

دریافت مقاله: ۱۳۹۴/۰۸/۲۶

پذیرش مقاله: ۱۳۹۶/۰۳/۳۱

آموزش شبکه عصبی مصنوعی با نسخه آشوب گونه الگوریتم جستجوی گرانشی و کاربرد آن در پیش بینی آلاینده های هوا: مطالعه قیاسی

منصور شیخان^{۱*}، زهرا جعفری نسب^۲

دانشیار، گروه مهندسی برق، دانشگاه آزاد اسلامی، واحد تهران جنوب، تهران، ایران

msheikhn@azad.ac.ir

کارشناس ارشد، گروه مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد تهران جنوب، تهران، ایران

z_jafari_n@yahoo.com

چکیده: امروزه پیش بینی آلودگی هوا در نواحی شهری، به دلیل تأثیر آن بر روی سلامتی انسان، یکی از موضوعات مهم در پژوهش های زیست محیطی است. با وجود اهمیت بسزای موضوع آلودگی هوا، داده های نقاط اندکی در دسترس بوده و اندازه گیری آن برای تمام نقاط مورد نظر ناممکن است. به همین علت، تاکنون مدل های مختلفی برای پیش بینی آلودگی هوا معرفی شده اند. در این مقاله، ۱۰ ویژگی مهم از مجموع ۲۰ ویژگی مربوط به داده های سازمان هواشناسی و مؤثر در پیش بینی آلودگی هوا، توسط نسخه دودویی الگوریتم جستجوی گرانشی انتخاب شده اند. در ادامه، با استفاده از نسخه آشوب گونه الگوریتم جستجوی گرانشی، یک شبکه عصبی به منظور پیش بینی آلاینده های هوای شهر تهران آموزش داده شده است. برای ارزیابی کارایی شبکه عصبی آموزش دیده، عملکرد آن در پیش بینی میزان آلاینده های هوای شهر تهران، در شرایط کاملاً یکسان با پنج روش دیگر که در آن ها شبکه عصبی با الگوریتم های جستجوی گرانشی استاندارد، پستشار خطا، اجتماع زنبورها، ترکیب الگوریتم های وراثتی با تبرید شبیه سازی شده، و بهینه سازی ازدحام ذرات آموزش دیده است، مقایسه می شود. نتایج تجربی گویای برتری روش پیشنهادی در آموزش شبکه عصبی با هدف پیش بینی آلودگی است.

واژه های کلیدی: آموزش شبکه عصبی، انتخاب ویژگی، نسخه آشوب گونه الگوریتم جستجوی گرانشی، آلودگی هوا.

۱. مقدمه

سیستم‌هایی با رفتار پیچیده دارند. پس می‌توان از آن‌ها برای پیش‌بینی میزان غلظت آلاینده‌های هوا با استفاده از پارامترهای هواشناسی استفاده کرد.

در ادامه، کارهای مرتبط در این حوزه و نوآوری روش پیشنهادی در این مقاله (در مقایسه با پژوهش‌های قبلی) ارائه خواهند شد. در ادامه، داده‌های ورودی به مدل پیش‌بین معرفتی و چگونگی پیش‌پردازش داده‌ها بیان خواهد شد. سپس الگوریتم‌های مورداستفاده برای آموزش شبکه عصبی معرفتی شده و در ادامه، نتایج شبیه‌سازی با مجموعه داده‌های سازمان هواشناسی ارائه و نتایج، مورد بحث و بررسی قرار می‌گیرند.

۲. کارهای مرتبط و نوآوری روش پیشنهادی

ارتباط بین میزان غلظت O_3 و شرایط آب‌وهوایی و به‌دنبال آن، امکان پیش‌بینی O_3 از داده‌های هواشناسی در بسیاری از پژوهش‌های قبلی گزارش شده است [۴ و ۵]. در این باره، روش‌های آماری [۸-۶]، رگرسیون خطی چندگانه [۹]، سامانه‌های فازی [۱۰] و [۱۱]، شبکه عصبی مصنوعی پرسپترون چندلایه (MLP) [۱۲-۱۴]، شبکه عصبی پایه-شعاعی (RBF) [اصلاحی ۵]، ترکیبی از چند شبکه عصبی MLP [۱۵]، مدل ANFIS [۱۲] و شبکه عصبی بازگشتی (RNN) [۱۶] مورداستفاده قرار گرفته‌اند. در موضوع تخمین غلظت آلاینده‌ها در کلان‌شهرهای کشور نیز می‌توان برای نمونه به پژوهش‌هایی درخصوص تخمین و ارزیابی غلظت آلاینده‌ها در مجاورت بزرگراه‌ها [۱۷] و تأثیر طوفان‌های گرد و غبار خاورمیانه بر غلظت آلاینده‌ها [۱۸] اشاره کرد. همچنین م. ارحامی و همکاران [۱۹] برای بهینه‌سازی مدل شبکه عصبی پیش‌بین غلظت آلاینده‌های هوای شهر تهران، تحلیل عدم قطعیت‌ها براساس شبکه عصبی و شبیه‌سازی‌های مونت‌کارلو را انجام دادند. متغیرهای ورودی به شبکه عصبی مصنوعی نیز سرعت باد، دما، رطوبت نسبی و جهت باد در نظر گرفته شده‌اند.

آلودگی هوا را می‌توان این‌گونه تعریف کرد: «وجود هر نوع آلاینده اعم از جامد، مایع، گاز یا تشعشع پرتوزا و غیرپرتوزا در هوا، به مقدار و در مدت‌زمانی که کیفیت زندگی را برای انسان و دیگر جانداران به خطر اندازد و یا به آثار باستانی و اموال خسارت وارد آورد» [۱]. امروزه مشکلات ناشی از آلودگی هوا در کلان‌شهرها، به یک چالش زیست‌محیطی تبدیل شده است. این مشکل در کلان‌شهرها به دلیل حجم ترافیکی سنگین ناشی از ترابری، استفاده از خودروهای غیراستاندارد و احتراق ناقص سوخت‌های مورداستفاده خودروها اهمیت ویژه‌ای دارد [۲].

آلاینده‌های اصلی هوا که برای آن‌ها استاندارد وجود دارد و اندازه‌گیری می‌شوند، عبارت است از [۳]: منوکسیدکربن (CO)، اکسیدهای ازت (NOx)، تمامی هیدروکربن‌ها (THC) یا ترکیبات آلی فرار (VOC)، اکسیدهای گوگرد (SOx)، تمامی ذرات معلق (TSP)، سرب (Pb) و ازن (O_3).

یکی از عوامل مهم آلودگی هوا در اغلب شهرهای بزرگ، گاز ازن در سطح زمین است. ازن در سطح زمین هنگامی تشکیل می‌شود که گازهای آلاینده حاصل از خودروها و سایر وسایلی که سوخت مصرف می‌کنند، با نور خورشید واکنش می‌دهد؛ در نتیجه گاز ازنی به وجود می‌آید که برای انسان سمی است.

باتوجه به ماهیت غیرخطی آلودگی هوا و تأثیرپذیری از پارامترهای طبیعی مختلف، پیش‌بینی میزان آلاینده‌ها بسیار دشوار بوده و نیاز به استفاده از پارامترهای هواشناسی و مکانی مختلفی دارد. برای این منظور، روابط و روش‌های گوناگونی پیشنهاد شده‌اند که هرکدام در یک منطقه یا در یک شرایط خاص، نتایج مطلوبی ارائه کرده‌اند. مدل‌های پیش‌بینی آلودگی هوا وسیله بسیار مؤثری هستند که به کمک آن‌ها می‌توان برای مدیریت آلودگی هوای کلان‌شهرها، رویکردهای مناسب و کم‌هزینه را طرح‌ریزی کرد.

شبکه‌های عصبی مصنوعی می‌توانند با استفاده از داده‌های تجربی، رفتار سیستم را یافته و به‌عنوان مدل پیش‌بین به‌کار روند. این شبکه‌ها قابلیت بسیار خوبی در مدل‌سازی

1. Multi-Layer Perceptron
2. Radial-Basis Function
3. Adaptive Neuro-Fuzzy Inference System
4. Recurrent Neural Network

درباره نوآوری پژوهش حاضر در مقایسه با پژوهش‌های مرتبط در سالیان اخیر، می‌توان به موارد زیر اشاره کرد:

الف. در مرجع [۱۶]، از ۶ ویژگی (میزان غلظت NO_2 ، دما، رطوبت نسبی، سرعت باد، جهت باد و میزان تشعشع خورشیدی) به صورت مستقیم به عنوان ورودی استفاده شده است. حال آنکه در پژوهش اخیر، از گستره وسیع تری از داده‌های ورودی (۲۰ ویژگی) بهره گرفته شده است که از این بین، به کمک الگوریتم جستجوی گرانشی دودویی (BGSA)^۱، ۱۰ ویژگی بهینه انتخاب شده‌اند. شایان ذکر است که این موضوع فراتر از ازدیاد صرف و بدون رویه تعداد ورودی‌هاست. مشکل موجود در موضوع این مقاله، وابسته بودن پیش‌بینی به تعداد زیادی از متغیرها و درعین حال، عدم تمایل به داشتن مدلی با تعداد ورودی‌های زیاد (که آن را پیچیده کند) می‌باشد. لذا انتخاب ویژگی با روش پیشنهادی، یکی از نوآوری‌های این مقاله برای غلبه بر این مشکل است که در پژوهش‌های قبلی گزارش نشده است.

ب. شاید نزدیک‌ترین پژوهش‌های انجام‌شده به پژوهش حاضر، از لحاظ به کارگیری روش‌های ترکیبی و پیش‌بینی غلظت آلاینده‌های هوا توسط شبکه‌های عصبی، روش‌های پیشنهادی در مراجع [۲۰ و ۲۱] باشد. در مرجع [۲۰]، یک واحد استخراج معرفت (KD)^۲ همراه با شبکه عصبی به کار گرفته شده است تا سهم/اهمیت هریک از ورودی‌ها را در میزان متغیر موردپیش‌بینی (از طریق محاسبه صرف یک ضریب اهمیت) تعیین کند (شبیه همان کاری که در پژوهش حاضر با به کارگیری واحد انتخاب ویژگی مبتنی بر روش هوشمند BGSA انجام می‌شود). به همین ترتیب در مرجع [۲۱]، از روش سنتی تحلیل مؤلفه اصلی (PCA)^۳ برای انتخاب ورودی‌ها به شبکه عصبی و ماشین بردار پشتیبان (SVM)^۴ به عنوان دو ابزار پیش‌بینی میزان غلظت آلاینده O_3 استفاده شده است. حال آنکه در مقاله حاضر، از روش هوشمند و نوین تر BGSA برای انتخاب ویژگی و نیز نسخه آشوب‌گونه الگوریتم جستجوی گرانشی (CGSA)^۵ برای

آموزش شبکه عصبی مصنوعی (به جای روش استاندارد پس‌انتشار خطا) استفاده شده و مقایسه نتایج با روش‌های مرتبط رقیب نیز صورت پذیرفته است.

