

**تحلیل آماری و برآورد فاصله اطمینان پیش‌بینی شبکه عصبی
ترکیبی به منظور مقایسه با مدل خطی ARIMA:
مطالعه موردی مصرف ماهانه گاز طبیعی در بخش خانگی ایران**

سید کمال صادقی^۱ سید مهدی موسویان^۲
تاریخ دریافت: ۹۳/۲/۲۴ تاریخ پذیرش: ۹۳/۸/۲۰

چکیده

مصرف گاز طبیعی به عنوان یکی از مهم‌ترین حامل‌های انرژی، طی سال‌های اخیر روند صعودی را داشته و مدیریت مصرف و برنامه‌ریزی جهت تأمین نیازهای آن، نیازمند شناخت وضعیت مصرف کنونی و پیش‌بینی روند آتی آن می‌باشد. با معرفی و کاربرد گسترده مدل‌های مختلف همچون شبکه‌های عصبی مصنوعی جهت برآورد روند آتی مصرف و از طرفی تصادفی بودن آن‌ها، آگاهی از دقت این مدل‌ها جهت نیل به هدف پیش‌بینی دقیق‌تر، اهمیت بیشتری یافته است. پژوهش حاضر سعی دارد با به‌کارگیری شبکه‌های عصبی مصنوعی به عنوان مدلی غیرخطی و مدل خطی ARIMA در پیش‌بینی مصرف ماهانه گاز طبیعی در بخش خانگی ایران به عنوان عمده‌ترین بخش مصرف‌کننده، به مقایسه دقیق‌تر این پیش‌بینی‌ها با استفاده از باز نمونه‌گیری از نمونه‌ها بپردازد. بدین منظور ابتدا آموزش شبکه با استفاده از الگوریتم‌های ژنتیک و ازدحام

۱. دانشیار دانشکده اقتصاد، مدیریت و بازرگانی دانشگاه تبریز،

Email: sadeghiseyedkamal@gmail.com

۲. دانشجوی دکتری علوم اقتصادی دانشگاه تبریز، (نویسنده مسئول)،

Email: me.mousavian@gmail.com

ذرات صورت گرفته و مقایسه آن‌ها با استفاده از روش «10-fold» حاکی از عملکرد بهتر الگوریتم ازدحام ذرات جهت آموزش شبکه عصبی با استفاده از باز نمونه‌گیری با جایگذاری از داده‌های اردیبهشت‌ماه ۱۳۸۱ تا اسفندماه ۱۳۸۸ به تعداد ۲۰۰۰ بار توسط الگوریتم ازدحام ذرات آموزش داده شد و مصرف ماهانه گاز طبیعی در بخش خانگی طی سال‌های ۱۳۸۹ و ۱۳۹۰ توسط آن‌ها پیش‌بینی و فاصله اطمینان ۹۵ درصدی برای پیش‌بینی‌ها محاسبه شد. نتایج بررسی معنی‌داری اختلاف پیش‌بینی مدل ترکیبی شبکه عصبی با مدل ARIMA و همچنین مقادیر واقعی، بر اساس فاصله اطمینان به دست آمده حاکی از عملکرد بهتر شبکه عصبی ترکیبی نسبت به مدل ARIMA در اغلب ماه‌ها بود.

واژه‌های کلیدی: الگوریتم PSO، بوت استرپ، پیش‌بینی، شبکه‌های عصبی، گاز طبیعی.

JEL :O13, Q40, C53, E27

۱. مقدمه

پیش‌بینی و برآورد صحیح مصرف آتی، چه در حوزه اقتصاد انرژی و چه سایر بخش‌ها، همواره مورد توجه پژوهشگران و سیاست‌گذاران بوده است. اهمیت این مقوله منجر به مطالعات وسیعی در این حیطة و معرفی مدل‌های مختلف جهت نیل به هدف پیش‌بینی بهتر گردیده است. با گسترش و معرفی روش‌های جدید پیش‌بینی، مدل‌های غیرخطی نیز در این زمینه استفاده شدند. طی دهه‌های اخیر، شبکه‌های عصبی مصنوعی (ANN^۱) به طور گسترده‌ای در علوم مختلف مورد استفاده قرار گرفته است. یکی از کاربردهای این شبکه‌ها پیش‌بینی متغیرهای سری زمانی می‌باشد که در این زمینه مطالعات بسیاری در حوزه اقتصاد انرژی انجام شده است. در این مطالعات، عموماً شبکه‌های عصبی به عنوان یک الگوی غیرخطی، با مدل‌های سری زمانی همچون مدل خود رگرسیون میانگین متحرک انباشته (ARIMA^۲) خطی، مورد مقایسه قرار گرفته است. این مقایسه‌ها عمدتاً بر اساس معیارهای آماری همچون حداقل مجذور خطا (RMSE^۳) صورت گرفته اما با توجه به خاصیت تصادفی بودن شبکه‌های عصبی مصنوعی (به سبب انتخاب تصادفی وزن‌های^۴ اولیه)، نتایج پیش‌بینی نیز تصادفی می‌باشد و با هر بار آموزش^۵ شبکه عصبی مقادیر پیش‌بینی به دلیل تغییر وزن‌ها تغییر می‌کند. از این رو با آموزش مجدد شبکه عصبی، خطای پیش‌بینی نیز تغییر کرده که این تغییرات می‌تواند نتایج مقایسه بین مدل‌ها را تحت الشعاع خود قرار داده و اریب در نتایج را سبب گردد. در این مطالعه، با هدف بررسی و مقایسه دقیق‌تر شبکه‌های عصبی با مدل ARIMA سعی شده با باز نمونه‌گیری از داده‌ها، یک فاصله اطمینان برای پیش‌بینی‌های شبکه عصبی برآورد گردد و نتایج آن با مدل ARIMA مقایسه شود.

-
1. Artificial Neural Networks
 2. Autoregressive Integrated Moving Average
 3. Root Mean Square Error
 4. Weights
 5. Train

بخش خانگی با اختصاص ۲۹ درصد از کل مصرف گاز طبیعی طی سال ۱۳۸۹، یکی از عمده‌ترین بخش‌های مصرفی گاز طبیعی در ایران می‌باشد (ترازنامه انرژی ۱۳۹۰). طی سال‌های اخیر، سرعت توسعه در شبکه‌های فرعی گاز فراتر از سرعت طرح‌های توسعه تولید و حتی طرح‌های توسعه انتقال بوده است که این مسئله منجر به بروز مشکلاتی در تأمین گاز طی فصول سرد سال شده است؛ به نحوی که با اولویت دادن به تأمین گاز بخش خانگی، کمبود گاز سایر بخش‌های مصرفی به واسطه سوخت‌های مایع توأم با سوخت‌رسانی مستقیم و استفاده از ذخایر سوخت، جبران شده است (ترازنامه هیدروکربنی کشور، ۱۳۸۷). پیش‌گیری از مشکلات ذکر شده نیازمند برنامه‌ریزی تولید همسو با تغییرات مصرف است. از طرفی با آگاهی از وضعیت منابع و مصارف گاز طبیعی در کشور می‌توان امکان برنامه‌ریزی برای استفاده بهینه و کارآمد از این منبع و صادرات باثبات آن را فراهم نمود. در این راستا و با توجه به اهمیت گاز طبیعی در تأمین انرژی، در این پژوهش پیش‌بینی مصرف گاز در بخش خانگی به عنوان مطالعه موردی مد نظر قرار گرفته است.

در ایران، مطالعات نسبتاً زیادی پیرامون مصرف انرژی صورت گرفته است، اما مصرف گاز طبیعی سهم کمتری از این مطالعات را تشکیل می‌دهد. عمده‌ی این مطالعات به صورت سالانه بوده و به بررسی رابطه مصرف انرژی و متغیرهایی همچون درآمد و قیمت پرداخته و توجه نسبتاً کمتری به متغیرهایی همچون تغییرات آب و هوایی شده است. این پژوهش ضمن توجه به اثر تغییرات دما و لحاظ متغیرهای تأثیرگذار همچون قیمت و تعداد مشترکین، با باز نمونه‌گیری از داده‌ها و ساختن فاصله اطمینان برای پیش‌بینی‌های کوتاه مدت شبکه عصبی، به بررسی معنی‌دارای اختلاف بین این نتایج با پیش‌بینی مدل ARIMA پرداخته است. در ادامه، ابتدا مروری بر مطالعات انجام شده در زمینه پیش‌بینی گاز طبیعی و موارد مشابه انجام شده است، در پی آن مبانی نظری تحقیق بیان شده سپس الگوریتم ازدحام ذرات (PSO) و الگوریتم ژنتیک (GA) به صورت خلاصه

-
1. Particle Swarm Optimization
 2. Genetic Algorithm

بیان شده‌اند. در پایان نیز مدل مناسب برای پیش‌بینی مصرف برآورد و نتایج با هم مقایسه شده‌اند.

۲. ادبیات موضوع

۲-۱. پیشینه تحقیق

مطالعات اولیه پیرامون تقاضای گاز طبیعی از اواسط دهه ۱۹۶۰ شروع شد. در دهه‌های ۱۹۷۰ و ۱۹۸۰ مطالعات بیشتری در این حوزه به انجام رسید و در دهه ۱۹۹۰ افزایش تقاضای بین‌المللی برای گاز طبیعی و مطرح شدن آن به عنوان یک منبع انرژی نسبتاً پاک و ارزان و همچنین پیشرفت فناوری تبدیل گاز به صورت‌های مختلف با قابلیت حمل و نقل بهتر، گسترش بیشتر این مطالعات را در پی داشت (مشیری، شاهمرادی ۱۳۸۵). در این بخش به برخی از این مطالعات اشاره می‌شود.

تا کنون متغیرهای مختلفی جهت برآورد مصرف گاز طبیعی مورد استفاده قرار گرفته است. این مطالعات مصرف گاز طبیعی را به صورت تابعی از قیمت آن و درآمد (اردوئندو ۲۰۱۰؛ یو، لیم و کواک^۱ ۲۰۰۹، کشاورز حداد و میرباقری جم ۱۳۸۶)، قیمت‌های نسبی یا قیمت جانشین‌های آن (کشاورز حداد و میرباقری جم ۱۳۸۶)، دمای هوا یا تعداد روزهای سرد (تیمر و لمب ۲۰۰۷؛ لیو و لین ۱۹۹۱؛ هربرت و کریل ۱۹۸۹؛ آراس و آراس^۲ ۲۰۰۴؛ کریمی، صادقی مقدم و رهنما ۱۳۸۹)، تعداد انشعابات یا جمعیت (بالسترا و نرلاو^۳ ۱۹۶۶) و سایر متغیرهای مرتبط در نظر گرفته‌اند؛ اما در بسیاری از مطالعات صورت گرفته در خصوص تقاضای گاز و همچنین به زعم عمده‌ی کارشناسان این بخش، متغیر دما نقش تعیین‌کننده‌ای را در مدل‌سازی تقاضای گاز طبیعی دارد (میدی، محمدی، سلطان‌العلمایی ۱۳۸۹).

