۹۴	۰/۴۸۵
N ₂ *PF*TF*N ₁	۱۸	۱/۵۶E+۲۵	۸/۶۴E+۲۳	۱/۱۳	۰/۳۱۵
NT*N ₂ *PF*TF*N ₁	۱۸	۱/۸۱E+۲۵	۱/۰۱E+۲۴	۱/۳۲	۰/۱۶۷
Error	۹۶۰	۷/۳۳E+۲۶	۷/۶۴E+۲۳		
Total	۱۱۹۹	۹/۴۰E+۲۶			

بر اساس مقادیر P-Value در جدول آنالیز واریانس و بر اساس خطای نوع اول ۵ درصد، نتایج ذیل استناد می شود.

- نوع شبکه بر عملکرد شبکه مؤثر است.
- تعداد نرونها در لایه پنهان دوم بر عملکرد شبکه مؤثر است.
- تابع تبدیل بر عملکرد شبکه مؤثر است.
- تعداد نرونهای لایه پنهان اول، در کنار سایر عوامل تأثیر کمتری بر عملکرد شبکه دارد.

- تعداد نرونهای لایه پنهان اول و لایه پنهان دوم بر هم اثر متقابل دارند. به این معنا که لزوماً، اضافه کردن نرون در هر دو لایه به بهبود عملکرد شبکه منتهی نمی‌شود.
- تعداد نرونهای موجود در لایه پنهان دوم و تابع تبدیل بر هم اثر متقابل دارند. به این معنا که اضافه کردن تعداد نرون در کنار یک نوع تابع تبدیل به بهبود عملکرد شبکه منجر شده، در حالی که در کنار تابع تبدیل دیگر موجب تنزل عملکرد شبکه می‌شود.
- تعداد نرونهای لایه پنهان اول، اگرچه بر عملکرد شبکه، اثر قابل توجهی ندارد، اما در کنار تابع تبدیل بر عملکرد شبکه مؤثر است و به عبارتی با آن اثر متقابل دارد.
- به همین ترتیب سایر آثار معنادار نیستند.

چنانچه مشاهده می‌شود، برخی عوامل، علاوه بر تأثیر بر عملکرد شبکه با سایر عوامل دارای اثر متقابل هستند و در طراحی بهینه شبکه عصبی نمی‌توان آنها را در نظر نگرفت. این در حالی است که Zhang در بررسی میزان تأثیر نرونهای ورودی و نرونهای لایه پنهان در پیش‌بینی سری‌های زمانی آنها را مستقل از هم در نظر می‌گیرد و از اثر متقابل بین آنها که می‌تواند بر نتایج تأثیرگذار باشد، چشم پوشی می‌کند [۳۸۱-۳۹۶].

۳-۵-۴- آزمایش کفایت مدل

برای تعیین فرض نرمال بودن باقیمانده‌ها بر اساس تست Kolmogorov-Smirnov مقدار آماره KS برابر $1/0.5$ و مقدار P-Value برابر $17/0$ می‌شود که با توجه به مقدار P-Value فرض نرمال بودن قابل قبول است.

برای آزمون برابری واریانس‌ها با انجام تست Bartlett مقدار آماره B برابر $1/47$ و مقدار P-Value برابر $29/0$ می‌شود که مقدار نسبتاً بالای P-Value فرض برابری واریانس‌ها را تأکید می‌کند

۳-۵-۵- تعیین ساختار بهینه

در این مرحله با توجه به آثار معنادار در جدول آنالیز واریانس، یک چندجمله‌ای درجه دو کامل بر داده‌های آزمایش انطباق داده می‌شود و آنگاه با استفاده از بهینه‌سازی غیرخطی بدون محدودیت، جواب بهینه به دست می‌آید. در مسأله تحت بررسی، از آنجا که معادله چندجمله‌ای

دارای ۶۶ متغیر است برای رعایت اختصار از ذکر رابطه طولانی آن اجتناب می‌شود. پس از بهینه‌سازی معادله مزبور ساختار بهینه معماری شبکه عصبی که در پیش‌بینی مصرف ماهانه نفت‌گاز استفاده شده است، بشرح جدول ۳ به دست می‌آید.