شایان ذکر است که آموزش شبکه عصبی مصنوعی با الگوریتمی مناسب از چالش‌های به کارگیری این مدل در کاربردهای پیش‌بینی است. با توجه به مشکل روش سنتی پس‌انتشار خطا (EBP)^۶ در آموزش (یعنی گرفتاری در کمینه‌های محلی)، استفاده از روش‌های هوش محاسباتی در بسیاری از پژوهش‌های اخیر توصیه شده است. به کارگیری CGSA در این مقاله، نتایج مناسبی را در مقایسه با روش‌های رقیب ارائه نموده است.

پ. موضوع به کارگیری CGSA در بهینه‌سازی ساختار مدل‌ها نیز در مورد مدل رگرسیون بردار پشتیبان^۷ (SVR) [۲۲]، مدل فازی تاکاگی-سوگنو^۸ [۲۳] و شناسایی پارامترهای یک سیستم آشوبناک^۹ [۲۴] گزارش شده است. اما براساس اطلاع نویسندگان مقاله، موضوع به کارگیری تلفیق CGSA و BGSA در انتخاب ورودی‌ها و آموزش شبکه عصبی با هدف پیش‌بینی میزان غلظت آلاینده‌ها، تاکنون در مرجعی گزارش نشده است.

۳. داده‌های ورودی به مدل پیش‌بین

برای پیش‌بینی مقدار پارامترهای کیفی هوا، مجموعه‌ای از داده‌ها به عنوان ورودی به شبکه عصبی داده می‌شود که نمونه‌هایی از آن‌ها عبارت است از [۲۵]:

- الف. روزهای هفته (برای مثال، کاهش تردد خودروها در آخر هفته در اطراف مکان‌های اداری)؛
- ب. روزهای ماه (برای مثال، افزایش حجم ترافیک در مهرماه به علت بازگشایی مدارس)؛
- پ. سال اندازه‌گیری پارامترهای آلودگی هوا (برای مثال به دلیل آغاز فعالیت یک کارخانه در یک سال نسبت به سال قبل، آلودگی روند رو به رشدی دارد)،

6. Error Back-Propagation
7. Support Vector Regression
8. Takagi-Sugeno
9. Identification
10. Chaotic

1. Binary Gravitational Search Algorithm
2. Knowledge Discovery
3. Principal Component Analysis
4. Support Vector Machine
5. Chaotic Gravitational Search Algorithm

ر. فشار هوا (فشار هوا نیرویی است که در هر نقطه بر حسب وزن ستونی از هوا که در بالای آن نقطه وجود دارد، بر سطح وارد می‌شود. این فشار در سطح آب‌ها و دریاهای آزاد در حدود یک کیلوگرم بر سانتی‌متر مربع است. با افزایش ارتفاع از سطح زمین، به دلیل کاسته شدن ارتفاع ستون هوای قرار گرفته در بالای سطح، فشار هوا کاهش می‌یابد)؛

ز. میزان بارندگی؛

ژ. میزان تبخیر؛

س. غبار (ریزگرد به پدیده‌ای هواشناختی گفته می‌شود که در اثر گرد و غبار، ذرات خشک معلق و دود به وجود می‌آید). در این پژوهش، از داده‌های آلودگی هوا، مربوط به ایستگاه مهرآباد شهر تهران، برای یک دوره دوساله (سال‌های ۲۰۱۱ و ۲۰۱۲) استفاده شده است. داده‌های خام از سازمان هواشناسی ایران تهیه شده و دارای ۷۳۰ رکورد است.

۱.۳.۱. پیش‌پردازش داده‌ها

داده‌های موجود در جدول سازمان هواشناسی، بدین ترتیب برای استفاده در شبکه عصبی، پردازش شدند: ابتدا داده‌های نامعتبر پاک‌سازی شدند، سپس عمل هنجارسازی داده‌ها^۴ انجام شده و در نهایت با استفاده از BGSA ویژگی‌ها انتخاب شدند.

۱.۳.۱.۱. پاک‌سازی داده‌ها

در برخی روزها، داده‌های هواشناسی و آلودگی هوا ثبت نشده است. این اتفاق می‌تواند به علت وجود اشکال در دستگاه ثبت اطلاعات باشد. برای تخمین و تکمیل اطلاعات آلودگی، پاک‌سازی داده‌ها چنین انجام شد: ابتدا کنترل داده‌های مشکوک و مقایسه آن با داده‌های روزهای قبل و بعد انجام شد. سپس داده‌های هواشناسی با مقدار معمول مقایسه شد. مثلاً در برخی روزها سرعت باد محاسبه شده از مقدار معقول سرعت باد در تهران، بسیار بیشتر است، که به هیچ‌وجه قابل قبول نیست.

از آنجاکه هدف از این پژوهش، تحقق شبکه عصبی مصنوعی برای پیش‌بینی میزان آلاینده ازن (O_3) است، ستون‌های مربوط به آلاینده‌های CO ، PM_{10} ، NO_2 و SO_2 حذف شدند.

ت. سرعت باد (وزش باد در پراکندگی آلودگی هوا تأثیر محسوسی دارد و با افزایش ارتفاع، سرعت باد روند افزایشی دارد)؛

ث. جهت باد (جهت باد را از سمتی که می‌وزد گزارش می‌کنند، مثلاً باد شمال از شمال به جنوب می‌وزد. جهت باد معمولاً یا با چهار جهت اصلی و یا با درجه سمت^۱ گزارش می‌شود. برای مثال، بادی که از جنوب می‌وزد، باد ۱۸۰ درجه و بادی که از شرق می‌وزد، باد ۹۰ درجه نامیده می‌شود)؛

ج. باد غالب (هر منطقه باد غالب و برجسته‌ای دارد که در فصل خاص و گاهی در تمام فصول می‌وزد. باد غالب، باد خاصی است که وزش آن طولانی‌تر بوده و در جهت خاصی می‌وزد)؛

چ. رطوبت نسبی (به میزان بخارآب موجود در جو، رطوبت می‌گویند. نسبت بخارآب موجود در حجم معینی از هوا در دمای t به وزن بیشینه بخارآبی را که می‌تواند در همین حجم در دمای t داشته باشد، رطوبت نسبی می‌گویند)؛

ح. قابلیت دید^۲ (قابلیت دید در هوا، بزرگ‌ترین فاصله‌ای است که امکان دیدن و تشخیص دادن با چشم غیر مسلح وجود دارد. دید در هواشناسی اندازه فاصله‌ای است که در آن یک شیء یا نور می‌تواند تشخیص داده شود)؛

خ. فشار سطح دریا^۳ (فشار اتمسفری، فشار جو یا فشار هوا، مقدار نیرویی است که از طرف جو به واحد سطح اعمال می‌شود. با افزایش ارتفاع از سطح زمین فشار هوا کاهش می‌یابد. بنابراین مطابق قرارداد، برای استانداردسازی واحدها، از فشار معادل سطح دریا در گزارشات هواشناسی استفاده می‌شود)؛

د. دما (بزرگ‌ترین تأثیری که پارامتر دما بر آلودگی هوا دارد، ایجاد اختلاف فشار در مناطق مختلف و در نتیجه ایجاد حرکت در هوا و جابه‌جایی و پراکنش آلودگی است که می‌تواند با خود کاهش غلظت‌ها را به دنبال داشته باشد. عدم امکان وقوع این پدیده در ایجاد وارونگی دمایی نقش اصلی را دارد)؛

ذ. مقدار غلظت آلاینده در روز قبل؛

1. Azimuth
2. Visibility
3. Sea level pressure

4. Data normalization

است (ویژگی های ۱۶ و ۲۰)؛ لذا آن‌ها نیز به‌عنوان ورودی‌های مستقل در نظر گرفته شده‌اند.

جدول (۱): داده‌های ورودی مدل پیش‌بینی غلظت آلاینده O₃ هوای

تهران

شرح	نام داده	ردیف
کمینه میزان دید	Visibility_min	۱
بیشینه دما	Temperature_max	۲
کمینه دما	Temperature_min	۳
میانگین دما	Temperature_avg	۴
کمینه فشار سطح دریا	Sea level pressure_min	۵
میانگین فشار سطح دریا	Sea level pressure_avg	۶
میانگین فشار سطح ایستگاه	Station level pressure_avg	۷
غبار	Nebulosity	۸
کمینه رطوبت	Humidity_min	۹
بیشینه رطوبت	Humidity_max	۱۰
میانگین رطوبت	Humidity_avg	۱۱
جهت باد در ارتفاع ۳ متری	Wind direction_03	۱۲
جهت باد در ارتفاع ۹ متری	Wind direction_09	۱۳
جهت باد در ارتفاع ۱۵ متری	Wind direction_15	۱۴
میانگین جهت باد	Wind direction_avg	۱۵
بیشینه جهت باد	Wind direction_max	۱۶
سرعت باد در ارتفاع ۳ متری	Wind speed_03	۱۷
سرعت باد در ارتفاع ۹ متری	Wind speed_09	۱۸
سرعت باد در ارتفاع ۱۵ متری	Wind speed_15	۱۹
بیشینه سرعت باد	Wind speed_max	۲۰

۳.۱.۲. هنجارسازی داده‌ها

هنجارسازی داده‌ها با به‌کارگیری رابطه (۱) انجام شده است:

$$X = 2 * \frac{x - \min(x)}{\max(x) - \min(x)} - 1 \quad (1)$$

که $\min(x)$ مین کمینه و $\max(x)$ مبین بیشینه بردار ورودی x بوده و X مقدار هنجارشده آن است.

۳.۲. انتخاب ویژگی

در این پژوهش، از الگوریتم جستجوی گرانشی دودویی برای

ستون‌های Rain24 (میانگین بارندگی در ۲۴ ساعت گذشته) و Evapor_transpiration (میزان تبخیر) به‌علت دارا بودن تعداد بسیار زیاد نقص داده‌ها (به‌ترتیب ۴۷۸ و ۲۰۸) حذف شدند. در سایر ستون‌های جدول داده‌ها، نقص داده‌ها با میانگین سایر داده‌های موجود در همان ستون، جای‌گذاری شدند. درنهایت با اعمال موارد فوق، داده‌های ۶۸۱ روز قابل قبول شناخته شد که شامل ۲۰ ویژگی ورودی به‌صورت جدول (۱) و مقدار خروجی ازن هستند. بدین ترتیب برای پیش‌بینی میزان آلاینده ازن در هر روز، از داده‌های روز گذشته استفاده شده است. گستره مقدار غلظت ازن در دادگان موجود نیز در بازه ۲۵ تا ۹۷۸ جزء در هزار میلیون^۱ (ppb) است.