سیلر و مونز^۴ (۱۹۹۷)، همبستگی بین متغیر دما و مصرف ماهانه گاز طبیعی را در بخش‌های مسکونی و تجاری چند ایالت آمریکا بررسی کردند. آن‌ها متغیر دما را بر اساس

1. Erdogdu (2010), Yoo, Lim, Kwak (2009)

2. Timmer & Lamb (2007), Liu & Lin (1991), Herbert & Kreil (1989), Aras & Aras (2004)

3. Balestra & Nerlove

4. Sailor & Muñoz

میانگین وزنی متوسط دما محاسبه کرده و ارتباط آن با مصرف سرانه گاز طبیعی را مورد مطالعه قرار دادند. در مطالعه آن‌ها، به سبب وابسته بودن میزان مصرف گاز به تعداد مصرف‌کنندگان، از جمعیت هر ناحیه به عنوان وزن برای محاسبه دما استفاده شده است. آن‌ها مدل‌های جداگانه‌ای را مورد استفاده قرار دادند و بر اساس نتایج، اکثر مدل‌ها ضریب همبستگی پیرسون بزرگ‌تر از ۰/۹ را نشان می‌دادند، که استفاده از متغیر دما به عنوان تنها متغیر مستقل را حمایت می‌کرد. آن‌ها همچنین اثر یک درجه افزایش متوسط دمای ماهانه در کل آمریکا را، کاهش ۸,۱ درصدی مصرف گاز پیش‌بینی کردند.

سیلان و اوزتورک^۱ (۲۰۰۴) در مقاله‌ای تحت عنوان؛ برآورد تقاضای انرژی در ترکیه بر اساس شاخص‌های اقتصادی و استفاده از الگوریتم ژنتیک، به برآورد تقاضای انرژی پرداختند. شاخص‌های اقتصادی مورد استفاده آن‌ها شامل تولید ناخالص داخلی^۲، جمعیت، صادرات و واردات می‌شد. آن‌ها دو مدل خطی و نمایی را برای تخمین تقاضای دوره ۱۹۷۰ تا ۲۰۰۱ به کار بردند و با استفاده از الگوریتم ژنتیک پارامترهای مدل را برآورد کردند؛ سپس مصرف انرژی را از سال ۲۰۰۲ تا سال ۲۰۲۵ بر اساس سه سناریوی مد نظر خود پیش‌بینی کردند. بر اساس نتایج آن‌ها، یک رابطه قوی بین شاخص‌های اقتصادی آنان و مصرف انرژی وجود داشت.

وانگ، زنگ و چن^۳ (۲۰۱۵)، با ترکیب الگوریتم ADE^۴ و شبکه‌های عصبی مصنوعی، به مقایسه نتایج پیش‌بینی این شبکه با پیش‌بینی شبکه عصبی پس انتشار خطا^۵ و مدل ARIMA پرداختند. آن‌ها با مقایسه این نتایج در مورد پیش‌بینی تقاضای ماهانه الکتریسیته در منطقه شمال شرقی چین، عملکرد بهتر شبکه ترکیبی نسبت به شبکه پس انتشار خطا و همچنین خطای کمتر هر دو نوع شبکه نسبت به مدل ARIMA را نتیجه گرفتند.

-
1. Ceylan H, Ozturk H K
 2. Gross Domestic Product (GDP)
 3. Wang, Zeng & Chen
 4. Adaptive Differential Evolution
 5. Back Propagation

لطفعلی پور و باقری (۱۳۸۲) در مقاله‌ای با عنوان "تخمین تابع تقاضای گاز طبیعی مصارف خانگی شهر تهران"، توابع تقاضای کل گاز طبیعی و متوسط مصرف گاز طبیعی هر خانوار در شهر تهران را تخمین زده و در نهایت کشش‌های درآمدی و قیمتی گاز طبیعی را محاسبه کرده‌اند. داده‌های آن‌ها به صورت فصلی از سال ۱۳۷۴ تا ۱۳۷۸ را شامل می‌شد. الگوی آن‌ها شامل یک الگوی لگاریتمی و یک الگوی خطی بود و برآوردها با استفاده از روش ^۱ OLS صورت گرفته است. متغیرهای مورد استفاده در الگوهای آن‌ها شامل متوسط دما، قیمت گاز طبیعی و درآمد سرانه بود. مطابق نتایج به دست آمده؛ کشش تقاضا برای گاز طبیعی در بخش خانگی ناچیز بوده و کشش درآمدی تقاضا نیز کمتر از واحد و مبنی ضروری بودن مصرف گاز در سبد خانوار بود؛ همچنین مهم‌ترین عامل اثرگذار بر مصرف متوسط دما معرفی شد.

کشاوری حداد و میرباقری جم (۱۳۸۶)، تابع تقاضای گاز طبیعی را در دو بخش خانگی و تجاری و با استفاده از مدل ساختار سری زمانی (STSM)^۲ بررسی کردند. آن‌ها در مطالعه خود از داده‌های فصلی مقدار مصرف گاز طبیعی، دمای هوا، قیمت نسبی گاز طبیعی به برق و درآمد مصرف‌کننده برای سال‌های ۱۳۷۴ تا ۱۳۸۳ استفاده کردند. بر اساس نتایج به دست آمده، کشش مصرف سرانه گاز طبیعی نسبت به دما $-0/26$ برآورد شد و کشش‌های قیمتی و درآمدی نیز به ترتیب حدود $-0/13$ و $0/17$ به دست آمد.

کریمی و همکاران (۱۳۸۹)، در مقاله‌ای تحت عنوان «بررسی اثر تغییرات دما بر مصرف گاز طبیعی در ایران»، به برآورد مصرف گاز طبیعی با استفاده از رگرسیون داده‌های تلفیقی پرداختند. آن‌ها از داده‌های فروردین ۱۳۸۶ تا اسفند ۱۳۸۷ برای ۲۵ استان و در ۷ بخش مصرفی استفاده کردند. همچنین مصرف کل سالانه استان‌ها را نیز با داده‌های سالانه از سال ۱۳۷۹ تا ۱۳۸۷ برآورد نمودند. بر اساس نتایج آن‌ها، بین متغیر دما و مصرف در بخش‌های حمل و نقل، تجاری، عمومی، کشاورزی و خانگی رابطه معکوس وجود داشته و این رابطه در بخش خانگی دارای ضریب تعیین ۷۰ درصد بود.

1. Ordinary Least Square
2. Structure Time Series Model

در جمع‌بندی مطالعات صورت گرفته پیرامون مصرف گاز طبیعی و موضوعات مرتبط با انرژی می‌توان گفت؛ مطالعاتی که به پیش‌بینی مصرف این منابع با استفاده از تکنیک‌های نوین همچون الگوریتم ازدحام ذرات پرداخته‌اند، کمتر به خاصیت تصادفی بودن این الگوریتم‌ها توجه کرده‌اند. این در حالی است که مقایسه نتایج مربوط به این الگوریتم‌ها با یکدیگر و یا سایر روش‌ها بدون در نظر گرفتن این موضوع چندان خالی از اشکال نبوده و سطح اتکا به این نتایج را کاهش می‌دهد.

۲-۲. مبانی نظری

تحلیل تقاضای یکی از قدرتمندترین ابزارها به منظور درک و پیش‌بینی پدیده‌ها و متغیرهای اقتصادی می‌باشد (امامی میدی و همکاران، ۱۳۸۹). از طرفی عرضه گاز طبیعی در ایران تحت کنترل دولت قرار داشته و این موضوع سبب تعیین شدن قیمت و عرضه به صورت برون‌زا برای مصرف‌کنندگان می‌شود. ویژگی دولتی بودن و برون‌زا عمل کردن قیمت و عرضه انرژی، نیاز به در نظر گرفتن همزمان عرضه و تقاضا را از بین می‌برد (عباسی نژاد، صادقی، ۱۳۷۸)؛ بنابراین الگوسازی تقاضا در این بازار را می‌توان به صورت مجزا انجام داد.

چنانچه فرض کنیم خانوار به دنبال تخصیص مخارج انرژی خود بین گاز طبیعی و سایر حامل‌های انرژی باشد؛ مطلوبیت خانوار را می‌توان به صورت تابعی از مصرف گاز طبیعی و سایر حامل‌ها در نظر گرفت. در این پژوهش تابع مطلوبیت را به صورت رابطه (۱) فرض می‌کنیم.

$$U = U(Q_G, Q_{NG}) \quad (1)$$

در رابطه (۱)، Q_G بیانگر سطح مصرف گاز طبیعی در بازه زمانی ماهانه و Q_{NG} سطح مصرف سایر حامل‌های انرژی می‌باشد. در ادامه فرض می‌شود که خانوار به دنبال حداکثر سازی مطلوبیت خود با در نظر گرفتن قید بودجه زیر می‌باشد.

$$P_G Q_G + P_{NG} Q_{NG} = I \quad (2)$$

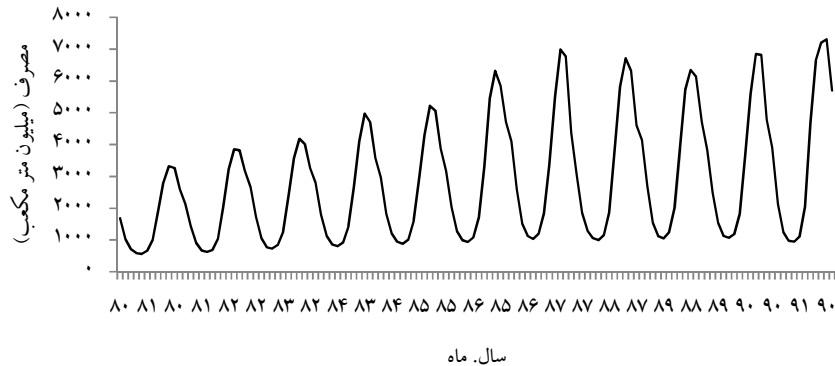
در معادله (۲)، P_G و P_G به ترتیب نماینده قیمت گاز طبیعی و قیمت سایر حامل‌های انرژی می‌باشد و I نیز سطح درآمد خانوار است. با حداکثر سازی مطلوبیت نسبت به قید بودجه، تقاضای مارشالی برای هر کدام از کالاها (گاز طبیعی و سایر حامل‌ها) به دست می‌آید (لیارد و والترز^۱، ۱۳۷۷). بنابراین تابع تقاضای گاز طبیعی را می‌توان به صورت زیر نشان داد:

$$Q = Q(P_G, P_{NG}, Y, Z) \quad (۳)$$

در رابطه (۳)، Z بیانگر سایر متغیرهای اثرگذار بر مصرف می‌باشد. با توجه به ویژگی‌های گاز طبیعی، در پیش‌بینی مصرف آن علاوه بر درآمد و قیمت به نظر می‌رسد موارد زیر نیز تحت عنوان متغیر Z بر مصرف گاز طبیعی اثرگذار باشند:

مصرف گاز طبیعی به عنوان خوراک و سوخت صنایع، به طور معمول روند ثابتی را در طی سال می‌گذراند ولی مصرف آن در بخش‌های خانگی و نیروگاهی طی فصول مختلف با نوسانات زیادی روبه‌رو می‌شود (تراز هیدروکربنی کشور ۱۳۸۷)؛ بنابراین مصرف گاز طبیعی به شرایط آب و هوایی نیز بستگی داشته به نحوی که در فصول سرد سال مصرف آن افزایش و در فصل بهار با کاهش روبه‌رو می‌شود. به عنوان نمونه، حدود ۳۸٫۶ درصد از کل مصرف گاز طبیعی سال ۱۳۸۷ در ۸ ماه ابتدای سال و ۶۱٫۴ درصد باقی‌مانده تنها در طی ۴ ماه پایان سال مصرف شده است (بر اساس آمار تراز هیدروکربنی کشور ۱۳۸۷).