جدول ۳ ترکیب بهینه عوامل طراحی معماری شبکه عصبی

ارزش	عامل طراحی
شبکه پیش‌خوراند	نوع شبکه
۸	تعداد نرونهای لایه پنهان اول
صفر	تعداد نرونهای لایه پنهان دوم
MinMax	تابع پیش‌پردازش
Tansig	تابع تبدیل لایه پنهان

بنابر جدول ۳ معماری شبکه حاصل عبارت است از یک شبکه پیش‌خوراند با یک لایه پنهان و ۸ نرون که در آن از تابع تبدیل Tansig در لایه پنهان و از تابع MinMax برای پیش‌پردازش داده‌ها استفاده شده است.

۳-۳- اجرای مدل پیش‌بینی

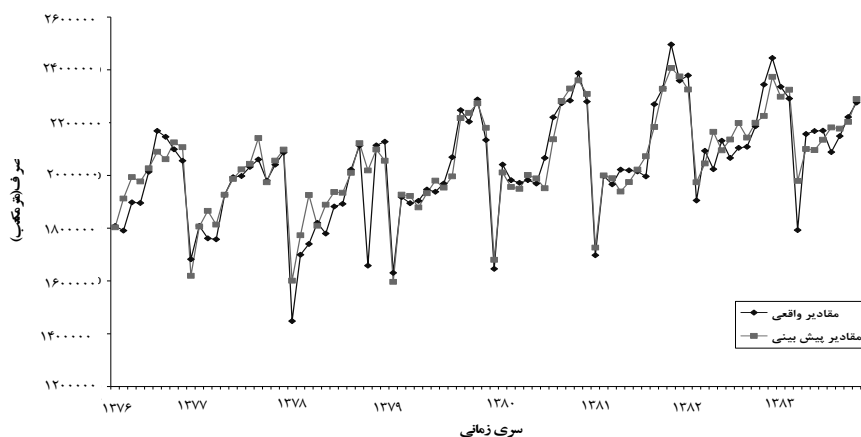
با تنظیم معماری شبکه عصبی براساس جدول ۳ و ایجاد و اجرای شبکه عصبی، نتایج ذیل بر حسب میانگین درصد خطا (MAPE) به دست می‌آید جدول ۴.

جدول ۴ ترکیب بهینه عوامل طراحی معماری شبکه عصبی

MAPE	شرح
٪۴٫۷	مجموعه آموزش
٪۶٫۴	مجموعه تست
٪۵	مجموعه کل داده‌ها

در شکل ۴ به صورت گرافیکی مقایسه بین پیش‌بینی حاصل از شبکه عصبی و مقادیر

واقعی آمده است.



شکل ۴ نمودار مقایسه بین پیش‌بینی و مقادیر واقعی مصرف ماهانه نفت‌گاز

■ خطوط پرننگ که با علامت ♦ متمایز شده‌اند، مقادیر واقعی و خطوط کم‌رنگ که با علامت متمایز شده‌اند مربوط به مقادیر پیش‌بینی است. چنانچه در شکل ۴ مشاهده می‌شود، مقادیر حاصل از پیش‌بینی بر مقادیر واقعی تطابق بسیار خوبی دارند.

۴- مقایسه و ارزیابی

در این بخش، ابتدا عملکرد مدل پیشنهادی در طراحی معماری شبکه عصبی با شبکه‌ای حاصل از رویکرد سعی و خطا، مقایسه و مزیت آن نشان داده می‌شود. آنگاه دقت و کیفیت پیش‌بینی حاصل از شبکه عصبی، صرف‌نظر از موضوع طراحی معماری شبکه با دو نوع مدل آماری، یعنی رگرسیون و ARIMA مقایسه می‌شود.

