

## Investigating air pollutants in Lorestan province and predicting their concentration using multi-layer neural network with stable online training (case studies: Khorramabad and Poldakhter)

Ahmadi, G.<sup>1</sup>  | Akbari, Z.<sup>2</sup>

1. Department of Mathematics, Payame Noor University.
2. Deputy Minister of Development and Meteorology of Lorestan Province.

**Corresponding Author E-mail:** [g.ahmadi@pnu.ac.ir](mailto:g.ahmadi@pnu.ac.ir)

(Received: 22 Jun 2024, Revised: 01 Sep 2024, Accepted: 08 Oct 2024, Published online: 08 Oct 2024)

### Abstract

With the uncontrolled expansion of large cities and the development of industries, air pollution has become a serious issue in urban management. In recent years, repeated droughts and widespread dam construction in neighboring countries have turned the problem of dust particles into a major challenge for Iran and other countries in the region. In many large cities, the concentration of air pollutants exceeds standards, which has widespread negative effects on human health, including an increase in cardiovascular and respiratory diseases, diabetes, hypertension, dementia, miscarriages, and premature deaths. Air pollution also negatively impacts the health of other living organisms, economic and social activities, agriculture, and the environment. The most important air pollutants include suspended particles with a diameter of less than 10 and 2.5 microns, nitrogen dioxide, nitrogen monoxide, sodium sulfate, carbon monoxide, and ozone. In this research, we first examine the issue of air pollution in Lorestan province based on data collected from the air pollution monitoring stations of the Environmental Protection Organization, and we calculate the correlation between the concentration of air pollutants and meteorological variables. Then, we use a multilayer neural network with a stable online learning algorithm to predict short-term pollutant concentration levels. Predicting air pollution is of high importance, and timely predictions can be effective in reducing the negative effects of air pollution. There are various methods for predicting air pollution, some of which have low accuracy. However, the use of artificial neural networks, as a more advanced method, has recently gained attention for modeling and predicting air quality. These models are especially useful for predicting time series and environmental data. Considering the quality of data collected at the air pollution monitoring stations of the Environmental Protection Organization in Lorestan province, the significant impacts of the dust phenomenon, and the geographical locations of the cities of Khorramabad and Poldokhtar, these cities were selected for case studies. According to the obtained results, there is a significant correlation between meteorological variables and pollutant concentrations, and the proposed neural network model demonstrates high accuracy in predicting pollutant concentrations.

**Keywords:** Air pollution, Time series forecasting, Air pollutant concentration forecasting, Multilayer neural network, Online stable learning algorithm

**Cite this article:** Ahmadi, G., & akbari, Z. (2024). Investigating air pollutants in Lorestan province and predicting their concentration using multi-layer neural network with stable online training (case studies: Khorramabad and Poldakhter). Journal of the Nivar, 48(126-127), 109-126. DOI: <https://doi.org/10.30467/nivar.2024.464234.1296>

**E-mail:** (2) [z.akbari91@yahoo.com](mailto:z.akbari91@yahoo.com)



## بررسی آلاینده‌های هوا در استان لرستان و پیش‌بینی غلظت آنها با استفاده از شبکه عصبی چندلایه با آموزش بهنگام پایدار (مطالعات موردی: خرم‌آباد و پلدختر)

قاسم احمدی<sup>۱</sup> | زینب اکبری<sup>۲</sup>

۱. استادیار گروه ریاضی، دانشگاه پیام‌نور، تهران، ایران.
۲. معاون توسعه و پیش‌بینی هواشناسی لرستان، خرم‌آباد، ایران.