شایان یادآوری است که پارامترهای زیادی از جمله دما، رطوبت، و وزش باد فشار سطح دریا را تغییر می‌دهند؛ لذا فشار مطلق سطح دریا در طول روزهای مختلف، متفاوت بوده و در دادگان موجود بین ۹۸۸/۵ و ۱۰۳۲/۵ میلی‌بار گزارش شده است.

گفتنی است که انواع متفاوتی برای جهت باد در دادگان مورداستفاده در این مقاله بدین ترتیب در نظر گرفته شده است: شمال به جنوب، شمال به شرق، شمال به غرب، جنوب به شمال، جنوب به شرق، جنوب به غرب، شرق به شمال، شرق به جنوب، شرق به غرب، غرب به شمال، غرب به جنوب، و غرب به شرق. براین‌اساس، در داده‌های سازمان هواشناسی این جهت‌ها با اعداد صفر تا ۱۱ شماره‌گذاری شده‌اند. لذا جهت باد، یک داده مقیاس‌پذیر در نظر گرفته شده و حتی میانگین آن نیز با یک عدد غیرصحیح در دادگان گزارش شده و به‌عنوان ورودی به شبکه عصبی اعمال می‌شود.

مقدار ویژگی پانزدهم (میانگین جهت باد) در دادگان ارائه‌شده در رکوردهای مختلف، لزوماً میانگین ۳ مقدار متناظر در ویژگی‌های دوازدهم تا چهاردهم (جهت باد در ارتفاع‌های ۳، ۹ و ۱۵ متری) نیست که حکایت از اندازه‌گیری در ارتفاع‌های مختلف می‌کند؛ لذا به‌عنوان ویژگی مستقل در نظر گرفته شده است. این وضعیت درباره ویژگی‌های مبین بیشینه نیز برقرار

1. Parts per billion

جدول (۲): مقادیر در نظر گرفته شده برای پارامترهای BGSA

نام پارامتر	مقدار
اندازه جمعیت	۴۰
تعداد تکرارها	۳۰
ضریب ثابت گرانشی	۱

جدول (۳): ویژگی‌های انتخاب شده توسط BGSA

ردیف	نام ویژگی
۱	کمینه دما
۲	میانگین دما
۳	کمینه فشار سطح دریا
۴	بیشینه رطوبت
۵	میانگین رطوبت
۶	جهت باد در ارتفاع ۳ متری
۷	جهت باد در ارتفاع ۱۵ متری
۸	بیشینه جهت باد
۹	سرعت باد در ارتفاع ۳ متری
۱۰	سرعت باد در ارتفاع ۹ متری

حل مسئله انتخاب ویژگی استفاده شده است. نسخه دودویی الگوریتم جستجوی گرانشی توسط راشدی و همکاران در سال ۲۰۱۰ معرفی شده است [۲۶]. در مسئله انتخاب ویژگی با BGSA هر جرم می‌تواند یک جواب برای مسئله باشد. هر جواب یا جرم به صورت رشته‌ای از صفر و یک‌ها تعریف می‌شود. طول رشته برابر کل ویژگی‌هاست. مقدار صفر و یا یک برای هر رشته - بیت به ترتیب بیانگر مشارکت و یا عدم مشارکت ویژگی متناظر با آن بیت است. پس از ایجاد جمعیت اولیه اجرام، برای هر جرم مقدار برازندگی آن با استفاده از تابع برازندگی که در اینجا میانگین مربعات خطا (MSE) است، محاسبه می‌شود. در هر مرحله، بهترین جرم (که با توجه به هدف مسئله مشخص می‌شود) انتخاب می‌شود. به تعداد تکرارهای مسئله، الگوریتم اجرا می‌شود و در نهایت با انتخاب جرم با کمترین مقدار تابع برازندگی ویژگی‌های مناسب به دست می‌آیند [۲۷].

تابع احتمال $S(V_i^d)$ طبق رابطه (۲) تعریف می‌شود [۲۶]:

$$S(V_i^d(t)) = |\tanh(V_i^d(t))| \quad (2)$$

که در آن، \tanh مبین تابع تانژانت هیپربولیک است. پس از محاسبه $S(V_i^d)$ اجرام طبق رابطه (۳) جابه‌جا می‌شوند:

if rand < $S(V_i^d(t+1))$ then

$$x_i^d(t+1) = \text{complement}(x_i^d(t)) \text{ else } \quad (3)$$

$$x_i^d(t+1) = x_i^d(t)$$

که در آن، rand یک متغیر تصادفی یکنواخت در بازه $[0,1]$ و complement مبین عملگر متمم منطقی (نفی) است. مقادیر در نظر گرفته شده برای پارامترهای CGSA به شرح جدول (۲) می‌باشد. بهترین جواب به دست آمده برای MSE مقدار 0.0022 بود. خروجی اجرای CGSA بر روی داده‌های سازمان هواشناسی، ۱۰ ویژگی برای پیش‌بینی آلاینده‌های هوای شهر تهران به صورت جدول (۳) بود که این ویژگی‌ها به عنوان ورودی به شبکه عصبی داده شدند.

۴. بررسی کارایی روش‌های بهینه‌سازی در آموزش

MLP

پس از آماده‌سازی داده‌ها، شبکه عصبی باید با استفاده از داده‌های ایستگاه مهرآباد شهر تهران آموزش ببیند تا بتواند با توجه به قابلیت تعمیم ذاتی خود، برای سایر نقاط تهران آلودگی را پیش‌بینی کند.

شبکه پیشنهادی مورد نظر یک شبکه MLP با یک لایه مخفی و تعداد ۴ نرون در لایه مخفی است. تابع تحریک مورد استفاده در این شبکه برای لایه پنهان، سیگموئید (Tansig) و برای لایه خروجی، خطی (Purelin) انتخاب شد. از بین ۶۸۱ رکورد موجود در جدول داده‌ها، ۷۰ درصد داده‌ها (۴۷۷ رکورد) به صورت تصادفی، برای آموزش و اعتبارسنجی^۲ شبکه (به ترتیب با ابعاد ۵۰ و ۲۰ درصد داده‌ها) و ۳۰ درصد باقی مانده (۲۰۴ رکورد)، برای آزمون آن به کار گرفته شد.

در شرایط یکسان، شبکه عصبی MLP مورد نظر برای پیش‌بینی غلظت آلاینده O_3 هوای شهر تهران توسط الگوریتم‌های

1. Mean Square Error

2. Validation

گام‌های اصلی این الگوریتم چنین هستند [۲۹]:

۱. منابع غذای اولیه برای تمامی زنبورهای کارگر تأمین می‌شود.
۲. هر زنبور کارگر براساس حافظه خود به سراغ یک منبع غذا می‌رود و منبع همسایه را مشخص می‌کند. سپس میزان شهد آن را مشخص می‌کند و در کندو به حرکت درمی‌آید.
۳. هر زنبور ناظر حرکت زنبورهای کارگر را مشاهده کرده و براساس آن یکی از منابع آن‌ها را انتخاب می‌کند و سپس به سمت آن منبع می‌رود. پس از انتخاب یک همسایه اطراف آن، میزان شهد آن را ارزیابی می‌کند.
۴. منابع غذای متروکه تعیین و با منابع غذای جدید که توسط زنبورهای پیشاهنگ کشف شده‌اند، جایگزین می‌شوند.
۵. بهترین منبع غذای یافت‌شده ثبت می‌شود.
۶. بازگشت به گام دوم و تکرار الگوریتم تا برآورده شدن شرایط توقف.

در این پژوهش، الگوریتم ABC برای آموزش شبکه MLP مورد استفاده قرار گرفت که جمعیت اولیه آن برابر ۴۰ و پس از ۱۰۰ بار تکرار الگوریتم، بهترین جواب به دست آمده برای MSE مقدار ۰/۰۰۲۳ بود.

۳.۴ آموزش MLP با الگوریتم GA+SA

در این پژوهش، الگوریتم وراثتی (GA) و شبیه‌سازی تبرید (SA)، برای آموزش شبکه عصبی MLP با یکدیگر ترکیب شده‌اند. در الگوریتم وراثتی، هر کروموزوم نشان‌دهنده یک نقطه در فضای جستجو و یک راه‌حل ممکن برای مسئله مورد نظر است. کروموزوم‌ها (راه‌حل‌ها) از تعداد ثابتی ژن (متغیر) تشکیل شده‌اند. برای نمایش کروموزوم‌ها معمولاً از کدگذاری‌های دودویی استفاده می‌شود. مجموعه‌ای از کروموزوم‌ها یک جمعیت را تشکیل می‌دهند. با تأثیر عملگرهای وراثتی بر روی هر جمعیت، جمعیت جدیدی با همان تعداد کروموزوم تشکیل می‌شود [۳۰].

بسیاری از الگوریتم‌های بهینه‌سازی با الگوگیری و شبیه‌سازی یکی از قوانین یا روابط موجود در طبیعت بنا نهاده می‌شوند. الگوریتم شبیه‌سازی تبرید نیز بر مبنای فرایند تبرید یا بازپخت فلزات بنا نهاده شده است [۳۱]. فرایندی که ذکر شد بدین صورت

EBP، اجتماع زنبورها (ABC)، ترکیب GA و SA^۲، PSO^۳، GSA^۵ و CGSA آموزش داده شد. تمام نتایج برای ۱۰۰ بار تکرار هر الگوریتم آورده شده است. همچنین با توجه به تصادفی بودن الگوریتم‌های فوق، تمام آزمایش‌ها ۱۰ مرتبه تکرار شده است.

۱.۴ آموزش MLP با الگوریتم پس‌انتشار خطا

در الگوریتم EBP در هر مرحله، مقدار خروجی محاسبه شده جدید با مقدار واقعی، مقایسه شده و با توجه به خطای به دست آمده به اصلاح وزن‌های شبکه پرداخته می‌شود، به نحوی که در انتهای هر تکرار، اندازه خطای حاصل کمتر از میزان به دست آمده در تکرار قبلی باشد. الگوریتم EBP را می‌توان متشکل از چهار گام اصلی در نظر گرفت [۲۸]. بدین ترتیب که پس از انتخاب تصادفی ضرایب وزن شبکه، گام‌های زیر طی می‌شود:

۱. محاسبات سطح فعالیت و خروجی نرون‌ها در مسیر ورودی به خروجی؛
 ۲. پس‌انتشار به لایه خروجی و تغییر ضرایب وزن لایه خروجی؛
 ۳. پس‌انتشار به لایه (های) مخفی و تغییر ضرایب وزن لایه (های) مخفی؛
 ۴. تکرار الگوریتم تا هنگامی که به آستانه خطای مورد پذیرش رسیده یا تعداد تکرارهای تعیین شده انجام شده باشد.
- در این پژوهش، الگوریتم EBP نیز برای آموزش شبکه MLP مورد استفاده قرار گرفت. بهترین جواب به دست آمده برای MSE مقدار ۰/۰۰۵۵ بود.