۴-۳- مقایسه عملکرد مدل پیشنهادی با دو رویکرد در طراحی معماری

در این بخش برای نشان دادن برتری مدل پیشنهادی به مقایسه عملکرد آن با شبکه عصبی که طراحی معماری آن از روش سعی و خطا به دست آمده و در غالب تحقیقات استفاده می‌شود، پرداخته می‌شود. در جدول ۴ ساختار شبکه حاصل از استفاده از روش سعی و خطا و نتایج حاصل از آن براساس



میانگین درصد خطای (MAPE) کل داده‌ها آمده است

جدول ۵ معماری شبکه عصبی با استفاده از روش سعی و خطا

ارزش	عامل طراحی
پیش خوراند	نوع شبکه
۵	تعداد نرونهای لایه پنهان اول
۰	تعداد نرونهای لایه پنهان دوم
StDev	تابع پیش پردازش
Tansig	تابع تبدیل لایه پنهان
میانگین درصد خطای کل داده‌ها = ۷	
میانگین درصد مجموعه تست = ۷/۱۵ درصد	

بر طبق جدول ۵ و با توجه به میانگین درصد خطا مشاهده می‌شود که استفاده از مدل پیشنهادی، نسبت به روش سعی و خطا دارای مزیت است.

۲-۴- مقایسه با روشهای آماری

۱-۲-۴- مقایسه با مدل رگرسیون

در مدل رگرسیون مانند شبکه عصبی از متغیرهای سال به عنوان نمایه‌ای از رشد سالیانه، فصل به عنوان نمایه‌ای از تغییرات فصلی، ماه خورشیدی به عنوان نمایه‌ای از مقتضیات رسمی کشور و حداکثر و حداقل دمای هوا در تمامی ایستگاههای سینوپتیک کشور و مقادیر تأخیریافته (LAG) به عنوان متغیر مستقل و مصرف واقعی به عنوان متغیر وابسته استفاده شده است. بر این اساس مدل رگرسیونی به شرح رابطه ۸ خواهد بود.

$$Y_t = \beta_0 \cdot X_{Year,t} + \beta_1 \cdot X_{Season,t} + \beta_2 \cdot X_{Month,t} + \beta_3 \cdot X_{MaxT,t} + \beta_4 \cdot X_{MinT,t} + \beta_5 \cdot Y_{t-1} + \epsilon_t \quad (3)$$

که در آن $X_{Year,t}$: متغیر سال: $X_{Season,t}$: متغیر فصل: $X_{Month,t}$: متغیر ماه: $X_{MaxT,t}$:

حداکثر دمای گزارش شده در کلیه ایستگاههای سینوپتیک کشور، $X_{\text{MinT},t}$: حداقل دمای گزارش شده در کلیه ایستگاههای سینوپتیک کشور: Y_{t-1} : مصرف ماه گذشته و ε_t : جمله خطا است. در رابطه (۶)، β_i ها ($i=1, \dots, 5$) ضرایب مجهول است که با روش کمترین مربعات برآورد می‌شود. خلاصه نتایج آماری مدل رگرسیون بشرح جدول ۶ است.

جدول ۶ نتایج حاصل از مدل رگرسیون

R Square		۰/۷۹۸۳۳۶۴۱۴	
Adjusted R Square		۰/۷۷۵۱۹۴۶۹	
Variable	Coefficients	t Stat	P-value
عرض از مبدأ	-۱۶۶۵۰۱۴	-۲/۶۳۹۱۱۲۳۶۴	۰/۰۱۰۵۳۶۶۴۲
سال	۳۶۷۸۴/۹	۳/۷۹۱۲۲۱۸۰۳	۰/۰۰۰۳۴۶۰۸۳
فصل	۶۵۳۹۴/۷	۳/۶۵۸۵۱۲۶۴۹	۰/۰۰۰۵۳۰۹۱۲
ماه	-۱۰۴۴۵۴/۵	-۱/۸۶۰۵۲۵۲۹۴	۰/۰۶۷۶۳۳۲۵۸
T حداکثر	-۱/۶۶۱۹۲	۱/۵۷۱۵۳۹	۰/۱۲۱۲۳۱
T حداقل	-۰/۶۶۹۰۷	-۰/۷۵۱۴۷۱	۰/۴۵۵۲۵۹
Lag ۱	-۰/۱۰۸۸۰۸	۱/۱۳۱۶۵۸	۰/۲۶۲۲۰۷

برطبق مقدار R Square در جدول ۷ در عین نیکویی مدل رگرسیونی، اما میانگین درصد خطای (MAPE) آن ۱۷ درصد می‌شود که در مقایسه با میانگین درصد خطای مربوط به کل داده‌ها در مدل شبکه عصبی پیشنهادی از کیفیت پایینتری برخوردار است.