نمودار ۱. مصرف ماهیانه گاز طبیعی در بخش خانگی



منبع: شرکت ملی گاز ایران

همان طور که در نمودار (۱) مشخص است، روند مصرف گاز طبیعی در بخش خانگی ایران دارای نوسانات فصلی است. به طوری که در فصول و ماه‌های سرد سال، و کاهش دما، میزان مصرف افزایش و با افزایش دما در فصول بهار و تابستان میزان مصرف آن کاهش می‌یابد. از این رو می‌توان انتظار داشت که بین مصرف گاز طبیعی و دمای هوا، ارتباط معکوس وجود دارد.

حال اگر روند مصرف گاز را به صورت اختلاف مصرف فعلی با مصرف همان ماه در سال قبل از آن در نظر بگیریم، اثر این نوسانات فصلی را می‌توان تا حدودی کاهش داد. از میان دیگر عوامل تأثیرگذار بر مصرف گاز طبیعی می‌توان به میزان دسترسی به آن و یا تعداد مشترکین اشاره کرد. طبیعتاً هر چه تعداد مشترکین افزایش یابد، در صورت ثبات سایر شرایط، می‌توان انتظار افزایش مصرف را داشت. همچنین بر اساس الگوی عادات مصرفی می‌توان مصرف فعلی را تابعی از مصرف دوره‌های قبل نیز در نظر گرفت.

۳. داده‌ها و روش‌شناسی

۳-۱. داده‌ها

در این مطالعه، به منظور پیش‌بینی مصرف گاز طبیعی دو نوع شبکه عصبی بر اساس متغیرهای ورودی مد نظر قرار گرفته است. ابتدا با توجه مباحث مطرح شده در بخش مبانی

نظری و داده‌های در دسترس، الگوی (۴) جهت پیش‌بینی شبکه عصبی استفاده شده است، سپس شبکه‌ای مشابه با مدل ARIMA تنها با در نظر گرفتن وقفه‌های مصرف به عنوان متغیرهای ورودی، آموزش داده شده است^۱:

$$NGC_t = F(T, N_t, NGC_{t-i}, P) \quad (۴)$$

در رابطه (۴)، NGC_t مصرف ماهیانه گاز طبیعی در بخش خانگی بر حسب میلیون مترمکعب در زمان t ، N_t تعداد مشترکین گاز طبیعی در بخش خانگی، T متوسط دما، NGC_{t-i} میزان مصرف ماهیانه گاز طبیعی در زمان $t-i$ و P قیمت حقیقی گاز طبیعی در بخش خانگی می‌باشد که به منظور محاسبه آن، قیمت اسمی گاز طبیعی بر اساس شاخص قیمت مصرف‌کننده (CPI^2) به قیمت ثابت سال ۱۳۸۳ تعدیل شده است. متوسط دما به صورت میانگین وزنی متوسط دمای ماهیانه مراکز استان‌های کشور محاسبه گردید. وزن‌ها نیز بر اساس تعداد نسبی مشترکین هر استان محاسبه شد. ضمن آنکه آمار مربوط به دما از سالنامه آماری ایران و متغیرهای مصرف گاز طبیعی و تعداد مشترکین از شرکت ملی گاز ایران اخذ شدند. قیمت گاز طبیعی نیز از ترازنامه انرژی وزارت نیرو استخراج گردید و شاخص قیمت مصرف‌کننده از بانک مرکزی ج.ا. ایران به دست آمد. متغیرهای ذکر شده شامل داده‌های فروردین‌ماه ۱۳۸۰ تا اسفندماه ۱۳۹۰ می‌باشند.

۲-۳. روش‌شناسی

روش‌های مورد استفاده برای پیش‌بینی را می‌توان به دو طبقه کلی تقسیم‌بندی نمود:
 ۱- روش‌های کیفی ۲- روش‌های کمی. روش‌های کیفی، ذهنی و مبتنی بر قضاوت افراد می‌باشند. این روش‌ها عمدتاً در مواقعی کاربرد دارند که داده‌های تاریخی کمی در دسترس باشند. از جمله‌ی این روش‌ها، می‌توان به روش‌های دلفی و سناریوسازی اشاره نمود. در روش‌های کمی از اطلاعات و داده‌های گذشته جهت پیش‌بینی تقاضا برای دوره‌های آینده استفاده می‌شود. روش‌هایی همچون رگرسیون، برآورد روند، سیستم‌های

۱. علت حذف متغیر درآمد از الگو، در دسترس نبودن تواتر ماهانه متغیر درآمد می‌باشد.

خبره و شبکه‌های عصبی مصنوعی در این دسته قرار می‌گیرند (رفعتی، ۱۳۹۰ و منتزور و کاهن^۱، ۱۹۹۵).

شبکه‌های عصبی مصنوعی که در حقیقت اقتباسی از فرایند یادگیری مغز انسان (هوش طبیعی) می‌باشند، می‌توانند روابط بین متغیرها را هرچند که پیچیده باشند یاد بگیرند و از آن برای پیش‌بینی مقادیر آتی استفاده نمایند. مزیت مهم شبکه‌های عصبی نسبت به سایر مدل‌های ساختاری و سری زمانی، عدم نیاز به اعمال فرض‌های آماری خاص در مورد رفتار متغیرها مانند فرض‌های مربوط به تابع توزیع احتمال آن‌ها می‌باشد (مشیری، ۱۳۸۰). بر این اساس در این پژوهش از شبکه‌های عصبی مصنوعی به منظور مقایسه با مدل پر کاربرد ARIMA، در پیش‌بینی مصرف گاز طبیعی استفاده شد. در ادامه، ضمن توضیح مختصر شبکه‌های عصبی مصنوعی، الگوریتم ژنتیک و الگوریتم ازدحام ذرات، به بیان نتایج تجربی پرداخته شده است.

۳-۲-۱. شبکه‌های عصبی مصنوعی

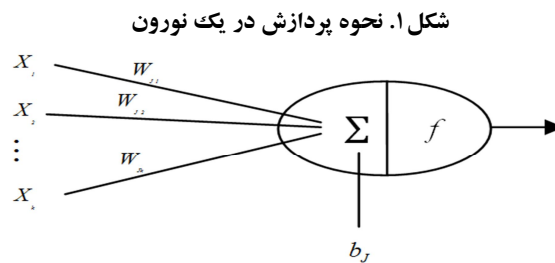
شبکه‌های عصبی مصنوعی دسته‌ای از مدل‌های غیرخطی انعطاف‌پذیر هستند که می‌توانند الگوهای اقتباس‌شده از داده‌ها را کشف کنند (آیدین^۲، ۲۰۱۴). این شبکه‌ها با وجود تنوعشان، از ساختار کلی مشابهی برخوردارند. هر شبکه عصبی به طور معمول از سه نوع لایه تشکیل می‌شوند: لایه ورودی^۳، لایه پنهان^۴ و لایه خروجی^۵. (مشیری، ۱۳۸۰). وجود لایه پنهان، قابلیت ارائه و محاسبه ارتباطات پیچیده بین الگوها را به شبکه می‌دهد (کانکال و همکاران^۶، ۲۰۱۱). در این لایه‌ها یک یا چند نورون^۷ قرار دارد که کار پردازش بر روی داده‌ها را انجام می‌دهند. هر نورون یک سیگنال ورودی را از سایر نورون‌ها و یا

1. Mentzer & Kahn
2. Aydin
3. Input Layer
4. Hidden Layer
5. Output Layer
6. Kankal & et al
7. Neuron

محرك‌های خارجی دريافت می‌کند، سپس آن را از طريق يك تابع فعال‌سازی^۱ و یا انتقال پردازش کرده و يك سیگنال خروجی را برای سایر نورون‌ها و یا خروجی تولید می‌کند (سولدو^۲، ۲۰۱۲). نحوه پردازش دريك نورون بدین صورت است که يك بردار ضرایب را در بردار ورودی خود ضرب کرده و حاصل را با يك جمله بایاس^۳ جمع می‌کند. در نهایت نیز نتیجه فرایند ذکر شده را به عنوان ورودی برای تابع فعال‌سازی در نظر گرفته و خروجی تابع به عنوان خروجی نورون به دست می‌آید. تابع فعال‌سازی بر اساس نیاز خاص مسئله‌ای که قرار است به وسیله شبکه عصبی حل شود، می‌تواند خطی یا غیر خطی انتخاب می‌شود (انوار، ۱۳۸۱). در شکل ۱ نحوه فعالیت يك نورون مصنوعی به تصویر کشیده شده است. خروجی يك نورون را می‌توان با فرمول زیر محاسبه کرد (رن و همکاران^۴، ۲۰۱۴):

$$\hat{Y}_j = f(Net), Net = \sum_{i=1}^k w_{ji} * X_{ji} + b_j \quad (5)$$

در فرمول شماره (۵)، X بردار ورودی نورون j ام، Y بردار خروجی، W بردار ضرایب و b جمله بایاس آن می‌باشد.

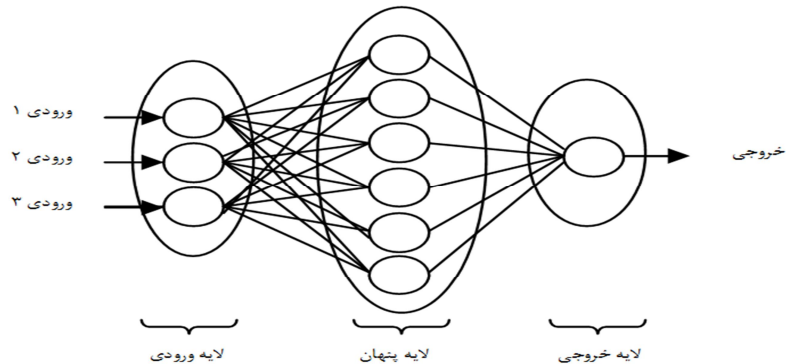


لایه ورودی شامل نورون‌هایی به تعداد متغیرهای توضیحی مدل می‌باشد، لایه‌های پنهان نیز شامل تعدادی نورون هستند که در آن‌ها عملیات جبری ذکر شده در فرمول (۵) بر روی اطلاعات ورودی که خروجی نورون‌های لایه‌های قبلی است، صورت گرفته و نتیجه

-
1. Activity Function
 - 2 . Soldo
 3. Bias
 4. Ren & et al

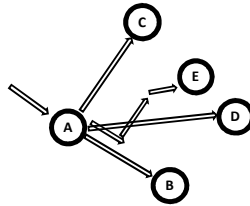
آن به صورت یک ورودی جدید به نورون‌های دیگر در لایه خروجی ارسال می‌گردد. تعداد نورون‌های لایه خروجی نیز، برابر با تعداد متغیرهای وابسته در نظر گرفته می‌شود. (مشیری، ۱۳۸۰). ارتباط بین لایه‌ها در شبکه عصبی می‌تواند کامل یا ناقص انتخاب شود. علاوه بر آن، جهت این ارتباطات نیز بسته به نوع شبکه می‌تواند روبه جلو (پیش‌خور^۱) و یا رو به عقب باشد. در شکل (۲) یک شبکه عصبی پیش‌خور با ۶ نورون در لایه پنهان و یک نورون در لایه خروجی نمایش داده شده است.