۲.۴ آموزش MLP با الگوریتم اجتماع زنبورها

در الگوریتم ABC موقعیت منبع غذایی یک راه‌حل مسئله بهینه‌سازی را نشان می‌دهد. یادآوری می‌شود که در الگوریتم ABC سه گروه زنبور وجود دارد: کارگر^۶، ناظر^۷ و پیشاهنگ^۸.

1. Artificial Bee Colony
2. Genetic Algorithm
3. Simulated Annealing
4. Particle Swarm Optimization
5. Gravitational Search Algorithm
6. Employed
7. Onlooker
8. Scout

گروه پرندگان مدل‌سازی شده است. این الگوریتم از حرکت دسته‌جمعی پرندگانی که به دنبال غذا هستند، الهام گرفته شده است. گروهی از پرندگان در فضایی به صورت تصادفی به دنبال غذا می‌گردند. هیچ‌یک از پرندگان محل غذا را نمی‌دانند. یکی از بهترین تدابیر می‌تواند دنبال کردن پرنده‌ای باشد که کمترین فاصله را تا غذا داشته باشد. این تدبیر در واقع جان‌مایه الگوریتم است. فضای مورد جستجو در الگوریتم PSO معادل فضای مورد جستجو در الگوی حرکت پرندگان است. هر راه‌حل در الگوریتم PSO، که به آن یک ذره گفته می‌شود، معادل یک پرنده است و تعداد ذرات (راه‌حل‌ها) معادل تعداد پرندگان است. هر ذره یک مقدار شایستگی دارد که توسط یک تابع برازش محاسبه می‌شود و هرچه ذره در فضای جستجو به هدف (یعنی غذا در مدل حرکت پرندگان) نزدیک‌تر باشد، برازش بیشتری دارد. همچنین هر ذره دارای یک جابه‌جایی است که هدایت حرکت ذره را بر عهده دارد و به کمک آن، مکان بعدی ذره مشخص می‌شود. هر ذره با دنبال کردن ذرات بهینه در حالت فعلی، به حرکت خود در فضای جستجو ادامه می‌دهد تا اینکه در نهایت به جواب بهینه دست یابد.

الگوریتم PSO از چهار گام اصلی به ترتیب زیر تشکیل شده است [۳۳]:

۱. ارزیابی برازش هر ذره؛
 ۲. به‌روزرسانی برازش‌های تکی و جمعی؛
 ۳. به‌روزرسانی سرعت و مکان هر ذره؛
 ۴. تکرار الگوریتم تا هنگامی که شرایط توقف برآورده شود.
- در این پژوهش، الگوریتم PSO برای آموزش شبکه MLP مورد استفاده قرار گرفت که جمعیت اولیه در آن برابر ۴۰ و پس از ۱۰۰ بار تکرار الگوریتم، بهترین جواب به‌دست‌آمده برای MSE مقدار ۰/۰۰۳۶ بود.

۵.۴ آموزش MLP با الگوریتم GSA

در سال ۲۰۰۹، روش بهینه‌سازی دیگری به نام الگوریتم جستجوی گرانشی توسط راشدی و همکاران معرفی شد [۳۴]. این الگوریتم، از خانواده الگوریتم‌های هوش جمعی است که از

است که مواد تا حد امکان گرم شده و به آهستگی سرد می‌شوند تا به کمترین حد انرژی خود برسند و به اصطلاح، به پایداری ترمودینامیکی برسند. الگوریتم شبیه‌سازی تبرید الگوریتمی مبتنی بر جمعیت نیست و براساس یک نقطه عمل می‌کند. این الگوریتم براساس جستجوی محلی عمل کرده و عملکرد خوبی در مسائل گسسته و پیوسته دارد [۳۱]. برای حل یک مسئله بهینه‌سازی، الگوریتم SA ابتدا از یک جواب اولیه شروع می‌کند و سپس در یک حلقه تکرار به جواب‌های همسایه حرکت می‌کند. اگر جواب همسایه بهتر از جواب فعلی باشد، الگوریتم آن را به‌عنوان جواب فعلی قرار می‌دهد (به آن حرکت می‌کند)، در غیر این صورت، الگوریتم آن جواب را با احتمال $\exp(-\Delta E/T)$ به‌عنوان جواب فعلی می‌پذیرد. در این رابطه، ΔE تفاوت بین تابع هدف جواب فعلی و جواب همسایه است و T یک پارامتر به نام دماست. در هر دما، چندین تکرار اجرا و سپس دما به آرامی کاهش داده می‌شود. در گام‌های اولیه دما خیلی بالا قرار داده می‌شود تا احتمال بیشتری برای پذیرش جواب‌های بدتر وجود داشته باشد. با کاهش تدریجی دما، در گام‌های پایانی احتمال کمتری برای پذیرش جواب‌های بدتر وجود خواهد داشت و بنابراین الگوریتم به سمت یک جواب خوب همگرا می‌شود [۳۲].

روال کار در این روش ترکیبی نیز به این صورت بوده که ابتدا توسط الگوریتم وراثتی، عملگرهای تقاطع و جهش انجام پذیرفته است. سپس برای تشکیل جمعیت جدید، از الگوریتم شبیه‌سازی تبرید استفاده شده است. بدین صورت که ابتدا جمعیت فرزندان و جمعیت جهش‌یافتگان با یکدیگر ترکیب شده و برحسب تابع هدف (در اینجا MSE) به تعداد جمعیت (در اینجا ۴۰) از آن‌ها را جدا می‌کنیم. این جمعیت ۴۰ عضوی را با استفاده از قانون SA با جمعیت قبلی مقایسه و جایگزین می‌کنیم. جمعیت اولیه در این الگوریتم برابر ۴۰ و پس از ۱۰۰ بار تکرار، بهترین جواب به‌دست‌آمده برای MSE مقدار ۰/۰۰۳۷ بود.

۴.۴ آموزش MLP با الگوریتم PSO

الگوریتم بهینه‌سازی ازدحام ذرات یک الگوریتم بهینه‌سازی تصادفی براساس جمعیت است که از شبیه‌سازی رفتار اجتماعی

$$R_{ij}(t) = \|X_i(t), X_j(t)\|^2 \quad (۶)$$

در رابطه (۵)، ε یک عدد بسیار کوچک است. نیروی وارد بر جرم i در جهت بُعد d در زمان t ($F_i^d(t)$)، برابر مجموع نیروهایی است که k جرم برتر جمعیت بر جرم وارد می‌کنند (رابطه ۷). مقصود از اجرام برتر، عامل‌هایی هستند که دارای برازندگی بیشتری باشند.

$$F_i^d(t) = \sum_{j \in kbest, j \neq i} rand_j(t) \times F_{ij}^d(t) \quad (۷)$$

در رابطه (۷)، $kbest$ بیانگر مجموعه k جرم برتر جمعیت است که بهترین مقدار برازندگی و بیشترین جرم را دارند. همچنین در این رابطه، $rand_j$ عددی تصادفی با توزیع یکنواخت در بازه $[0,1]$ است که برای حفظ خصوصیت تصادفی بودن جستجو در نظر گرفته می‌شود.

طبق قانون دوم نیوتن هر جرم در جهت بُعد d شتابی می‌گیرد که متناسب است با نیروی وارد بر جرم i در آن جهت، تقسیم بر جرم i . رابطه (۸) شتاب جرم i در جهت بُعد d در زمان t را با $a_i^d(t)$ نشان می‌دهد.

$$a_i^d(t) = F_i^d(t)/M_i(t) \quad (۸)$$

سرعت هر جرم برابر مجموع ضربی از سرعت فعلی جرم و شتاب جرم است که طبق رابطه (۹) تعریف می‌شود. موقعیت جدید بُعد d از جرم i نیز به صورت رابطه (۱۰) محاسبه می‌شود.

$$V_i^d(t+1) = rand_i \times V_i^d(t) + a_i^d(t) \quad (۹)$$

$$x_i^d(t+1) = x_i^d(t) + V_i^d(t+1) \quad (۱۰)$$

در روابط فوق، $V_i^d(t)$ سرعت بُعد d عامل i در زمان t و $rand_i$ عددی تصادفی با توزیع یکنواخت در بازه $[0,1]$ است. برای تنظیم ضریب گرانش از رابطه (۱۱) استفاده می‌شود.

$$G(t) = G_0 [\exp(-at/T)] \quad (۱۱)$$

که در آن، G_0 مقدار اولیه ثابت گرانش است؛ t تکرار جاری و T تعداد کل تکرارهاست. ثابت گرانش، پارامتری مناسب برای کنترل دو ویژگی کاوش و بهره‌وری در این الگوریتم به‌شمار می‌آید. مقادیر بزرگ آن موجب تقویت توانایی کاوش الگوریتم و مقادیر کوچک آن موجب افزایش توانایی بهره‌وری الگوریتم می‌شود. از آنجاکه در مراحل اولیه جستجو لازم است الگوریتم به جستجوی نقاط جدیدی در فضای مسئله پرداخته شود و در

قوانین جاذبه میان اجرام و حرکت نیوتنی الهام گرفته شده است. طبق قانون جاذبه نیوتن، هر جسم به اجسام دیگر نیرو وارد کرده و آن‌ها را به سمت خود جذب می‌کند. به‌وضوح، هرچه این اجسام بزرگ‌تر و نزدیک‌تر باشند، تأثیر این نیرو بیشتر خواهد بود. در نتیجه هر جسم با استفاده از نیروی جاذبه، محل و مقدار جرم سایر اجسام را درک می‌کند. بنابراین می‌توان از این نیرو به‌عنوان رسانه‌ای برای تبادل اطلاعات استفاده کرد. از الگوریتم جستجوی گرانشی در حل مسائل بهینه‌سازی استفاده می‌شود. در این الگوریتم، پاسخ‌های موردنظر، موقعیت اجرام در فضای مسئله‌اند. میزان اجرام نیز باتوجه به تابع هدف تعیین می‌شود [۳۵].

پس از تشکیل سیستم، قوانین حاکم بر آن مشخص می‌شوند. فرض می‌کنیم تنها قانون گرانش و قوانین حرکت حاکم هستند. صورت کلی این قوانین تقریباً شبیه قوانین طبیعت است و به‌صورت زیر تعریف شده‌اند [۳۶]:

قانون گرانش: هر جرم در سیستم مصنوعی، تمام اجرام دیگر را به سمت خود جذب می‌کند. مقدار این نیرو متناسب است با جرم گرانشی و عکس فاصله دو جرم.