۲-۲-۴- مقایسه با مدل ARIMA

به‌طور عموم مدل $ARIMA(p,q)$ عبارت است از ترکیب خطی مقادیر و خطاهای گذشته که به‌صورت رابطه ۷، بیان می‌شود.

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (7)$$



که در آن $\{\phi_i\}$ و $\{\theta_i\}$ ضرایب مدل و p و q به ترتیب مرتبه اتورگرسیو و میانگین متحرک است. برای تعیین p و q از نمودارهای PACF و ACF استفاده می شود. براساس نمودارهای PACF و ACF سری زمانی مصرف نفت گاز و پس از انجام آنالیز آماری، ضرایب مدل سری زمانی و سطوح معنادار بودن هر یک از ترمهای اتورگرسیو و میانگین متحرک به شرح جدول ۷ بدست می آید.

جدول ۷ نتایج حاصل از تحلیل سری زمانی

Type	Coef	SE Coef	T	P
AR(۱)	۰/۳۹۴۵	۰/۱۱۴۳	۳/۴۵	۰/۰۰۱
MA(۱)	۰/۹۲۹۰	۰/۰۴۱۲	۲۲/۵۲	۰/۰۰۰

میانگین درصد خطای حاصل از اجرای مدل $ARIMA(۱,۱,۱)$ برابر ۱۲ درصد می شود و مشاهده می شود در این حالت نیز در مقایسه با میانگین درصد خطای مربوط به کل داده ها در مدل شبکه عصبی پیشنهادی از کیفیت پایبندی برخوردار است.

۵- نتیجه گیری

یکی از گامهای مهم در توسعه شبکه های عصبی مصنوعی طراحی معماری شبکه است که تأثیر بسزایی بر کیفیت عملکرد شبکه دارد. در طراحی معماری شبکه های عصبی مصنوعی عواملی از قبیل تعداد لایه های پنهان، تعداد نرونها در هر لایه، توابع تبدیل و الگوریتم آموزش باید تعیین شوند. محققان در طراحی معماری شبکه، به طور عمده از طریق سعی و خطا عمل می کنند و یا اینکه اثر متقابل بین عوامل مختلف در طراحی معماری شبکه را در نظر نمی گیرند. در این تحقیق، یک مدل مبتنی بر تکنیک طراحی آزمایشها برای طراحی بهینه معماری شبکه عصبی با یادگیری هدایت شده ارائه شده است. در مدل پیشنهادی نوع شبکه، تعداد لایه های پنهان، تعداد نرونهای هر لایه، تابع پیش پردازش، تابع تبدیل و اثر متقابل بین عوامل ذکر شده در نظر گرفته شده است. نتایج تجربی این تحقیق نشان می دهد که این عوامل علاوه بر تأثیر ثابتی که بر عملکرد شبکه دارند با عوامل دیگر اثر متقابل دارند. این در حالی است که در ادبیات موضوع، اثر متقابل بین عوامل طراحی در نظر گرفته نمی شود.

در این تحقیق از مدل پیشنهادی برای طراحی معماری شبکه عصبی به منظور پیش‌بینی مصرف نفت‌گاز کل کشور استفاده شده است. به‌منظور مقایسه عملکرد مدل پیشنهادی با استفاده از روشهای مرسوم در طراحی معماری، یک مدل شبکه عصبی برای پیش‌بینی مصرف نفت‌گاز توسعه داده شده و طی آن برتری مدل پیشنهادی نشان داده شده است. همچنین برای مقایسه عملکرد شبکه عصبی با روش‌های آماری، دو مدل با استفاده رگرسیون و مدل‌های باکس و جنکینز طراحی شده است. نتایج به‌دست آمده در پیش‌بینی مصرف نفت‌گاز در این بخش نشان می‌دهد که پیش‌بینی با شبکه عصبی طراحی شده جوابهای بهتری دارد.