شکل ۲. شمای کلی یک شبکه عصبی با یک لایه پنهان



در ادبیات شبکه‌های عصبی به جای عبارت تخمین، از عبارت آموزش استفاده می‌گردد (قدیمی، مشیری ۱۳۸۱). منظور از آموزش شبکه نیز، فرایند تعیین بردار وزن‌ها (ضرایب) در نورون‌ها می‌باشد. در واقع وزن‌های شبکه از طریق فرایند آموزش تعدیل می‌شوند (کانکال و همکاران، ۲۰۱۱). الگوریتم‌های مختلفی برای آموزش شبکه مورد استفاده قرار می‌گیرد. از آنجا که هدف حداقل کردن خطای پیش‌بینی شبکه می‌باشد و ممکن است مینیمم‌های محلی زیادی در این مسئله وجود داشته باشند، بایستی فرایند آموزش طوری انتخاب شود که امکان قرار گرفتن در مینیمم محلی را کاهش داد. در این پژوهش از الگوریتم‌های ژنتیک و ازدحام ذرات جهت آموزش شبکه استفاده شده و نتایج استفاده از آن‌ها با یکدیگر مقایسه می‌شود.

۲-۲-۳. الگوریتم ازدحام ذرات



شکل ۳. نحوه تغییر مکان ذرات در الگوریتم PSO

الگوریتم بهینه‌سازی ذرات (PSO)، یکی از تکنیک‌های بهینه‌یابی تصادفی، مبتنی بر جمعیت^۱ است که اولین بار توسط کندی و ابرهت^۲، با الهام گرفتن از پرواز همزمان پرندگان و رفتار حشرات، در سال ۱۹۹۵ معرفی شد (کران و همکاران^۳، ۲۰۱۲). در این الگوریتم همانند تکنیک‌های محاسبات تکاملی، چون الگوریتم ژنتیک کار با یک جمعیت اولیه تصادفی که جواب‌های بالقوه را تشکیل می‌دهند آغاز شده و سپس جمعیت به‌روز می‌شود. در این الگوریتم به هر جواب یک ذره ۴ گفته می‌شود. از ویژگی‌های این الگوریتم می‌توان به اشتراک‌گذاری اطلاعات توسط ذرات اشاره کرد. ذرات در هر بار تکرار الگوریتم، بر اساس اطلاعات موقعیت فعلی خود، بهترین موقعیتی که تا کنون در آن قرار داشته‌اند (کمترین خطا در پیش‌بینی را داشته‌اند) و بهترین موقعیت کل ذرات، تغییر یافته و در موقعیت جدید قرار می‌گیرند. در شکل (۳)، نحوه تغییر مکان ذره A نشان داده شده است. اگر فرض کنیم یک ذره در مکان A قرار دارد به نوعی که جهت بردار سرعت قبلی آن هم راستا با بردار B باشد می‌توان سه حالت حدی برای حرکت بعدی ذره در نظر گرفت. در حالت اول؛ ذره همان جهت حرکت قبلی خود را حفظ کرده و در همان راستا حرکت کند، در این صورت به نقطه B خواهد رسید. در حالت دوم؛ ذره در جهت بهترین مکانی که تا کنون تجربه کرده است

1. Population
2. James Kennedy and Russel Ebbhart
3. Kiran & et al
4. Particle

(کمترین میزان تابع هزینه ۱ یا بیش‌ترین میزان تابع برازندگی ۲ را داشته است) حرکت کند و به نقطه C برود. حالت سوم؛ حرکت در جهت مکان بهترین ذره و قرار گرفتن در نقطه D می‌باشد. حالت کلی حرکت در الگوریتم PSO و تعیین موقعیت جدید هر ذره را می‌توان بر اساس معادلات (۶) و (۷) محاسبه نمود (کران و همکاران، ۲۰۱۲):

$$V_i[t+1] = w * V_i[t] + C_1 * r_1 (pbest_i[t] - P_i[t]) + C_2 * r_2 (gbest[t] - P_i[t]) \quad (6)$$

$$P_i[t+1] = P_i[t] + V_i[t+1] \quad (7)$$

که در آن‌ها، V_i معرف سرعت ذره نام، t نماینده زمان، $pbest$ بهترین مقداری که ذره تا زمان t به آن رسیده است، $gbest$ بهترین جوابی که تا زمان t در گروه $pbest$ ها به دست آمده، P_i مکان ذره نام، w, C_1, C_2 و فاکتور وزنی و r_1 و r_2 اعداد تصادفی هستند که از بازه‌ی $[0,1]$ انتخاب می‌شوند. بنابراین می‌توان انتظار داشت ذره نام در زمان یا تکرار $t+1$ در مکان E قرار گیرد. بدین ترتیب همه‌ی ذرات به امید یافتن بهینه‌ترین نقطه مکان خود را تغییر خواهند داد. در این پژوهش این فرایند به منظور تعیین وزن‌های شبکه استفاده می‌شود؛ بدین صورت که هر جواب یا ذره، ماتریسی سطری به ابعاد تعداد وزن‌های شبکه عصبی بوده و هر درایه آن نماینده یکی از وزن‌ها می‌باشد.

۳-۲-۳. الگوریتم ژنتیک

الگوریتم ژنتیک در دهه ۷۰ میلادی، توسط جان هولاند^۳ مطرح گردید. این الگوریتم با الهام از فرایند انتخاب طبیعی و نظریه تکامل تدریجی داروین^۴، از فرایندهای زیست‌شناسی وراثت و جهش بهره می‌گیرد (کانی‌اورت و اوزتورک^۵، ۲۰۰۸). در این الگوریتم، کار با مجموعه‌ای از جواب‌های بالقوه که اصطلاحاً کروموزوم^۶ نامیده می‌شوند،

-
1. Cost Function
 2. Fitness Function
 3. John Holland
 4. Charles Robert Darwin
 5. Canyurt & Ozturk
 6. Chromosome

آغاز می‌شود. کل مجموعه کروموزوم‌ها یا هر بار تکرار الگوریتم، یک جمعیت یا نسل نامیده می‌شود. هر کروموزوم از تعدادی ژن^۱ تشکیل شده و هر ژن نماینده خصوصیتی از جواب مسئله (وزن شبکه عصبی) است.

در طبیعت، انتقال خصوصیات از نسلی به نسل دیگر از طریق کروموزوم‌ها و ژن‌ها صورت می‌گیرد. الگوریتم ژنتیک با این دیدگاه که در نهایت ژن‌ها و کروموزوم‌های بهتر و قوی‌تر باقی می‌مانند و ژن‌های ضعیف در طی زمان از بین می‌روند، یا همان قانون تکامل داروین که موجودات ضعیف‌تر از بین می‌روند و موجودات سازگارتر باقی خواهند ماند، به مدل‌سازی می‌پردازد. در این الگوریتم، ابتدا جمعیتی از جواب‌های بالقوه به صورت تصادفی ایجاد شده سپس برازندگی با استفاده از تابع هدف مسئله (حداقل سازی خطای پیش‌بینی) محاسبه و از میان آن‌ها تعدادی با توجه به میزان برازندگی آن‌ها جهت انجام عمل ترکیب^۲ و جهش^۳ انتخاب می‌شوند. عمل ترکیب به صورت انتخاب دو والد^۴ یا جواب و سپس ترکیب خصوصیات یا ژن‌ها به منظور تولید فرزندان^۵ یا جواب‌های جدید صورت می‌پذیرد. عملگر جهش نیز بدین صورت است که از میان خصوصیات یا ژن‌های انتخاب شده، یک یا چند ژن (وزن) به صورت تصادفی تغییر می‌کنند. در نهایت پس از انجام عمل ترکیب و جهش، جمعیت جدید یا نسل بعد تولید می‌شود. این فرایند به منظور یافتن بهترین جواب ادامه می‌یابد تا شرط توقف الگوریتم تأمین شود، در این پژوهش رسیدن به تعداد مشخصی نسل، شرط توقف در هر دو الگوریتم ژنتیک و الگوریتم ازدحام ذرات می‌باشد.

۳-۲-۴. مدل خود رگرسیو میانگین متحرک انباشته

مدل ARIMA که توسط باکس و جنکینز^۶ معرفی گردید، یکی از رایج‌ترین مدل‌ها در پیش‌بینی سری‌های زمانی می‌باشد (پای و لین^۷، ۲۰۰۵). در این مدل مقادیر آتی متغیر

-
1. Gene
 2. Crossover
 3. Mutation
 4. Parents
 5. Offspring
 6. Box & Jenkins
 7. Pai & Lin

به صورت ترکیب خطی از مقادیر گذشته آن و خطاهای گذشته در نظر گرفته می‌شود. مدل $ARIMA(p,d,q)$ را با فرض درجه انباشتگی صفر ($d=0$) می‌توان به صورت کلی زیر بیان کرد:

$$y_t = \theta_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (۸)$$

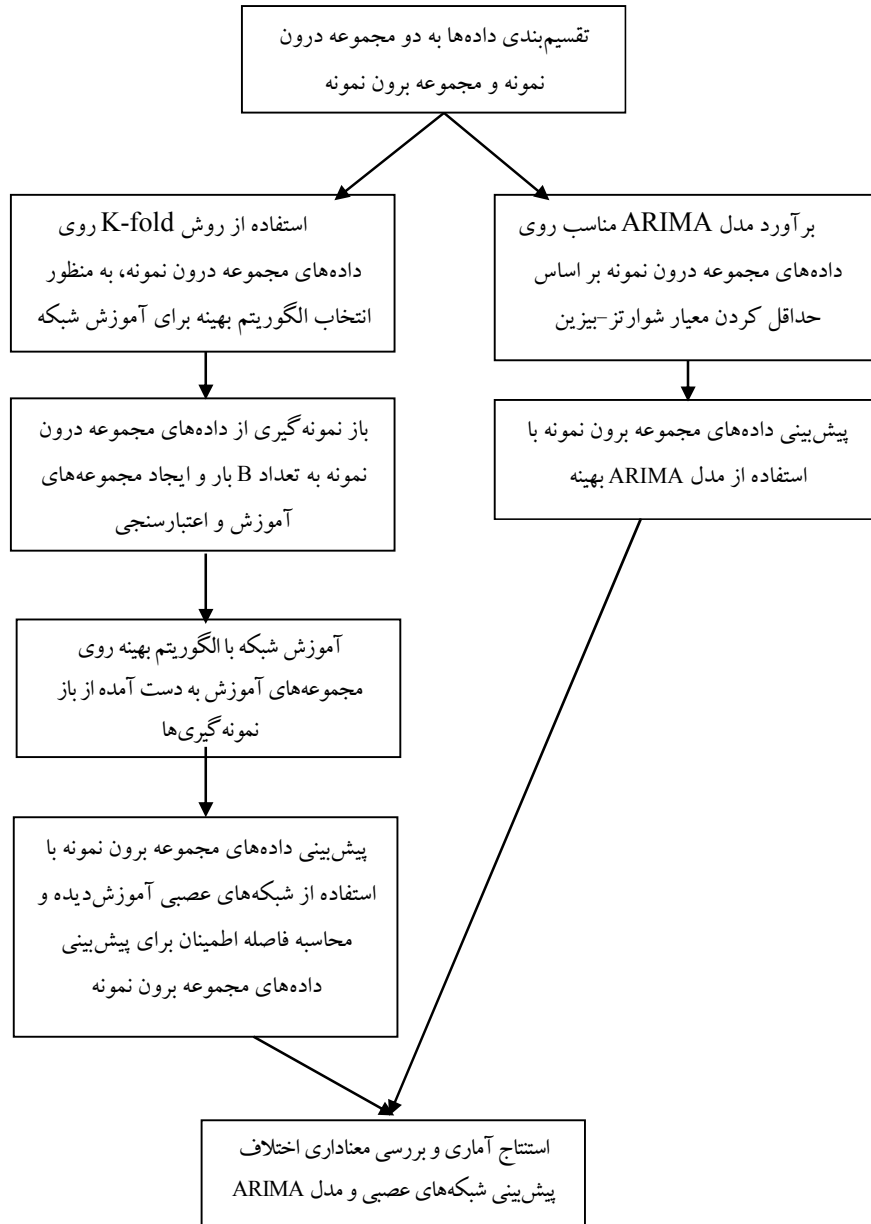
در فرمول (۸)، y_t مقدار واقعی متغیر y در زمان t می‌باشد و ε_t خطا در زمان t است؛ φ_i و θ_j ضرایب ثابت، p و q نیز به ترتیب تعداد جملات خود رگرسیو و تعداد جملات میانگین متحرک است. این روش به طور کلی دارای سه مرحله شناسایی، برآورد و تشخیص دقت می‌باشد. تعداد جملات خود رگرسیو و میانگین متحرک معمولاً با استفاده توابع خود همبستگی^۱ (AC) و خود همبستگی جزئی^۲ (PAC) بر اساس مراحل سه‌گانه ذکر شده محاسبه می‌شود اما از آنجا که ممکن است مدل‌های بهینه دیگری نیز وجود داشته باشند، انتخاب مدل نهایی بر اساس معیارهایی همچون شوارتز-بیزین^۳ صورت می‌پذیرد (صادقی و دیگران، ۱۳۹۰). d را می‌توان تعداد درجات لازم برای تفاضل‌گیری جهت مانا شدن سری در نظر گرفت. در صورتی که سری مورد استفاده نا مانا باشد، می‌توان با d بار تفاضل‌گیری آن را مانا کرد و پس از آن مقدار به دست آمده را به جای y در فرمول (۸) قرار داد. مدل $ARIMA$ اساساً یک مدل داده محور است که از ساختار خود داده‌ها اقتباس می‌شود با این حال وجود هرگونه روابط غیرخطی معنی‌دار در داده‌ها این مدل را محدود می‌کند (پای و لین، ۲۰۰۵).

در شکل (۴)، نحوه ارتباط بین الگوریتم‌ها در این تحقیق و تشکیل فاصله اطمینان برای پیش‌بینی شبکه‌های عصبی مصنوعی به منظور بررسی معناداری اختلافات در پیش‌بینی شبکه عصبی با مدل خطی $ARIMA$ و مقایسه دو مدل، ترسیم شده است.

1. Auto Coloration
2. Partial Auto Coloration
3. Schwarz-Bayesian

شکل ۴. فرایند تشکیل فاصله اطمینان برای پیش‌بینی شبکه‌های عصبی مصنوعی و مقایسه نتایج با

مدل خطی ARIMA



مراحل کلی این پژوهش از تعیین داده‌ها تا تشکیل فاصله اطمینان برای پیش‌بینی شبکه عصبی و استنتاج آماری را می‌توان به صورت ۵ مرحله زیر خلاصه کرد:

۱. ابتدا داده‌ها به دو دسته داده‌های درون نمونه و برون نمونه تفکیک شده به نحوی که از داده‌های مجموعه درون نمونه به منظور برآورد مدل‌ها استفاده شده و از داده‌های برون نمونه به منظور آزمون نهایی و مقایسه مدل‌ها استفاده شده است.
۲. در این مرحله، الگوریتم مناسب برای آموزش شبکه عصبی با بهره‌گیری از روش K-fold انتخاب می‌گردد. بدین منظور، مجموعه درون نمونه به K دسته مساوی افزای شده و سپس هر کدام از این دسته‌ها دقیقاً یک‌بار به عنوان مجموعه آزمون^۱ لحاظ شده و سایر مشاهدات به عنوان مجموعه آموزش^۲ استفاده می‌شود. در ادامه شبکه عصبی را با استفاده از الگوریتم ژنتیک و الگوریتم ازدحام ذرات، روی مجموعه‌های آموزش، آموزش داده و از میانگین خطای مجموعه‌های آزمون به منظور انتخاب الگوریتم بهینه برای آموزش استفاده شده است.
۳. پس از انتخاب الگوریتم مناسب برای آموزش شبکه، جهت تشکیل فاصله اطمینان برای پیش‌بینی شبکه عصبی اقدام به باز نمونه‌گیری از داده‌های مجموعه درون نمونه به تعداد B بار و ایجاد مجموعه‌های آموزش و اعتبارسنجی جدید شده است.
۴. در این مرحله، شبکه عصبی روی مجموعه‌های آموزش به دست آمده از مرحله سوم و با استفاده از الگوریتم بهینه به دست آمده از مرحله دوم، آموزش دیده و با پیش‌بینی داده‌های مجموعه برون نمونه توسط شبکه‌های آموزش دیده، فاصله اطمینان برای پیش‌بینی شبکه عصبی به دست آمده است.
۵. در نهایت از فاصله اطمینان به دست آمده در مرحله چهارم، جهت استنتاج آماری و بررسی معناداری اختلاف پیش‌بینی شبکه‌های عصبی و مدل ARIMA استفاده شده است.

1. Test Set
2. Training Set

۴. نتایج تجربی

همان طور که در بخش مبانی نظری گفته شد، مصرف گاز دارای نوسانات فصلی می‌باشد. از این رو به منظور حذف اثرات فصلی بر روی مصرف گاز، از اختلاف مصرف فعلی با مصرف ۱۲ ماه قبل آن به عنوان متغیر وابسته (DNGC) استفاده شد و از وقفه‌های آن به همراه اختلاف دمای فعلی با دمای ۱۲ ماه قبل (DT)، اختلاف تعداد مشترکین فعلی با وقفه دوازدهم آن (DN) و تغییرات قیمتی نسبت به ماه مشابه سال قبل^۱ (DP) به عنوان متغیرهای مستقل (ورودی‌های شبکه عصبی) استفاده شدند. پس از پیش‌بینی نیز مقدار پیش‌بینی شده با وقفه دوازدهم آن جمع شد و نتیجه به عنوان پیش‌بینی ماه مورد نظر در نظر گرفته شد.

به منظور مقایسه مدل‌های پیش‌بینی مصرف گاز طبیعی، ابتدا یک شبکه عصبی پیش‌خور با یک لایه پنهان ایجاد شد. سپس با آزمون و خطا و تغییر تعداد لایه‌های پنهان و تعداد نورون‌های لایه پنهان، در نهایت یک شبکه عصبی پیش‌خور با یک لایه پنهان، تعداد ۱۱ نورون در این لایه، تابع فعال‌سازی تانژانت زیگموئید^۲ در لایه پنهان و تابع خطی^۳ در لایه خروجی انتخاب شد. داده‌های مربوط به سال ۱۳۸۹ و ۱۳۹۰ به عنوان داده‌های آزمون نهایی کنار گذاشته شد و در فرایند آموزش مورد استفاده قرار نگرفت. آموزش شبکه نیز با استفاده از الگوریتم‌های ازدحام ذرات و ژنتیک صورت گرفت. ابتدا پارامترهای دو الگوریتم با آزمون و خطا، انتخاب و سپس نتایج آموزش با دو الگوریتم به منظور مقایسه دقیق‌تر، با استفاده از تکنیک "k-Fold" باهم مقایسه شدند که نتایج آن در جدول (۱) گزارش شده است. در این پژوهش $k=10$ انتخاب گردید و بدین ترتیب مشاهدات درون‌نمونه^۴ (۱۱۳۸۰ تا ۱۳۸۸) به ۱۰ زیرمجموعه جدا افراز و هر کدام یک‌بار به عنوان

۱. ابتدا تغییرات قیمت واقعی گاز طبیعی هر ماه به صورت درصدی، نسبت به ماه قبل محاسبه شده و سپس تغییرات آن نسبت به ماه مشابه سال قبل به عنوان DP محاسبه شده است.

2. Tangent sigmoid activity function
3. Linear Function
4. In Sample

مجموعه آزمون و ۹ بار به عنوان مجموعه آموزش لحاظ گردید و در کل ده بار شبکه آموزش دید و میانگین پیش‌بینی‌ها به عنوان پیش‌بینی نهایی با هم مقایسه گردیدند.

جدول ۱. مقایسه نتایج به دست آمده از الگوریتم‌های ژنتیک و PSO

الگوریتم ژنتیک		PSO الگوریتم		بارآزموهای الگوریتم
تعداد ذرات	۱۲۰	جمعیت	۱۲۰	
تعداد تکرار	۷۰	تعداد نسل‌ها	۷۰	
کسر ترکیب	۰,۷	w	$0.73*(0.9)^{\text{iter}-1}$	
تعداد نخه انتقالی	۶	طول گام	۰,۲	
کسر مهاجرت	۰,۲	C1	۱,۵	
تابع ترکیب	crossoverscattered	C2	۱,۵	
تابع جهش	mutationadaptfeasible			
تابع انتخاب	selectionstochunif			
میانگین خطای آموزش	میانگین خطای آزمون	میانگین خطای آموزش	میانگین خطای آزمون	
۱۶۴,۲۷	۱۷۵,۸۷	۱۵۶,۱۴	۱۶۵,۴۲	

منبع: نتایج تحقیق

مطابق نتایج به دست آمده از جدول (۱) و بر اساس معیار مجذور مربعات خطا (RMSE) الگوریتم ازدحام ذرات عملکرد بهتری نسبت به الگوریتم ژنتیک هم در مجموعه آزمون و هم در مجموعه آموزش داشته است. با توجه به نتایج به دست آمده، در ادامه به منظور پیش‌بینی نهایی برای ۲۴ ماه سال‌های ۱۳۸۹ و ۱۳۹۰ (خارج از نمونه^۱)، از الگوریتم ازدحام ذرات با پارامترهای مشخص شده در مرحله قبل استفاده شد و ۹۶ مشاهده از فروردین ۱۳۸۱ تا اسفند ۱۳۸۸ به عنوان مجموعه درون نمونه و تعیین مدل به کار رفتند. این بار با هدف افزایش دقت در اندازه‌گیری خطا، با الگوبرداری از روش باز نمونه‌گیری بوت استرپ^۲ اقدام به ۲۰۰۰ بار باز نمونه‌گیری از داده‌های درون نمونه شد^۳، همچنین به

1. Out of Sample
2. Bootstrap

۳. برای مطالعه بیشتر پیرامون روش‌های "K-Fold" و "Bootstrap" رجوع کنید به Efron & Tibshirani (1994)

تحلیل آماری و برآورد فاصله اطمینان پیش‌بینی شبکه عصبی ... □ ۹۵

منظور بالا بردن دقت شبکه، از داده‌های اعتبارسنجی^۱ نیز استفاده شد؛ بدین صورت که هر بار اقدام به باز نمونه‌گیری با جایگذاری از داده‌های مجموعه درون نمونه به تعدادی معادل با ۸۵ درصد تعداد داده‌ها صورت گرفت و نمونه به دست آمده، در فرایند آموزش مورد استفاده قرار گرفت. داده‌های باقیمانده از مجموعه درون نمونه نیز به عنوان داده‌های اعتبارسنجی لحاظ شدند. در نهایت نیز با شبکه آموزش دیده، مقادیر مصرف برای سال‌های ۱۳۸۹ و ۱۳۹۰ پیش‌بینی گردید و فاصله اطمینان ۹۵ درصدی برای پیش‌بینی‌های مذکور محاسبه شد. در جدول (۲) نتایج مربوط به پیش‌بینی بهترین شبکه عصبی آموزش دیده که کمترین خطای درون نمونه را در بین ۲۰۰۰ شبکه داشت بر اساس معیارهای مجذور میانگین مربعات خطا (RMSE)، میانگین درصد خطا (MAPE^۲) و میانگین قدر مطلق خطا (MAE^۳) گزارش شده است.