قانون حرکت: سرعت فعلی هر جرم برابر است با مجموع ضربی از سرعت قبلی جرم و تغییر سرعت آن. تغییر سرعت یا شتاب هر جرم نیز برابر است با نیروی وارد بر آن تقسیم بر جرم اینرسی.

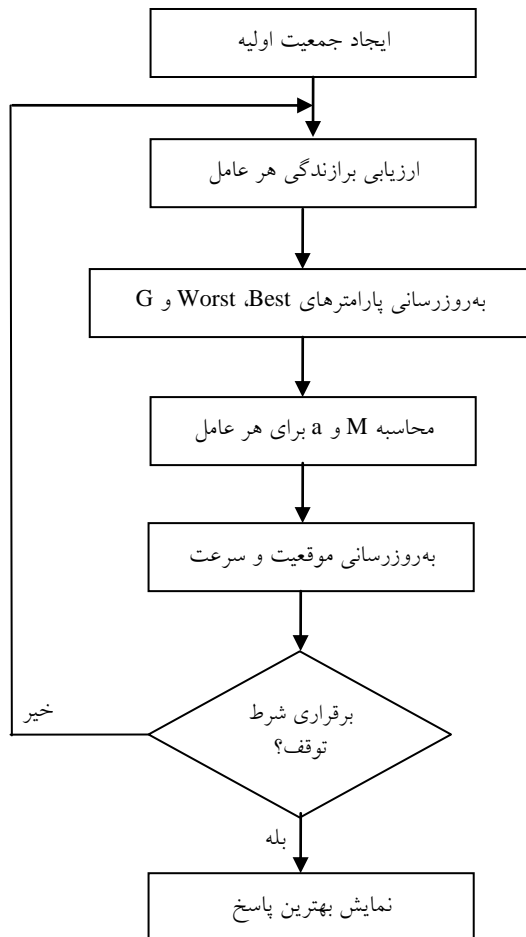
سیستم به‌صورت مجموعه‌ای از m جرم تصور می‌شود. موقعیت هر جرم می‌تواند جوابی برای مسئله باشد. در رابطه (۴) موقعیت بُعد d از جرم i با x_i^d نشان داده می‌شود:

$$x_i = (x_i^1, \dots, x_i^d, \dots, x_i^n), \quad i = 1, 2, \dots, m \quad (۴)$$

در رابطه (۴) نشان‌دهنده بُعد فضای پاسخ است. در این سیستم در زمان t ، به هر جرم i از سوی جرم j در جهت بُعد d نیرویی به اندازه $F_{ij}^d(t)$ وارد می‌شود. مقدار این نیرو طبق رابطه (۵) محاسبه می‌شود. $G(t)$ ثابت گرانش در زمان t و R_{ij} فاصله بین دو جرم i و j است. برای تعیین فاصله بین اجرام مطابق رابطه (۶) از فاصله اقلیدسی استفاده می‌شود.

$$F_{ij}^d(t) = G(t) \frac{M_i(t) \cdot M_j(t)}{R_{ij}(t) + \varepsilon} (x_j^d(t) - x_i^d(t)) \quad (۵)$$

استفاده شده است. از این‌رو با توجه به زمینه‌های کاربردی مورد استفاده، نسخه‌های متفاوتی از این الگوریتم ارائه شده است که می‌توان به BGSA [۲۶]، الگوریتم جست‌وجوی گرانشی نسخه‌گرای پیشرفته (AEGSA) [۴۱]، و الگوریتم جست‌وجوی گرانشی چندهدفه (MOGSA) [۴۲] اشاره کرد.



شکل (۱): روندنمای الگوریتم GSA [۳۴]

الگوریتم جست‌وجوی گرانشی باید به گونه‌ای هدایت شود که موقعیت اجرام با گذشت زمان بهبود پیدا کند. استراتژی به کار گرفته شده برای این منظور، مبنی بر تنظیم جرم عامل‌هاست. طبق روابط ارائه شده، تمام عامل‌ها متناسب با جرمشان روی هم تأثیر می‌گذارند و اثرگذاری آن‌ها روی عامل‌های همسایه بیشتر است. تأثیر عامل‌های با جرم سنگین‌تر بیشتر است و شعاع تأثیرگذاری بزرگ‌تری دارند. بنابراین، به اجرامی که تابع برازش

مراحل پایانی با افزایش توان بهره‌وری به بهبود جواب‌های دیده شده بپردازد، گزینه مناسب برای ثابت گرانشی بایستی با یک مقدار اولیه بزرگ شروع شده و با گذشت زمان، مقدار آن به تدریج کاهش یابد.

در رابطه (۱۲)، جرم عامل‌ها بر مبنای تابع هدف آن‌ها تنظیم می‌شود. به گونه‌ای که به عامل‌های با برازش بیشتر، جرم بیشتری نسبت داده می‌شود.

$$m_i(t) = [fit_i(t) - worst(t)] / [best(t) - worst(t)] \quad (12)$$

در این رابطه، $fit_i(t)$ بیانگر میزان برازندگی جرم i در زمان t است، $best(t)$ و $worst(t)$ به ترتیب بیانگر میزان برازش بهترین و بدترین عامل جمعیت در زمان t هستند. در نهایت، اندازه جرم عامل‌ها طبق رابطه (۱۳) هنجار می‌شود.

$$M_i(t) = m_i(t) / \sum_{j=1}^N m_j(t) \quad (13)$$

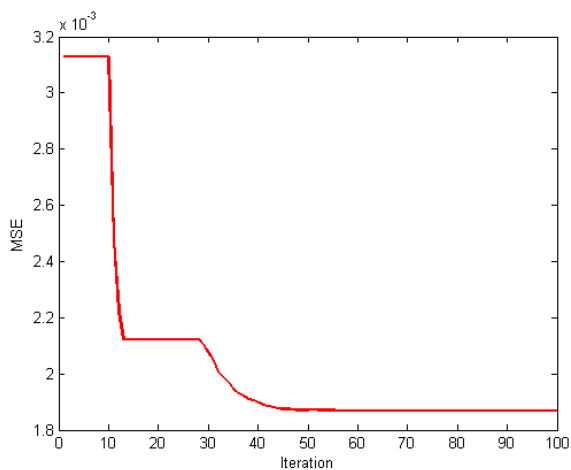
از آنجاکه در این پژوهش، هدف کاهش میزان MSE می‌باشد، این مسئله یک مسئله کمینه‌یابی بوده و می‌توان از روابط (۱۴) و (۱۵) برای محاسبه بهترین و بدترین عامل‌ها استفاده کرد.

$$best(t) = \min_{j \in [1, \dots, m]} fit_j(t) \quad (14)$$

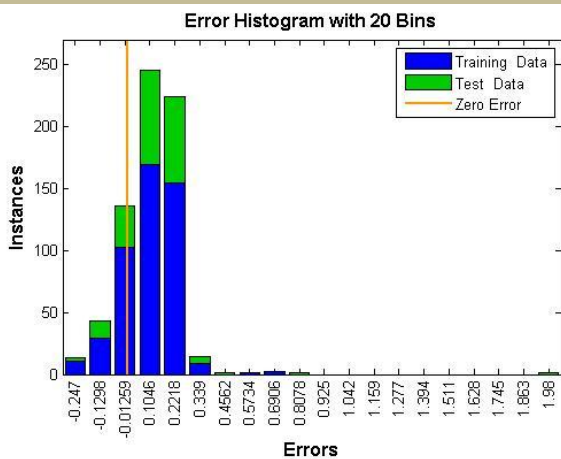
$$worst(t) = \max_{j \in [1, \dots, m]} fit_j(t) \quad (15)$$

در ابتدای تشکیل سیستم، هر جسم به صورت تصادفی در یک نقطه از فضا قرار می‌گیرد که جوابی از مسئله است. در هر لحظه از زمان، اجرام ارزیابی شده، سپس تغییر مکان هر جرم پس از محاسبه روابط (۴) تا (۱۵) محاسبه می‌شود. پارامترهای سیستم نیز در هر مرحله به‌روزرسانی می‌شوند. شرط توقف، رسیدن به حداکثر تعداد تکرار است. در شکل (۱) روندنمای GSA نشان داده شده است.

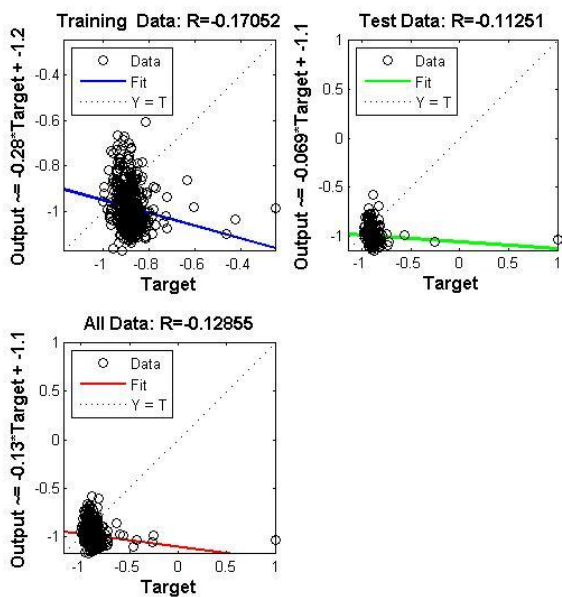
ویژگی‌های مثبت GSA همچون همگرایی سریع، عدم توقف در بهینه‌های محلی، کاهش حجم محاسباتی نسبت به الگوریتم‌های تکاملی و عدم نیاز به حافظه در مقایسه با دیگر الگوریتم‌های خانواده هوش جمعی، بستر جدیدی از تحقیقات را فراروی محققان قرار داده است. این الگوریتم تا به حال در کاربردهای متنوعی همچون طراحی فیلتر [۳۷]، انتخاب ویژگی [۳۸]، جبران‌سازی توان راکتو [۳۹] و آموزش شبکه عصبی [۴۰]



شکل (۲): میزان MSE در تکرارهای GSA



شکل (۳): میزان خطا در GSA



شکل (۴): تقابل میان خروجی شبکه و خروجی موردانتظار در GSA

بهتری دارند، جرم گرانشی بهتری نسبت داده می‌شود. در نتیجه هر جرم به اندازهٔ برآزش خود، سایر اجرام را به سمت خود جذب می‌کند. بنابراین با گذشت زمان، اجرام به سمت موقعیت‌های مناسب‌تر می‌روند [۳۵]. تغییر موقعیت یک جرم به معنای به‌روزرسانی وزن‌های شبکه در جهت کاهش MSE در تکرار جاری است.