۶- تشکر و قدردانی

شایسته است در این تحقیق به سبب حمایت و در اختیار گذاشتن اطلاعات مربوط به داده‌های مصرف نفت‌گاز از مدیریت محترم پژوهش و توسعه شرکت ملی پخش فراورده‌های نفتی و همکاران محترم تشکر و قدردانی به‌عمل آید.

۷- منابع

- [1] Rumelhart D. E., Hinton G. E., Williams R. J.; Learning internal representation by back-propagating errors; MIT Press, 1986.
- [2] Widrow B., Rumelhart D. E., Lehr M. A.; "Neural networks: Applications in industry, business and science"; *Communications of the ACM*, Vol.37, 1994.
- [3] Granger C. W. J., Terasvirta T.; Modeling nonlinear economic relationships; Oxford University Press, 1993.
- [4] Hwang H.B., "Insights into neural-network forecasting of time series corresponding to ARMA(p,q) structures"; *Omega* 29, 2001.
- [5] Zhang G. P., Patuwo B. E., Hu M. Y.; "A simulation study of artificial neural networks for nonlinear time-series forecasting"; *Computers & Operations Research*, Vol. 28, 2001.
- [6] Marquez L.; "Function approximation using neural networks: A simulation study"; Ph.D. Dissertation, University of Hawaii, 1992.
- [7] Ash T.; "Dynamic node creation in backpropagation networks"; *Connection*



Science, Vol.1, 1998.

- [8] Fahlman S.E., Lebiere C.; The cascade-correlation learning architecture; Advances in Neural Information Systems; Morgan-Kaufmann, Los Altos, CA, 1990.
- [9] Kwok T.Y., Yeung D.Y.; "Experimental analysis of input weight freezing in constructive neural networks"; IEEE International Conference on Neural Networks, San Francisco, CA, 1993.
- [10] Islam M. M., Murase, K., "A new algorithm to design compact two-hidden-layer artificial neural networks", Neural Networks, Vol.14, 2001, 1265-1278.
- [11] Miller G.F., Todd P.M., Hedge S. U.; "Designing neural networks using genetic algorithm"; In: Proceedings of the Third International Conference on Genetic Algorithms, 1989.
- [12] Schaffer J. D., Caruana R. A., Eshelman L.J., "Using genetic search to exploit the emergent behavior of neural networks"; Physica D. Vol.42, 1990.
- [13] Arifovic J., Gencay R., "Using genetic algorithms to select architecture of a feedforward artificial neural network"; Physica A, Vol.289, 2001.
- [14] Benardos P.G., Vosniakos G., "Optimization feedforward artificial neural network architecture"; Engineering applications of artificial intelligence, article in press, 2006.
- [15] Kwok T.Y., Yeung D.Y., "Constructive algorithms for structure learning in feedforward neural networks for regression problems"; IEEE Transactions on Neural Networks, Vol.8, 1997.
- [16] Khaw J.F.C., Lim B.S., Lim L.E.N. "Optimal design of neural networks using the Taguchi method"; Neurocomputing, Vol.7, 1995.
- [17] Lin N.H., Chen J.C., "Evaluation of fuzzy-nets training efficiency"; Proceedings of the 1996 Asian Fuzzy Systems Symposium, Taiwan, ROC, 1996.
- [18] Lin T.Y., Tseng C.H.; "Optimal design for artificial neural networks: an example in a bicycle derailleur systems"; Engineering Applications of Artificial Intelligence, Vol.13, No.1, 2000.
- [19] Teo M.Y., Sim S.K., "Training the neocognitron network using design of