جدول ۲. مقادیر پیش‌بینی شده توسط شبکه عصبی بهینه

شهریور	مرداد	تیر	خرداد	اردیبهشت	فروردین	۱۳۸۹
۱۱۸۷	۱۰۶۶	۱۱۳۰	۱۵۳۹	۲۴۷۸	۳۸۳۲	مقدار واقعی
۱۳۰۶	۱۲۷۷	۱۱۵۱	۱۴۱۳	۲۵۵۰	۴۳۰۷	پیش‌بینی
اسفند	بهمن	دی	آذر	آبان	مهر	۱۳۸۹
۴۷۷۸	۶۸۲۲	۶۸۴۵	۵۶۰۱	۳۶۳۸	۱۸۱۶	مقدار واقعی
۴۸۳۱	۶۸۱۵	۷۰۱۴	۵۴۸۳	۴۰۲۷	۱۸۱۸	پیش‌بینی
شهریور	مرداد	تیر	خرداد	اردیبهشت	فروردین	۱۳۹۰
۱۱۰۶	۹۴۳	۹۷۸	۱۲۵۰	۲۱۲۹	۳۹۰۲	مقدار واقعی
۱۱۴۸	۸۷۷	۹۰۱	۱۲۵۴	۲۱۶۹	۳۷۵۴	پیش‌بینی
اسفند	بهمن	دی	آذر	آبان	مهر	۱۳۹۰
۵۷۰۵	۷۳۱۱	۷۲۰۳	۶۶۳۷	۴۶۷۷	۲۰۴۷	مقدار واقعی
۵۴۰۹	۷۳۴۰	۷۱۲۶	۶۷۴۸	۴۵۰۰	۱۹۴۲	پیش‌بینی
خطای خارج از نمونه (۱:۱۳۸۹ تا ۱:۱۳۹۰)			خطای درون نمونه (۱:۱۳۸۰ تا ۱:۱۳۸۸)			
RMSE=۱۶۹,۲۹, MAPE=۴,۷۴, MAE=۱۲۲,۲۹			RMSE=۱۲۶,۱۵, MAPE=۳,۷۶, MAE=۹۱,۳۰			

منبع: نتایج تحقیق

1. Validation Set
2. Mean Absolute Percentage Error
3. Mean Absolute Error

در ادامه با هدف تعیین اهمیت هر یک از ورودی‌ها بر روی خروجی شبکه، از روش پارتیشن‌بندی وزنی گارسون^۱ (۱۹۹۱) استفاده شد. در این روش جهت تعیین اهمیت هر متغیر ورودی در پیش‌بینی شبکه عصبی، از معادله زیر استفاده می‌شود (فائور و کلورک^۲، ۲۰۰۳):

$$I_j = \frac{\sum_{m=1}^{N_h} \left(\left(\frac{|W_{jm}^{ih}|}{\sum_{k=1}^{N_i} |W_{km}^{ih}|} \right) \times |W_{mn}^{ho}| \right)}{\sum_{k=1}^{N_i} \left\{ \sum_{m=1}^{N_h} \left(\left(\frac{|W_{km}^{ih}|}{\sum_{k=1}^{N_i} |W_{km}^{ih}|} \right) \times |W_{mn}^{ho}| \right) \right\}} \quad (9)$$

که در آن I_j اهمیت نسبی پارامتر ورودی j ام برای خروجی است، ضمن اینکه N_i و N_h به ترتیب نشانگر تعداد ورودی و تعداد نورون در لایه مخفی هستند. زیروندهای i ، h و o به ترتیب نشان‌دهنده لایه‌های ورودی، پنهان و خروجی بوده و زیروندهای k ، m و n نشان‌دهنده نورون‌های ورودی، مخفی و خروجی هستند. W بیانگر وزن‌های بهینه ارتباطی بین لایه‌ها هستند که از شبکه بهینه در مرحله قبل به دست آمدند. نتایج این روش در جدول (۳) گزارش داده شده است.

جدول ۳. نتایج بررسی میزان اثرپذیری پیش‌بینی‌ها از متغیرهای ورودی

متغیر	دما	قیمت	تعداد مشترکین	وقفه اول	وقفه دوم
اهمیت	۲۲٫۸۴	۱۹٫۱۸	۱۷٫۳۰	۲۱٫۳۴	۱۹٫۳۴

منبع: نتایج تحقیق

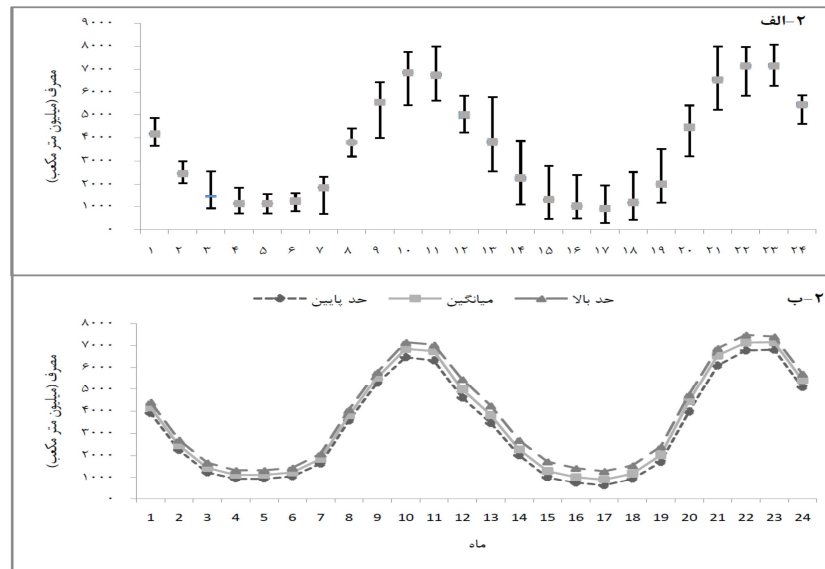
مطابق نتایج به دست آمده، دما بیشترین تأثیر و تعداد مشترکین کمترین تأثیر را بر مصرف گاز طبیعی در بخش خانگی داشته است. همچنین اثر قیمت نیز معادل ۱۹ درصد

1. Garson
2. Faur-Brasquet & Le Cloirec

بوده است. این نتایج مطابق با انتظارات می‌باشد به این دلیل که قیمت گاز طبیعی نسبتاً پایین است و جایگزین چندانی ندارد. بنابراین واکنش مصرف‌کنندگان به تغییرات قیمت کمتر از واکنش به متغیرهای تغییرات دمایی و وقفه‌های مصرفی می‌باشد. از طرفی با توجه به ماهانه بودن مشاهدات، و پایین بودن نسبت تعداد انشعابات جدید به کل انشعابات قبلی، سهم اثرگذاری تعداد مشترکین نیز کمتر از بقیه متغیرها است.

در نمودار (۲) نتایج مربوط به پیش‌بینی برون‌نمونه شبکه‌های عصبی، برای ماه‌های سال ۱۳۸۹ و ۱۳۹۰ به ترتیب از یک تا ۲۴ ترسیم شده است. نمودار (۲-الف) نمودار جعبه‌ای نتایج پیش‌بینی ۲۰۰۰ شبکه آموزش‌دیده روی ۲۰۰۰ بار باز نمونه‌گیری و نمودار (۲-ب) نیز میانگین مقادیر پیش‌بینی را به همراه فاصله اطمینان ۹۵ درصدی آن نشان می‌دهد.

نمودار ۱. پیش‌بینی مصرف گاز طبیعی در سال ۱۳۸۹



منبع: نتایج تحقیق

بر اساس نتایج به دست آمده از نمودار (۲)، با کاهش دما و افزایش مصرف گاز در ماه‌های سرد سال طول فاصله اطمینان بزرگ‌تر شده است. این در حالی است که درصد خطای پیش‌بینی در طول این ماه‌ها متفاوت بوده است. در ادامه ابتدا داده‌های مجموعه

برون‌نمونه با استفاده از مدل ARIMA پیش‌بینی شده است، سپس ۲۰۰۰ شبکه عصبی این بار تنها با وقفه‌های مصرف گاز طبیعی آموزش داده شده و در نهایت با ساختن فاصله اطمینان، نتایج با مدل ARIMA مقایسه شده است.

با توجه به این که مدل‌سازی ARIMA بر اساس متغیرهای مانا صورت می‌پذیرد، ابتدا مانایی سری تفاضل مرتبه ۱۲ مصرف گاز طبیعی با استفاده از آزمون‌های دیکی فولر تعمیم‌یافته (ADF)^۱، KPSS^۲ و فیلیپس پرون (PP)^۳ مورد بررسی قرار گرفته است. نتایج مربوط به این آزمون‌ها در جدول (۴) گزارش شده است.

جدول ۴. نتایج آزمون مانایی

متغیر	ADF		KPSS		PP	
	آماره آزمون	مقدار بحرانی	آماره آزمون	مقدار بحرانی	آماره آزمون	مقدار بحرانی
DNGC	-۳,۶۴	-۳,۴۵	۰,۰۶۲	۰,۱۴۶	-۳,۹۸	-۳,۴۵

منبع: نتایج تحقیق

فرض اولیه در آزمون‌های دیکی فولر و فیلیپس پرون، وجود ریشه واحد در متغیر است، اما در آزمون KPSS فرض اولیه عدم وجود ریشه واحد و مانایی است. مطابق نتایج به دست آمده از جدول (۴)، قدر مطلق آماره دیکی فولر و فیلیپس پرون بزرگ‌تر از مقادیر بحرانی متناظر خود در سطح خطای ۵ درصد بوده و در آزمون KPSS آماره آزمون کوچک‌تر از حد بحرانی خود است؛ بنابراین نتایج هر سه آزمون بیانگر مانایی سری تفاضل دوازدهم مصرف گاز طبیعی می‌باشد. پس از آزمون مانایی، برآورد مدل ARIMA نیز بر اساس حداقل کردن معیار شوارتز-بیزین صورت گرفت. با برآورد مدل‌های مختلف در مجموعه درون‌نمونه، نهایتاً مدل $ARIMA(0,0,2)$ به عنوان بهترین مدل (دارای کمترین آماره شوارتز-بیزین) جهت پیش‌بینی انتخاب گردید. نتایج مربوط به پیش‌بینی به صورت ایستا^۴ مدل ARIMA در جدول (۵) گزارش شده است. همچنین به منظور مقایسه این نتایج با مدل شبکه‌های عصبی مصنوعی، با استفاده از متغیرهای وقفه اول

1. Augmented Dickey-Fuller
2. Kwiatkowski-Philips-Schmidt-Shin
3. Philips-Perron
4. Static Forecast

تحلیل آماری و برآورد فاصله اطمینان پیش‌بینی شبکه عصبی ... □ ۹۹

و دوم مصرف گاز طبیعی، مجدداً به تعداد ۲۰۰۰ بار و به روش ذکر شده در آموزش شبکه-های نوع اول (وجود پنج متغیر در لایه ورودی) شبکه عصبی آموزش داده شد و پیش‌بینی مصرف در سال ۱۳۸۹ و ۱۳۹۰ انجام شد. مشخصات این شبکه‌ها مشابه با شبکه‌های نوع اول در نظر گرفته شد با این اختلاف که متغیرهای ورودی مشابه مدل ARIMA در نظر گرفته شدند و تعداد نورون‌های لایه پنهان نیز با آزمون و خطا برابر با ۶ در نظر گرفته شد. نتایج مربوط به پیش‌بینی بهینه‌ترین شبکه نیز بر اساس کمترین خطای درون‌نمونه، در جدول (۵) گزارش شده است.