در هر تکرار، تمام اجرام موقعیتشان را با محاسبهٔ سرعت جدید به‌روزرسانی و به سمت موقعیت جدید حرکت می‌کنند [۴۳]. از آنجاکه این الگوریتم حافظه‌دار نیست، هیچ‌یک از موقعیت‌های دیده‌شدهٔ قبلی را به‌خاطر نمی‌سپارد. با وجود این دیده می‌شود که عملکردی به‌خوبی الگوریتم‌های حافظه‌دار دارد [۳۵]. اما الگوریتم جستجوی گرانشی، پیچیدگی محاسباتی بالاتری نسبت به الگوریتم‌هایی مانند PSO داشته و از سوی دیگر پیچیدگی محاسباتی آن نسبت به الگوریتم وراثتی پایین‌تر است [۳۵]. یک راه برای ایجاد مصالحه بین بهره‌وری و کاوش، کاهش تعداد عامل‌ها با گذر زمان است. از این‌رو پیشنهاد می‌شود که مجموعه‌ای از عامل‌ها با جرم سنگین‌تر نیروی‌شان را بر دیگران وارد آورند. به‌رحال باید مراقب این راهکار باشیم؛ چراکه ممکن است کاوش را کاهش و بهره‌وری را افزایش دهد [۳۴].

برای جلوگیری از افتادن در نقاط بهینهٔ محلی، الگوریتم باید کاوش را در ابتدا در نظر بگیرد و با ادامهٔ تکرارها، بهره‌وری باید ظاهر شده و کاوش کم‌رنگ شود. برای افزایش کارایی GSA با کنترل کاوش و بهره‌وری، تنها تعداد k_{best} عامل، عوامل دیگر را جذب خواهند کرد. k_{best} تابعی از زمان است که مقدار اولیهٔ آن K_0 بوده و در گذر زمان کاهش می‌یابد. به‌عبارتی، در ابتدا تمام عوامل نیروی خود را به عوامل دیگر وارد می‌کنند و با گذر زمان k_{best} به‌صورت خطی کاهش یافته و در انتها، تنها یک عامل خواهد بود که به دیگران نیرو وارد می‌کند [۳۴].

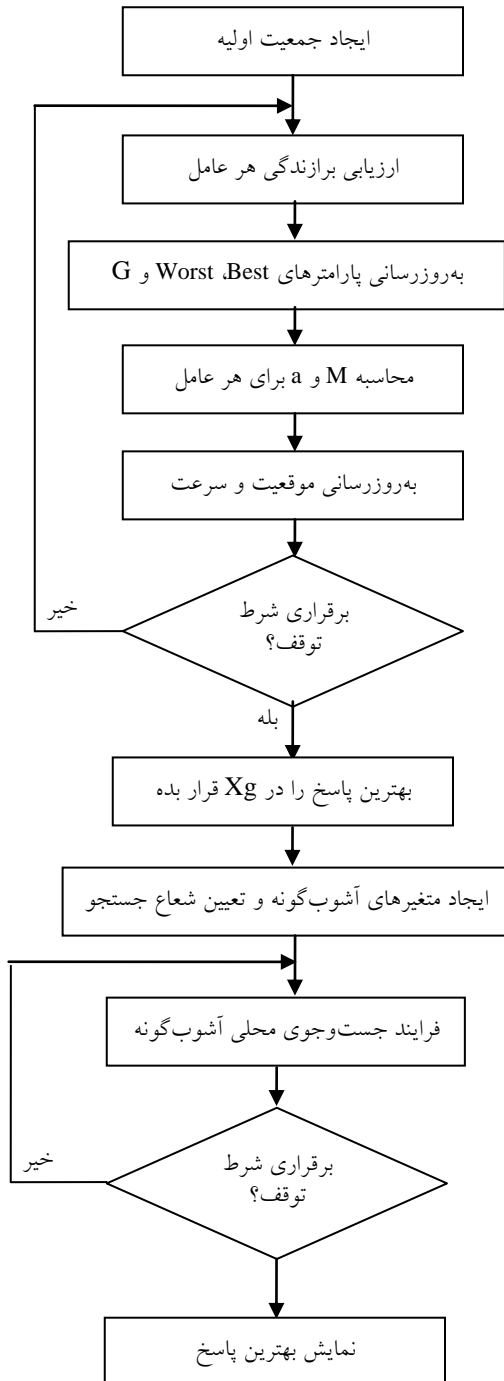
در این پژوهش، GSA برای آموزش شبکهٔ MLP مورد استفاده قرار گرفت که جمعیت اولیهٔ در آن برابر ۴۰ و پس از ۱۰۰ بار تکرار الگوریتم، بهترین جواب به‌دست‌آمده برای MSE مقدار ۰/۰۰۱۹ بود. نتایج خروجی این الگوریتم در شکل‌های (۲) تا (۴) نمایش داده شده است.

۶.۴ آموزش MLP با CGSA

- گام دوم: متغیرهای آشوب‌گونه برای تکرار بعدی طبق

رابطه (۲۰) تعیین می‌شوند:

$$Cx^{(k+1)} = \sin[a/Cx^{(k)}] \quad (20)$$



شکل (۵): روندنمای CGSA [۲۴]

از سال ۲۰۰۹ تاکنون، نسخه‌های متعددی از GSA معرفی شده است. CGSA در سال ۲۰۱۲ معرفی شد که در آن برای افزایش کارایی GSA، سازوکار آشوب به آن افزوده می‌شود که عبارت است از یک جستجوی محلی در اطراف جواب بهینه یافته‌شده توسط GSA [۲۴].

ایده اصلی جستجوی محلی آشوب‌گونه، استفاده از متغیرهای آشوب‌گونه است. تصادفی بودن متغیرهای آشوب‌گونه باعث می‌شود بهینه‌سازی آشوب‌گونه، نقاط بهینه سراسری را به سرعت بیابد [۲۴]. برای بهبود کارایی GSA در زمینه سرعت هم‌گرایی، جستجوی محلی آشوب‌گونه بر روی عامل بهینه یافته‌شده توسط GSA صورت می‌گیرد. به عبارتی پس از آنکه GSA، جواب بهینه (X_g) را یافت، از آنجاکه احتمال وجود جوابی بهتر از X_g در اطراف X_g بالا می‌باشد، جستجوی محلی آشوب‌گونه، به جستجوی اطراف X_g برای یافتن جوابی بهتر می‌پردازد. روندنمای CGSA در شکل (۵) نمایش داده شده است [۲۴]. روال جستجوی محلی آشوب‌گونه چنین است:

ابتدا متغیرهای آشوب‌گونه ساخته می‌شوند و سپس توسط آن‌ها عمل جستجوی محلی آغاز می‌شود. شعاع جستجو (σ) محدوده جستجوی محلی را تعیین می‌کند و در هر تکرار نسبت به تکرار قبلی کاهش می‌یابد. در CGSA، فرایند جستجوی محلی آشوب‌گونه در اطراف جواب بهینه یافته‌شده توسط فرایند GSA برای یافتن جوابی بهتر از آن صورت می‌پذیرد.

- گام اول: فرض می‌شود که $k=0$ و متغیرهای آشوب‌گونه به صورت تصادفی تعیین می‌شوند. سپس شعاع جستجو (σ) به صورت دلخواه و محدوده جستجوی (X_{min_i} و X_{max_i}) X_{best} و F_{best} توسط روابط (۱۶) تا (۱۹) تعیین می‌شوند:

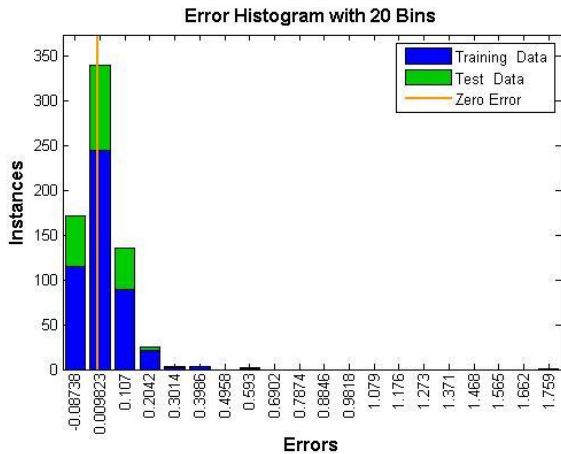
$$X_{max_i} = X_g + \sigma_i/2 \quad (16)$$

$$X_{min_i} = X_g - \sigma_i/2 \quad (17)$$

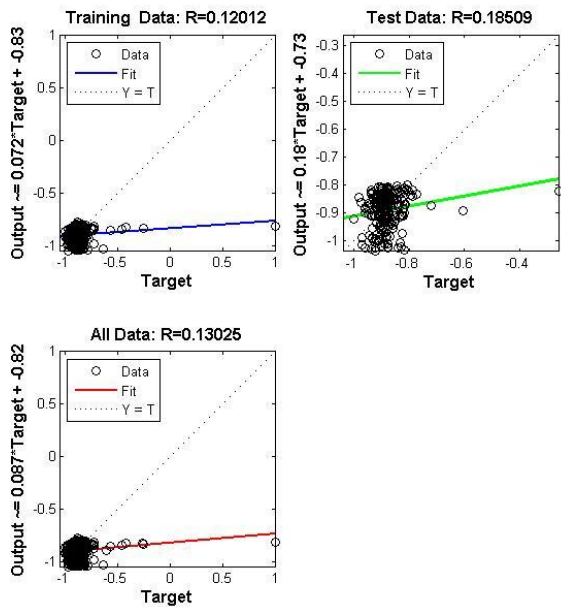
$$X_{best} = X_g \quad (18)$$

$$F_{best} = J_g \quad (19)$$

که در آن، J_g عبارت است از مقدار تابع برازندگی برای بهترین جواب یافته‌شده توسط GSA (X_g).



شکل (۷): میزان خطا در CGSA



شکل (۸): تقابل میان خروجی شبکه و خروجی موردانتظار در CGSA

۵. نتیجه گیری

باتوجه به ماهیت غیرخطی آلودگی هوا و تأثیرپذیری از پارامترهای طبیعی مختلف، پیش‌بینی میزان آلاینده‌ها بسیار دشوار بوده و نیاز به استفاده از پارامترهای هواشناسی و مکانی مختلفی دارد. برای این منظور، روابط و روش‌های گوناگونی پیشنهاد شده‌اند که هرکدام در یک منطقه یا در شرایطی خاص، نتایج مطلوبی ارائه کرده‌اند. مدل‌های پیش‌بینی آلودگی هوا وسیله بسیار مؤثری است که به کمک آن، می‌توان برای مدیریت آلودگی هوای کلان‌شهرها، تدابیر مناسب و کم‌هزینه طراحی کرد. لذا در این مقاله، سعی

- گام سوم: از رابطه (۲۱) برای به‌روزرسانی موقعیت اجرام استفاده می‌شود:

$$X_i^{(k+1)} = \frac{Xmax_i + Xmin_i}{2} + \frac{Xmax_i - Xmin_i}{2} Cx_i^k \quad (21)$$

- گام چهارم: موقعیت جدید اجرام، توسط تابع هدف (در این پژوهش، MSE) ارزیابی شده و در J_{k+1} نگهداری می‌شود.