- experiment”; *Artificial Intelligence in Engineering*, Vol.9, No.2, 1995.
- [20] Wang Z., Massimo C. D., Tham M. T., Morris A. J.; “A procedure for determining the topology of multilayer feedforward neural networks”; *Neural Networks*, Vol.7, No.2, 1994.
- [21] Murata N., Yoshizawa S., Amari S.; “Network information criterion-determining the number of hidden units for an artificial neural network model”; *IEEE Transactions on Neural Networks*, Vol.5, No.6, 1994.
- [22] Roy A., Kim L. S., Mukhopadhyay S.; “A polynomial time algorithm for the construction and training of a class of multilayer perceptrons”; *Neural Networks*, Vol.6, 1993.
- [23] Zhang G., Patuwo B.E., Hu M. Y.; “Forecasting with artificial neural networks: The state of the art”; *International Journal of Forecasting*, Vol.14, 1998.
- [24] Kang, S.; “An Investigation of the Use of Feedforward Neural Networks for Forecasting”; Ph.D. Thesis, Kent State University, 1991.
- [25] Tang Z., Almeida C., Fishwick P. A.; “Time series forecasting using neural networks vs Box-Jenkins methodology”; *Simulation*, Vol.57, No.5, 1991.
- [26] Kermanshahi B.; “Recurrent neural network for forecasting next 10 years loads of nine Japanese utilities”; *Neurocomputing*, Vol.23, 1998.
- [27] Kuan C.L.; “Forecasting exchange rates using feed forward and recurrent neural network”; *Journal of Applied Econometrics*, Vol.10, 1995.
- [28] Connor J.T., Martin R.D., Atlas L. E.; “Recurrent neural network and robust time series prediction”; *IEEE Transaction on Neural Network*, Vol.16, 1994.
- [29] Mediros M.C.P.; “What are the effects of forecasting linear time series with neural network?”; *Engineering Intelligent Systems* 4, 2001.
- [30] Parrlos A.G., Rais O.T., Atiya A. F.; “Multi step ahead prediction using dynamic recurrent neural network”; *Neural Network* 13, 2000.
- [31] Husken M.S.; “Recurrent neural network for time series classification”; *Neurocomputing*, Vol.50, 2003.
- [32] Vermark J.B.; “Recurrent neural network for short term load forecasting”;



- IEEE Transaction on Power System, Vol.1, No.13, 1998.
- [33] Barron A.R.; "A comment on Neural networks: A review from a statistical perspective"; Statistical Science, Vol.9 , No.1, 1994.
- [34] Srinivasan D., Liew A.C., Chang C. S.; "A neural network short-term load forecaster"; Electric Power Systems Research, Vol.28, 1994.
- [35] Zhang X.; "Time series analysis and prediction by neural networks"; Optimization Methods and Software 4, 1994.
- [36] Fausett L.; Fundamentals of Neural Networks; Prentice-Hall, 1994.
- [37] Lapedes A.,Farber R.; How neural nets work; In: Anderson, D.Z., (Ed.), Neural Information Processing Systems, American Institute of Physics, New York, 1988.
- [38] Sharda R., Patil R.B.; "Neural networks as forecasting experts: An empirical test"; In: Proceedings of the International Joint Conference on Neural Networks, Washington, D.C.,Vol. 2, 1990.
- [39] Makridakis S., Wheelwright S.C., McGee V. E., Forecasting: Methods and applications, 2nd ed, John Wiley, New York, 1983.
- [40] Yokum J.T., Armstrong J. S.; "Beyond accuracy: comparison of criteria used to select forecasting methods"; *International Journal of Forecasting*, Vol.11, 1995.
- [41] Montgomery D. C.; Design and analysis of experiments; Wiley, New York, 1991.
- [42] Ediger V. S., Akar S.; "Arima forecasting of primary energy demand by fuel in Turkey"; Energy Policy, Article in Press, 2006.
- [43] Ghosh S.; "Future demand of petroleum products in India"; Energy Policy, 34, 2006.
- [44] Pokharel S.; "An econometric analysis of energy consumption in Nepal"; Energy Policy, Article in Press,2006.
- [45] Nasr G.E., Badr E.A., Joun C.; "Backpropagation neural networks for modeling gasoline consumption"; Energy Conversion and Management ,Vol.44, 2003.
- [46] Jebaraj S., Iniyar S.; "A review of energy models, renewable and sustainable energy review";Vol.15, 2004.
- [47] Box G. E., Jenkins G. M.; Time series analysis: Forecasting and control; San Francisco, 1976.

- [48] Tang Z., Fishwick P. A.; “Feedforward neural nets as models for time series forecasting”; *ORSA Journal on Computing* . Vol.5, No.4, 1993.
- [49] Granger, C.W. J.; “Strategies for modeling nonlinear time-series relationship”; *The Economic Record*, Vol.69, 1993.