جدول ۵. مقایسه نتایج حاصل از پیش‌بینی شبکه عصبی بهینه (۲) با مدل ARIMA

ANN ₂			ARIMA			مدل
RMSE	MAPE	MAE	RMSE	MAPE	MAE	معیار
۲۶۹,۲۳	۶,۵۶	۱۹۹,۹۳	۳۲۲,۷۶	۱۱,۴۴	۲۵۸,۵۳	پیش‌بینی خارج از نمونه
۱۸۱,۵۲	۳,۹۰	۱۱۱,۹۳	۱۸۴,۵۶	۴,۷۴	۱۱۹,۹۴	پیش‌بینی درون نمونه

منبع: نتایج تحقیق

با توجه به نتایج به دست آمده در جدول (۵)، مدل شبکه‌های عصبی نسبت مدل خطی ARIMA، هم در پیش‌بینی درون نمونه و هم در پیش‌بینی برون نمونه، خطای کمتر و در نتیجه عملکرد بهتری داشته است. به منظور بررسی دقیق‌تر، نتایج حاصل از پیش‌بینی به روش شبکه‌های عصبی مصنوعی و مدل ARIMA، از لحاظ آماری با یکدیگر مقایسه شدند تا روش قوی‌تر از لحاظ پیش‌بینی مشخص گردد. بدین منظور، نتایج مربوط به آماره استاندارد شده متوسط دو هزار پیش‌بینی نقطه‌ای شبکه‌های عصبی و مقدار پیش‌بینی توسط روش ARIMA در جدول (۶) ارائه شده است.

جدول ۶. نتایج مربوط به آزمون معنادار بودن اختلاف پیش‌بینی‌ها و مقادیر واقعی با یکدیگر

شهریور	مرداد	تیر	خرداد	اردیبهشت	فروردین	۱۳۸۹
۰,۸۴	۰,۶۷	-۰,۸۶	۳,۶۸	-۲,۱۹	۱,۰۳	$ARIMA - ANN_1$
-۰,۶۱	-۰,۶۷	-۰,۰۲	۰,۸۱	۰,۰۹	-۲,۷۷	$Actual - ANN_1$
۰,۰۵	۰,۸۸	-۲,۲۸	۲,۴۷	-۱,۳۵	-۱,۳۹	$ARIMA - ANN_2$
-۱,۷۵	-۰,۷۵	-۱,۴۹	-۰,۲۲	۰,۵۰	-۵,۶۶	$Actual - ANN_2$
اسفند	بهمن	دی	آذر	آبان	مهر	۱۳۸۹
۰,۳۶	۰,۶۷	-۲,۴۳	۰,۵۴	۰,۰۵	۰,۸۷	$ARIMA - ANN_1$
-۰,۹۰	۰,۳۹	-۰,۰۶	۰,۳۸	-۱,۳۹	-۰,۴۰	$Actual - ANN_1$
-۱,۳۴	-۰,۱۰	۰,۳۶	-۰,۴۰	۰,۴۳	-۰,۳۱	$ARIMA - ANN_2$
-۳,۷۲	-۰,۳۹	۴,۲۹	-۰,۵۹	-۱,۳۲	-۱,۹۱	$Actual - ANN_2$
شهریور	مرداد	تیر	خرداد	اردیبهشت	فروردین	۱۳۹۰
۰,۴۰	-۰,۲۰	۲,۳۹	-۲,۵۷	۳,۴۱	-۱,۰۶	$ARIMA - ANN_1$
-۰,۴۸	۰,۱۶	-۰,۲۶	-۰,۳۲	-۰,۸۱	۰,۳۹	$Actual - ANN_1$
۰,۴۴	-۲,۱۵	۲,۵۸	-۲,۷۷	۴,۵۹	-۰,۸۱	$ARIMA - ANN_2$
-۰,۹۵	-۱,۵۴	-۰,۹۲	-۰,۱۱	-۶,۷۵	۱,۰۷	$Actual - ANN_2$
اسفند	بهمن	دی	آذر	آبان	مهر	۱۳۹۰
۰,۲۵	-۱,۲۶	۰,۲۷	۱,۳۰	-۲,۳۲	-۰,۸۵	$ARIMA - ANN_1$
۱,۶۹	۰,۹۶	۰,۳۴	۰,۴۲	۱,۱۰	۰,۱۶	$Actual - ANN_1$
۲,۰۹	-۰,۵۲	-۱,۰۴	۱,۶۲	-۰,۷۶	۰,۳۰	$ARIMA - ANN_2$
۴,۳۷	۱,۳۱	-۰,۹۸	۰,۶۹	۵,۳۷	۲,۰۹	$Actual - ANN_2$

منبع: نتایج تحقیق

در جدول (۶) مقادیر مربوط به ردیف ARIMA-ANN مقدار آماره محاسبه شده برای بررسی معنی‌داری اختلاف نتایج پیش‌بینی مدل ARIMA با میانگین پیش‌بینی شبکه‌های عصبی را برای ماه‌های سال‌های ۱۳۸۹ و ۱۳۹۰ نشان می‌دهند؛ ردیف Actual-ANN نیز مقدار این آماره به منظور بررسی اختلاف میانگین پیش‌بینی شبکه‌های عصبی با مقادیر واقعی بیان می‌کنند. همچنین زیر وندهای یک و دو نیز به ترتیب مشخص‌کننده شبکه عصبی نوع اول با متغیرهای ورودی دما، تعداد مشترکین، قیمت گاز طبیعی و وقفه‌های اول و دوم مصرف گاز و شبکه عصبی نوع دوم تنها با ورودی وقفه‌های اول و

دوم مصرف گاز می‌باشد. بر اساس آماره‌های به دست آمده از جدول (۶) و مقایسه آن‌ها با مقادیر متناظر در جدول t در سطح خطای پنج درصد نتایج زیر به دست آمد:

پیش‌بینی شبکه‌های عصبی نوع اول، تنها در فروردین ماه ۱۳۸۹ با مقدار واقعی اختلاف معناداری داشته، اما در سایر ماه‌ها اختلاف معناداری با هم ندارند. همچنین نتایج پیش‌بینی این شبکه‌ها با پیش‌بینی مدل ARIMA در ۷ ماه باهم اختلاف معناداری داشته‌اند که در این مدت پیش‌بینی‌های شبکه‌های عصبی به مقدار واقعی نزدیک‌تر بوده است. بین میانگین پیش‌بینی شبکه‌های عصبی نوع دوم و مقادیر واقعی، در ۱۷ ماه از ۲۴ ماه پیش‌بینی شده، اختلاف معناداری نیست و فرض اولیه مبنی بر تساوی این دو رد نمی‌شود. همچنین بین مقادیر پیش‌بینی شده این شبکه‌ها و پیش‌بینی مدل ARIMA تنها در هفت ماه اختلاف معناداری وجود دارد که در دو ماه آن پیش‌بینی شبکه عصبی با مقدار واقعی اختلاف معناداری داشت و در اسفندماه ۱۳۹۰ نیز پیش‌بینی مدل ARIMA به مقدار واقعی نزدیک‌تر بود؛ اما در سایر ماه‌ها عموماً شبکه عصبی بهتر پیش‌بینی کرده است. مطابق این نتایج، شبکه‌های عصبی مصنوعی به عنوان الگویی غیرخطی، نسبت به مدل خطی ARIMA بهتر عمل کرده است و خطای کمتری دارد.

۵. جمع‌بندی و نتیجه‌گیری

امروزه، انرژی نقش ویژه‌ای در بهبود کیفیت زندگی، توسعه و رفاه اقتصادی جوامع به عهده دارد. این امر در مورد ایران که از ذخایر عظیم انرژی بهره می‌برد، اهمیت خاصی یافته است. ایران با داشتن حدود ۳۴ میلیون متر مکعب ذخایر گازی و سهمی معادل ۱۸ درصد کل ذخایر گازی جهان، رتبه نخست این ذخایر را در سال ۲۰۱۲ داشته است (بر اساس آمار منتشرشده توسط بریتیش پترولیوم (BP^۱) در سال ۲۰۱۲) که این امر لزوم استفاده بهینه و سیاست‌گذاری مناسب در این حوزه را افزایش داده است. بدون شک هرگونه برنامه‌ریزی در این حوزه نیازمند آگاهی از وضعیت مصرف کنونی و روند آتی آن

می‌باشد. نیاز به پیش‌بینی روند آتی منجر به کاربرد مدل‌ها و روش‌های مختلف با هدف پیش‌بینی دقیق‌تر شده است. در دهه‌های اخیر نیز به موازات مدل‌های سنتی پیش‌بینی مانند مدل‌های سری زمانی، روش‌های فرا‌ابتکاری مانند شبکه‌های عصبی مصنوعی نیز به طور گسترده‌ای به کاررفته‌اند (مشیری ۱۳۸۰) و در مطالعات بسیاری به مقایسه آن‌ها با مدل‌های دیگر همچون ARIMA پرداخته شده است؛ اما فرایند جستجوی تصادفی در این شبکه‌ها منجر به بروز نوعی نا اطمینانی در نتایج آن‌ها شده است. از طرفی شناخت مدل بهتر جهت پیش‌بینی، به سبب تأثیرگذاری که در اجرای سیاست‌ها و میزان موفقیت آن‌ها دارند، حائز اهمیت بسیاری می‌باشد. در این راستا و با توجه به اهمیت گاز طبیعی در تأمین انرژی، در این پژوهش، ضمن ترکیب شبکه‌های عصبی مصنوعی با الگوریتم ژنتیک و الگوریتم ازدحام ذرات به منظور بررسی و مقایسه دقت آن‌ها در پیش‌بینی مصرف ماهانه گاز طبیعی در بخش خانگی، با باز نمونه‌گیری از داده‌ها، یک فاصله اطمینان برای پیش‌بینی شبکه‌های عصبی آموزش دیده با الگوریتم بهینه استخراج گردید تا مقایسه دقیق‌تر بین شبکه‌های عصبی و مدل مرسوم ARIMA در پیش‌بینی مصرف گاز صورت پذیرفته و روش بهتر معرفی شود.