- گام پنجم: اگر $J_{k+1} < F_{best}$ در آن صورت $X_{best} = X_{k+1}$ و $F_{best} = J_{k+1}$

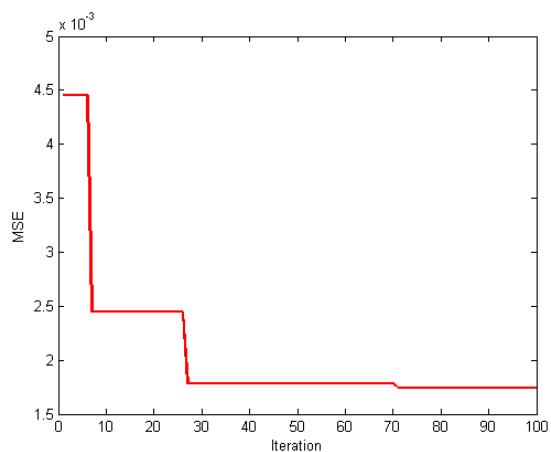
- گام ششم: شعاع جستجو طبق رابطه (۲۲) کاهش می‌یابد.

$$\sigma(k+1) = w \times \sigma(k) \quad (22)$$

که در آن، w یک عدد مثبت کوچک‌تر از یک است (برای مثال ۰/۹۸).

- گام هفتم: اگر تعداد تکرارهای مجاز به پایان نرسیده است، به گام دوم می‌رویم.

در این پژوهش، CGSA برای آموزش شبکه MLP مورد استفاده قرار گرفت که جمعیت اولیه در آن برابر ۴۰ و پس از ۱۰۰ بار تکرار الگوریتم، بهترین جواب به‌دست‌آمده برای MSE مقدار ۰/۰۰۱۷ بود که در مقایسه با GSA استاندارد، ۰/۰۰۰۲ بهبود داشت. نتایج خروجی این الگوریتم در شکل‌های (۶) تا (۸) نمایش داده شده است.



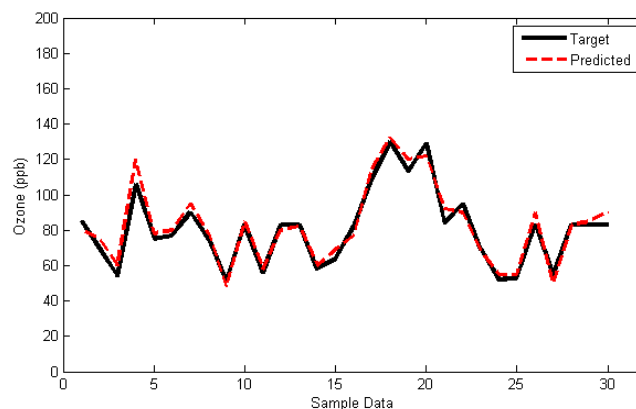
شکل (۶): میزان MSE در تکرارهای CGSA

انتخاب ویژگی‌های بهینه برای ورودی این سیستم پیش‌بین به‌کار گرفته شد. بدین ترتیب می‌توان با انتخاب گستره وسیعی از عوامل مؤثر در پیش‌بینی میزان غلظت آلاینده‌ها، به‌صورتی نظام‌مند و هوشمند عوامل مؤثرتر را انتخاب و با شبکه‌ای که از لحاظ ساختاری نیز حجیم و پیچیده نباشد، تخمین درست‌تری را ارائه کرد. برای نمونه، عملکرد مدل پیشنهادی در پیش‌بینی غلظت آلاینده O_3 که مقایسه‌ای شهودی از مقادیر پیش‌بینی شده و خروجی‌های واقعی است، در شکل (۹) نشان داده شده است.

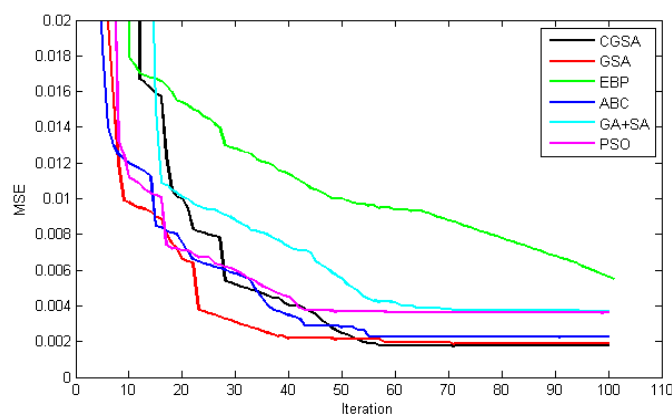
باتوجه به شکل (۹) عملکرد قابل قبول این مدل اثبات می‌شود. عملکرد پیش‌بینی‌کننده ارائه‌شده با پنج روش دیگر که در آن‌ها شبکه عصبی MLP با الگوریتم‌های پس‌انتشار خطا (EBP)، ABC، GA+SA، PSO و GSA استاندارد آموزش دیده بودند، مورد مقایسه قرار گرفت.

شد که با ارائه مدلی نوین برای پیش‌بینی میزان غلظت آلاینده‌های هوا، در این راستا گامی برداشته شود. شبکه‌های عصبی مصنوعی مدل‌های قدرتمندی برای پیش‌بینی آلاینده‌های هوا هستند. بسیاری از محققان تمایل دارند از این ابزار استفاده کنند، اما با مسئله آموزش شبکه عصبی مواجه می‌شوند. الگوریتم جست‌وجوی گرانشی از جمله روش‌های جدید برای الگوریتم‌های ابتکاری است که با الهام از قانون جاذبه و مفهوم جرم در طبیعت بنا نهاده شده است.

در این پژوهش، پیشنهاد استفاده از نسخه اصلاحی الگوریتم جست‌وجوی گرانشی به نام CGSA برای آموزش شبکه عصبی MLP به‌منظور پیش‌بینی میزان غلظت آلاینده O_3 در هوای شهر تهران ارائه شد. بدین ترتیب می‌توان با جست‌وجوی بهترین وزن‌ها برای شبکه عصبی، به پیش‌بینی‌کننده مناسب برای آلاینده‌های هوا دست یافت. همچنین نسخه دودویی این الگوریتم (BGSA) برای



شکل (۹): عملکرد مدل پیشنهادی در پیش‌بینی غلظت آلاینده ازن - مقایسه مقادیر پیش‌بینی و خروجی واقعی



شکل (۱۰): رفتار همگرایی به‌هنگام آموزش شبکه عصبی پیش‌بین توسط الگوریتم‌های EBP، ABC، GA+SA، PSO، GSA و CGSA در ۱۰۰ بار تکرار الگوریتم

جدول (۴): MSE در شرایط آموزش شبکه عصبی پیش‌بین توسط الگوریتم‌های GSA, PSO, GA+SA, ABC, EBP و CGSA

MSE	نام الگوریتم
۰/۰۰۵۵	پس‌انتشار خطا (EBP)
۰/۰۰۲۳	اجتماع زنبور عسل مصنوعی (ABC)
۰/۰۰۳۷	ترکیب وراثتی و شبیه‌سازی تبرید (GA+SA)
۰/۰۰۳۶	بهینه‌سازی ازدحام ذرات (PSO)
۰/۰۰۱۹	جستجوی گرانشی (GSA)
۰/۰۰۱۷	نسخه آشوب‌گونه جستجوی گرانشی (CGSA)
۰/۰۰۳۸	CGSA برای آموزش شبکه و عدم استفاده از BGSA برای انتخاب ویژگی

در این‌باره، میزان MSE در تخمین توسط هریک از روش‌های مذکور با فرض به‌کارگیری BGSA برای انتخاب ویژگی، در جدول (۴) ارائه شده است. همچنین MSE در شرایطی که BGSA برای انتخاب ویژگی به‌کار گرفته نشود و از CGSA برای آموزش شبکه استفاده شود، در سطر آخر جدول (۴) گزارش شده است. همان‌گونه که مشاهده می‌شود، در شرایط شبیه‌سازی مشابه، CGSA عملکرد بهتری را نتیجه می‌دهد. همچنین حذف ویژگی‌های با اهمیت کمتر توسط

BGSA منجر به بهبود عملکرد مدل پیش‌بین شده است. شکل (۱۰) نیز رفتار همگرایی روش‌های مذکور را طی ۱۰۰ بار تکرار نمایش می‌دهد. در این مورد نیز CGSA رفتار همگرایی را نشان می‌دهد.

باتوجه به اینکه ابزار پیش‌بینی در این مقاله، شبکه عصبی مصنوعی است و آموزش استاندارد این شبکه با الگوریتم EBP مشکل گرفتاری در کمینه‌های محلی را در برخی مسائل تقریب تابع و نیز طولانی بودن روال همگرایی به جواب را به دنبال دارد؛ لذا روش‌های بهینه‌سازی هوشمند که عملکرد بهتری را از خود نشان داده‌اند، در قالب یک مطالعه قیاسی در این مقاله مورد استفاده قرار گرفتند. نتایج نیز نشان داد که CGSA عملکرد موفق‌تری را از لحاظ میزان خطا نسبت به سایر روش‌ها دارد (جدول ۴). شایان ذکر است که پارامترهای این روش‌ها با توجه به مراجع مربوط در بازه مناسب پیشنهادی انتخاب شده؛ لذا بررسی این روش‌ها در قالب این مطالعه قیاسی به منظور برآورد عملکرد آن‌ها در رسیدن به هدف نهایی این پژوهش در پیش‌بینی موفق میزان غلظت آلاینده O₃ انجام شد.

مراجع

- [1] Sajani, S. Z., Scotto, F., Lauriola, P., Galassi, F., Montanari, A., "Urban air pollution monitoring and correlation properties between fixed-site stations", Air & Waste Management Association Vol. 54, No. 10, pp. 1236-1241, 2004.
- [2] اشرفی. خسرو، هشیاری پور. غلامعلی، نجار اعرابی. بایک، کشاورزی شیرازی. هما، «پیش‌بینی روزانه غلظت کربن منوکسید با استفاده از مدل تلفیقی انتخاب پیشرو - عصبی فازی براساس تحلیل پایداری جو، بررسی موردی: تهران»، مجله فیزیک زمین و فضا، دوره ۳۸، شماره ۲، ص ۱۸۳-۲۰۱، ۱۳۹۱.
- [3] منصور. غیاث‌الدین، «آلودگی هوا. منابع، اثرات و کنترل»، انتشارات دانشگاه تهران، چاپ اول، ۱۳۸۵.
- [4] Dempsey, F., "Observations of stratospheric O₃ intrusions in air quality monitoring data in Ontario, Canada", Atmospheric Environment, Vol. 98, pp. 111-122, Dec. 2014.
- [5] Ha, Q.P., Wahid, H., Duc, H., Azzi, M., "Enhanced radial basis function neural networks for ozone level estimation", Neurocomputing, Vol. 155, pp. 62-70, May 2015.
- [6] Rasmussen, D.J., Fiore, A.M., Naik, V., Horowitz, L.W., McGinnis, S.J., Schultz, M.G., "Surface ozone-temperature relationships in the Eastern US: a monthly climatology for evaluating chemistry-

- climate models*", Atmospheric Environment, Vol. 47, pp. 142-153, Feb. 2012.
- [7] Castellano, M., Franco, A., Cartelle, D., Febrero, M., Roca, E., "Identification of NO_x and ozone episodes and estimation of ozone by statistical analysis", Water, Air, & Soil Pollution, Vol. 198, No. 1, pp. 95-110, 2009.
- [8] Davis, J., Cox, W., Reff, A., Dolwick, P., "A comparison of CMAQ-based and observation-based statistical models relating ozone to meteorological parameters", Atmospheric Environment, Vol. 45, No. 20, pp. 3481-3487, 2011.
- [9] Agirre-Basurko, E., Ibarra-Berastegi, G., Madariaga, I., "Regression and multilayer perceptron-based models to forecast hourly O_3 and NO_2 levels in the Bilbao area", Environmental Modelling & Software, Vol. 21, No. 4, pp. 430-446, 2006.
- [10] Heo, J.S., Kim, D.S., "A new method of ozone forecasting using fuzzy expert and neural network systems", Science of The Total Environment, Vol. 325, No. 1-3, pp. 221-237, 2004.
- [11] Domanska, D., Wojtylak, M., "Explorative forecasting of air pollution", Atmospheric Environment, Vol. 92, pp. 19-30, Aug. 2014.
- [۱۲] علیاری شوره‌دلی. مهدی، تشنه‌لب. محمد، خاکی صدیق. علی، «پیش‌بینی کوتاه‌مدت آلودگی هوا به کمک شبکه‌های عصبی پرسپترون چندلایه، خط حافظه‌دار تأخیر، گاما و ANFIS با روش‌های ترکیبی آموزش مبتنی بر PSO»، مجله کنترل، جلد ۲، شماره ۱، ص ۱-۱۹، ۱۳۸۷.
- [۱۳] بوداقپور. سیامک، چرخستانی. امیر، «پیش‌بینی میزان غلظت آلاینده‌های هوای تهران با استفاده از شبکه عصبی مصنوعی»، علوم و تکنولوژی محیط زیست، دوره ۱۳، شماره ۱، ص ۱-۱۰، ۱۳۹۰.
- [14] Ozbay, B., Keskin, G.A., Dogruparmak, S.C., Ayberk, S., "Predicting tropospheric ozone concentrations in different temporal scales by using multilayer perceptron models", Ecological Informatics, Vol. 6, No. 3-4, pp. 242-247, 2011.
- [15] Salcedo-Sanz, S., Camacho, J.L., Pérez-Bellido, Á.M., Ortiz-Garcia, E.G., Portilla-Figuera, A., Hernández-Martín, E., "Improving the prediction of average total ozone in column over the Iberian Peninsula using neural networks banks", Neurocomputing, Vol. 74, No. 9, pp. 1492-1496, 2011.
- [16] Biancofiore, F., Verdecchia, M., Di Carlo, P., Tomassetti, B., Aruffo, E., Busilacchio, M., Bianco, S., Di Tommaso, S., Colangeli, C., "Analysis of surface ozone using a recurrent neural network", Science of The Total Environment, Vol. 514, pp. 379-387, May 2015.
- [17] Nayeb Yazdi, M., Delavarrafiee, M., Arhami, M., "Evaluating near highway air pollutant levels and estimating emission factors: Case study of Tehran, Iran", Science of The Total Environment, Vol. 538, pp. 375-384, Dec. 2015.
- [18] Givehchi, R., Arhami, M., Tajrishy, M., "Contribution of the Middle Eastern dust source areas to PM_{10} levels in urban receptors: Case study of Tehran, Iran", Atmospheric Environment, Vol. 75, pp. 287-295, Aug. 2013.
- [19] Arhami, M., Kamali, N., Rajabi, M.M., "Predicting hourly air pollutant levels using artificial neural networks coupled with uncertainty analysis by Monte Carlo simulations", Environmental Science Pollution Research, Vol. 20, No. 7, pp. 4777-4789, 2013.
- [20] Chan, K.Y., L. Jian, "Identification of significant factors for air pollution levels using a neural network based knowledge discovery system", Neurocomputing, Vol. 99, pp. 564-569, Jan. 2013.
- [21] Luna, A.S., Paredes, M.L.L., de Oliveira, G.C.G., Corrêa, S.M., "Prediction of ozone concentration in tropospheric levels using artificial neural networks and support vector machine at Rio de Janeiro, Brazil", Atmospheric Environment, Vol. 98, pp. 98-104, Dec. 2014.
- [22] Ju, F.Y., Hong, W.C., "Application of seasonal SVR with chaotic gravitational search algorithm in electricity forecasting", Applied Mathematical Modelling, Vol. 37, No. 23, pp. 9643-9651, 2013.
- [23] Li, C., Zhou, J., Xiao, J., Xiao, H., "Hydraulic turbine governing system identification using T-S fuzzy model optimized by chaotic gravitational search algorithm", Engineering Applications of

- Artificial Intelligence, Vol. 26, No. 9, pp. 2073-2082, 2013.
- [24] Li, C., Zhou, J., Xiao, J., Xiao, H., "Parameters identification of chaotic system by chaotic gravitational search algorithm", *Chaos, Solitons & Fractals*, Vol. 45, No. 4, pp. 539-547, 2012.
- [۲۵] ذوقی. محمدجواد، سعیدی. محسن، «پیش‌بینی روزانه غلظت دی‌اکسید گوگرد در یک ناحیه شهری با استفاده از شبکه عصبی مصنوعی»، دومین همایش و نمایشگاه تخصصی مهندسی محیط زیست، دانشگاه تهران، ۱۳۸۷.
- [26] Rashedi, E., Nezamabadi-pour, H., Saryazdi, S., "BGSA: Binary Gravitational Search Algorithm", *Natural Computing*, Vol. 9, No. 3, pp. 727-745, 2010.
- [27] Hooyberghs, J., Mensink, C., Dumont, G.; Fierens, F., Brasseur, O., "A neural network forecast for daily average PM₁₀ concentrations in Belgium", *Atmospheric Environment*, Vol. 39, No. 18, pp. 3279-3289, 2005.
- [28] Rojas, R., *Neural Networks: A Systematic Introduction*, Springer-Verlag, Berlin, 1996.
- [29] Karaboga, D. *An Idea Based on Honey Bee Swarm for Numerical Optimization*, Technical Report-TR06, Erciyes University, Oct. 2005.
- [30] Goldberg, D.E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Longman Publishing Co., Boston, MA, USA, 1989.
- [31] Kirkpatrick, S., Gelatto, C. D., Vecchi, M. P., "Optimization by simulated annealing", *Science*, Vol. 220, No. 4598, pp. 671-680, 1983.
- [32] Granville, V., Krivanek, M., Rasson, J.-P., "Simulated annealing: A proof of convergence", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 6, pp. 652-656, 1994.
- [33] Kennedy, J., Eberhart, R., Shi, Y., *Swarm Intelligence*, Morgan Kaufmann, San Francisco, CA, USA, 2001.
- [34] Rashedi, E., Nezamabadi-pour, H., Saryazdi, S., "GSA: A Gravitational Search Algorithm", *Information Sciences*, Vol. 179, No. 13, pp. 2232-2248, 2009.
- [۳۵] ده‌باشیان. مریم، ظهیری. سیدحمید، «آموزش شبکه عصبی MLP در فشرده‌سازی تصاویر با استفاده از روش GSA»، فصل‌نامه فناوری اطلاعات و ارتباطات ایران، سال دوم، شماره‌های ۵ و ۶، صفحه ۴۵-۵۳، ۱۳۸۹.
- [۳۶] قنبری. نجمه، رضوی. سیدمحمد، نبوی کهریزی. سید حسن، «ارائه یک روش هوشمند انتخاب ویژگی مبتنی بر الگوریتم جستجوی گرانشی باینری در سیستم بازشناسی ارقام دستنویس فارسی»، نشریه مهندسی برق و مهندسی کامپیوتر ایران، سال ۹، شماره ۱، ص ۲۹-۳۶، ۱۳۹۰.
- [۳۷] راشدی. عصمت، نظام‌آبادی‌پور. حسین، «طراحی فیلترهای IIR به وسیله الگوریتم جستجوی گرانشی»، شانزدهمین کنفرانس مهندسی برق ایران، دانشگاه تربیت مدرس، جلد اول، ص ۴۶۹-۴۷۴، تهران، اردیبهشت ۱۳۸۷.
- [۳۸] راشدی. عصمت، نظام‌آبادی‌پور. حسین، «انتخاب ویژگی با استفاده از الگوریتم جستجوی گرانشی»، سومین کنفرانس فناوری اطلاعات و دانش، دانشگاه فردوسی مشهد، ص ۵۱-۶۱، ۱۳۸۶.
- [39] Rashedi, E., Nezamabadi-pour, H., Saryazdi, S., "Allocation of static VAR compensator using gravitational search algorithm", First Joint Congress on Fuzzy and Intelligence Systems, Ferdowsi University of Mashhad, Iran, Aug 2007.
- [40] Sheikhan, M., Jadidi, Z., "Flow-based anomaly detection in high-speed links using modified GSA-optimized neural network", *Neural Computing and Applications*, Vol. 24, No. 3, pp. 599-611, 2014.
- [۳۶] ده‌باشیان. مریم، ظهیری. سیدحمید، «الگوریتم جستجوی گرانشی نخبه‌گرای پیشرفته»، اولین کنفرانس انرژی‌های تجدیدپذیر و تولید پراکنده ایران، دانشگاه بیرجند، ص ۳۵-۴۵، ۱۳۸۸.
- [۳۷] ده‌باشیان. مریم، ظهیری. سیدحمید، «MOGSA: روشی جدید در بهینه‌سازی چندهدفه مبتنی بر الگوریتم جستجوی گرانشی»، شانزدهمین کنفرانس ملی سالانه انجمن کامپیوتر ایران، دانشگاه صنعتی شریف، ص ۵۰۲-۵۰۷، ۱۳۸۹.

[۳۸] ده‌باشیان. مریم، ظهیری. سیدحمید، «آموزش شبکه

عصبی MLP در طبقه‌بندی داده‌ها با استفاده از روش

GSA»، نشریه مهندسی برق و مهندسی کامپیوتر ایران،

سال ۸، شماره ۳، ص ۲۶۷-۲۷۴، ۱۳۸۹.