در این مطالعه، ابتدا شبکه‌ای با ورودی وقفه‌های یک و دوی مصرف گاز طبیعی و متغیرهای دیگر شامل؛ قیمت گاز طبیعی، دما و تعداد مشترکین در نظر گرفته شد و آموزش شبکه با هر دو الگوریتم صورت پذیرفت. طبق نتایج به دست آمده بر اساس روش "10-fold"، الگوریتم ازدحام ذرات بهتر و سریع‌تر از الگوریتم ژنتیک توانست شبکه را آموزش دهد. همچنین نتایج مربوط به تحلیل حساسیت نیز حاکی از تأثیرپذیری بالای مصرف از تغییرات دمایی بود. در ادامه با استفاده از الگوریتم ازدحام ذرات، مجدداً شبکه عصبی با ورودی وقفه‌های اول و دم مصرف گاز، مشابه با مدل ARIMA به تعداد ۲۰۰۰ بار روی داده‌های قبل از ۱۳۸۹ به صورت باز نمونه‌گیری با جایگذاری آموزش داده شد و فاصله اطمینان ۹۵ درصدی برای پیش‌بینی‌ها (ماه‌های سال ۱۳۸۹ و ۱۳۹۰) محاسبه شد. بر مبنای فاصله اطمینان به دست آمده، مدل ترکیبی شبکه عصبی با الگوریتم ازدحام ذرات از لحاظ آماری با مدل ARIMA مقایسه گردید. مطابق نتایج به دست آمده، می‌توان گفت

در حالت کلی مدل غیرخطی شبکه‌های عصبی نسبت به مدل خطی ARIMA، روند موجود در مصرف گاز طبیعی بخش خانگی را بهتر مدل‌سازی کرده، در نتیجه دقت بالاتری دارند. همچنین لحاظ متغیرهای تأثیرگذار تغییرات دمایی، قیمتی و تعداد مشترکین می‌تواند نتایج پیش‌بینی را به طور معنی‌داری بهبود بخشد. از این رو توصیه می‌شود دولت و سایر نهادهای مسئول در پیش‌بینی مصرف گاز طبیعی از الگوهای غیرخطی همچون شبکه عصبی استفاده کنند. ضمن این که اثرپذیری کمتر نتایج پیش‌بینی از قیمت گاز طبیعی نسبت به متغیر دما و وقفه‌های مصرف، بیانگر ضرورت توجه به سیاست‌های بهینه‌سازی مصرف انرژی به منظور افزایش کارایی در مصرف می‌باشد. بنابراین اصلاح قیمت‌های گاز طبیعی به تنهایی کافی نبوده و اجرای سیاست‌هایی چون ایجاد استانداردهای کارآمد در حوزه مصرف انرژی بخش مسکن توصیه می‌گردد.

منابع و مآخذ

- امامی میبدی، علی، محمدی، تیمور و سلطان‌العلمایی، سید محمدهادی. (۱۳۸۹). تخمین تابع تقاضای داخلی گاز طبیعی به روش فیلتر کالمن (مطالعه موردی تقاضای بخش خانگی شهر تهران). فصلنامه اقتصاد مقداری. شماره ۳. صص ۴۱-۲۳.
- انوار، سید حسام‌الدین. (۱۳۸۱)؛ پیش‌بینی نرخ ارز با استفاده از شبکه‌های عصبی؛ پایان‌نامه کارشناسی ارشد، دانشگاه تربیت مدرس.
- ترازنامه انرژی کشور: دفتر برنامه‌ریزی کلان‌برق و انرژی، وزارت نیرو، سال‌های مختلف (www.pep.moe.gov.ir).
- ترازنامه هیدروکربنی کشور: وزارت نفت، (۱۳۸۷) (www.mop.ir).
- رفعتی، محسن، آذرینفر، یداله و محمدزاده، رویا. (۱۳۹۰). انتخاب الگوی مناسب پیش‌بینی سطح زیرکشت، تولید و قیمت چغندر قند در ایران. نشریه اقتصاد و توسعه کشاورزی (علوم و صنایع کشاورزی). شماره ۲. صص ۱۶۰-۱۴۹.
- سالنامه آماری کشور: مرکز آمار ایران، سال‌های مختلف (www.amar.org.ir).
- شاکری، عباس، محمدی، تیمور، جهانگرد، اسفندیار و موسوی، میرحسین. (۱۳۸۹). تخمین مدل ساختاری تقاضای بنزین و نفت گاز در بخش حمل و نقل ایران. فصل‌نامه مطالعات اقتصاد انرژی، سال هفتم، شماره ۲۵، صص ۳۱-۱.
- شرکت ملی گاز ایران: مدیریت گاز رسانی، (www.nigc.ir).
- صادقی، حسین، ذوالفقاری، مهدی و الهامی نژاد، مهدی. (۱۳۹۰). مقایسه عملکرد شبکه‌های عصبی و مدل ARIMA در مدل‌سازی و پیش‌بینی کوتاه مدت قیمت سبب نفت خام اوپک (با تأکید بر انتظارات تطبیقی). فصلنامه مطالعات اقتصاد انرژی، سال هشتم، شماره ۲۸، صص ۴۷-۲۵.
- عباسی نژاد، حسین، و صادقی، حسین. (۱۳۷۸). پایداری رابطه تقاضای انرژی با قیمت‌های انرژی و سطح فعالیت‌های اقتصادی در اقتصاد ایران. مجله تحقیقات اقتصادی، شماره ۵۴، صص ۵۱-۳۳.
- قدیمی، محمدرضا، و مشیری، سعید. (۱۳۸۱). مدل‌سازی و پیش‌بینی رشد اقتصادی در ایران با استفاده از شبکه‌های عصبی مصنوعی (ANN). فصلنامه پژوهش‌های اقتصادی ایران. شماره

۱۲. صص ۹۷-۱۲۶

▪ کریمی، تورج، صادقی مقدم، محمدرضا و رهنما، روح‌الله. (۱۳۸۹). بررسی اثر تغییرات دما بر مصرف گاز طبیعی در ایران. فصلنامه مطالعات اقتصاد انرژی. شماره ۲۴. صص ۲۱۸-۱۹۳.

▪ کشاورز حداد، غلامرضا، و میرباقری جم، محمد. (۱۳۸۶). بررسی تابع تقاضای گاز طبیعی (خانگی و تجاری) در ایران. فصلنامه پژوهش‌های اقتصادی ایران. شماره ۳۲. صص ۱۶۰-۱۳۷.

▪ لطفعلی پور، محمدرضا، و باقری، احمد. (۱۳۸۲). تخمین تابع تقاضای گاز طبیعی مصارف خانگی شهر تهران. فصلنامه پژوهش‌های اقتصادی ایران. شماره ۱۶. صص ۱۵۱-۱۳۳.

▪ لیارد، پی. آر. جی؛ ا. ا. والترز. (۱۳۷۷). تئوری اقتصاد خرد. ترجمه عباس شاکری، تهران، نشر نی.

▪ مشیری، سعید. (۱۳۸۰). پیش‌بینی تورم با استفاده از مدل‌های ساختاری، سری‌های زمانی و شبکه‌های عصبی. فصلنامه تحقیقات اقتصادی. شماره ۵۸. صص ۱۸۴-۱۴۷.

▪ مشیری، سعید. و شاه‌مرادی، اصغر. (۱۳۸۵). برآورد تقاضای گاز طبیعی و برق خانوارهای کشور: مطالعه خرد مبتنی بر بودجه خانوار. فصلنامه تحقیقات اقتصادی. شماره ۷۲. صص ۳۳۶-۳۰۵.

- Aras, H. Aras, N. (2004). Forecasting residential natural gas demand. *Energy Sources*, 26(5), 463-472.
- Aydin, G. (2014). Modeling of energy consumption based on economic and demographic factors: The case of Turkey with projections. *Renewable and Sustainable Energy Reviews*, 35, 382-389.
- Balestra, P. Nerlove, M. (1966). Pooling cross section and time series data in the estimation of a dynamic model: The demand for natural gas. *Econometrica: Journal of the Econometric Society*, 585-612.
- Canyurt, O. E. & Ozturk, H. K. (2008). Application of genetic algorithm (GA) technique on demand estimation of fossil fuels in Turkey. *Energy Policy*, 36(7), 2562-2569.
- Ceylan, H. Ozturk, H. K. (2004). Estimating energy demand of Turkey based on economic indicators using genetic algorithm approach. *Energy Conversion and Management*, 45(15), 2525-2537.
- Efron, B. Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. (Chapman & Hall/CRC Monographs on Statistics & Applied Probability).
- Erdogdu, E. (2010). Natural gas demand in Turkey. *Applied Energy*, 87(1), 211-219.
- Faur-Brasquet, C. & Le Cloirec, P. (2003). Modelling of the flow behavior of activated carbon cloths using a neural network approach. *Chemical Engineering and Processing: Process Intensification*, 42(8), 645-652.

- Garson, G. D. (1991). Interpreting neural-network connection weights. *AI expert*, 6(4), 46-51.
- Herbert, J. H. Kreil, E. (1989). Specifying and evaluating aggregate monthly natural-gas demand by households. *Applied Economics*, 21(10), 1369-1382.
- Kankal, M. Akpınar, A. Kömürcü, M. İ. & Özşahin, T. Ş. (2011). Modeling and forecasting of Turkey's energy consumption using socio-economic and demographic variables. *Applied Energy*, 88(5), 1927-1939.
- Kıran, M. S. Özceylan, E. Gündüz, M. & Paksoy, T. (2012). A novel hybrid approach based on particle swarm optimization and ant colony algorithm to forecast energy demand of Turkey. *Energy conversion and management*, 53(1), 75-83.
- Liu, L. M. Lin, M. W. (1991). Forecasting residential consumption of natural gas using monthly and quarterly time series. *International Journal of Forecasting*, 7(1), 3-16.
- Mentzer, J. T. Kahn, K. B. (1995). Forecasting technique familiarity, satisfaction, usage, and application. *Journal of Forecasting*, 14(5), 465-476.
- Pai, P. F. Lin, C. S. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega*, 33(6), 497-505.
- Ren, C., An, N., Wang, J., Li, L., Hu, B., & Shang, D. (2014). Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting. *Knowledge-Based Systems*, 56, 226-239.
- Sailor, D. J. Muñoz, J. R. (1997). Sensitivity of electricity and natural gas consumption to climate in the USA—methodology and results for eight states. *Energy*, 22(10), 987-998.
- Soldo, B. (2012). Forecasting natural gas consumption. *Applied Energy*, 92, 26-37.
- Timmer, R. P. Lamb, P. J. (2007). Relations between temperature and residential natural gas consumption in the Central and Eastern United States. *Journal of Applied Meteorology and Climatology*, 46(11), 1993-2013.
- Wang, L., Zeng, Y., & Chen, T. (2015). Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. *Expert Systems with Applications*, 42(2), 855-863.
- Yoo, S. H. Lim, H. J. Kwak, S. J. (2009). Estimating the residential demand function for natural gas in Seoul with correction for sample selection bias. *Applied Energy*, 86(4), 460-465.
- Yu, S.; Wei, Y. M; Wang, K. (2012). A PSO-GA optimal model to estimate primary energy demand of China. *Energy Policy*, 42, 329-340.
- www.bp.com