رایانامه نویسنده مسئول: [g.ahmadi@pnu.ac.ir](mailto:g.ahmadi@pnu.ac.ir)

(دریافت: ۱۴۰۳/۰۴/۰۲، بازنگری: ۱۴۰۳/۰۶/۱۱، پذیرش: ۱۴۰۳/۰۷/۱۷، انتشار آنلاین: ۱۴۰۳/۰۷/۱۷)

### چکیده

با گسترش بی‌رویه شهرهای بزرگ و توسعه صنایع، مسئله آلودگی هوا به موضوعی جدی در مدیریت شهری تبدیل شده است. در سال‌های اخیر، خشکسالی‌های مکرر و سدسازی‌های گسترده در کشورهای همسایه، مسئله ریزگردها را به یک چالش بزرگ برای ایران و دیگر کشورهای منطقه تبدیل کرده است. در بسیاری از شهرهای بزرگ، غلظت آلاینده‌های هوا بالاتر از استانداردهاست که اثرات منفی گسترده‌ای بر سلامتی انسان‌ها، از جمله افزایش بیماری‌های قلبی عروقی، تنفسی، دیابت، فشار خون، زوال عقل، سقط جنین و مرگ‌های زودرس دارد. آلودگی هوا همچنین بر سلامت جانداران دیگر، فعالیت‌های اقتصادی و اجتماعی، کشاورزی، و محیط زیست نیز تأثیرات منفی می‌گذارد. مهم‌ترین آلاینده‌های هوا شامل ذرات معلق با قطر کمتر از ۱۰ و ۲/۵ میکرون، دی‌اکسید نیتروژن، مونوکسید نیتروژن، سولفات سدیم، مونوکسید کربن و ازن هستند. در این پژوهش، ابتدا موضوع آلودگی هوا در استان لرستان را بر مبنای داده‌های جمع‌آوری شده در ایستگاه‌های پایش آلودگی هوای سازمان محیط زیست بررسی و میزان همبستگی غلظت آلاینده‌های هوا و متغیرهای هواشناسی را محاسبه می‌کنیم. سپس، از شبکه عصبی چندلایه با یک الگوریتم یادگیری بهنگام پایدار برای پیش‌بینی کوتاه‌مدت میزان غلظت آلاینده‌ها استفاده می‌کنیم. پیش‌بینی آلودگی هوا از اهمیت بالایی برخوردار است و پیش‌بینی به‌موقع می‌تواند در کاهش اثرات منفی آلودگی هوا مؤثر باشد. روش‌های مختلفی برای پیش‌بینی آلودگی هوا وجود دارند که برخی از آن‌ها از دقت پایینی برخوردارند. اما استفاده از شبکه‌های عصبی مصنوعی به عنوان یک روش پیشرفته‌تر، در سال‌های اخیر برای مدل‌سازی و پیش‌بینی کیفیت هوا مورد توجه قرار گرفته است. این مدل‌ها به‌ویژه برای پیش‌بینی سری‌های زمانی و داده‌های محیطی کاربرد دارند. با توجه به کیفیت داده‌های جمع‌آوری شده در ایستگاه‌های پایش آلودگی هوای سازمان محیط زیست استان لرستان، تأثیرات عمده پدیده ریزگردها و موقعیت مکانی شهرهای خرم‌آباد و پلدختر، این شهرها جهت مطالعات موردی انتخاب شدند. مطابق نتایج به‌دست آمده، همبستگی معنی‌داری میان متغیرهای هواشناسی و میزان غلظت آلاینده‌ها وجود دارد و مدل عصبی پیشنهاد شده برای پیش‌بینی میزان غلظت آلاینده‌ها از دقت بالایی برخوردار است.

**کلیدواژه‌ها:** آلودگی هوا، پیش‌بینی سری‌های زمانی، پیش‌بینی غلظت آلاینده هوا، شبکه عصبی چندلایه، الگوریتم یادگیری بهنگام پایدار

۱. مقدمه  
موضوع آلودگی هوا برای بشریت از اهمیت ویژه‌ای برخوردار است. از زمان برپایی آتش توسط انسان آلودگی هوا وجود داشته، ولی از زمان انقلاب صنعتی بحث آلودگی هوا جدی‌تر شده است. در قرن بیستم، پس از مشکلات عدیده‌ای که در شهرهای صنعتی و شلوغ پیش آمد، آلودگی

**استناد:** احمدی، قاسم، و اکبری، زینب. (۱۴۰۳). بررسی آلاینده‌های هوا در استان لرستان و پیش‌بینی غلظت آنها با استفاده از شبکه عصبی چندلایه با آموزش بهنگام پایدار (مطالعات موردی: خرم‌آباد و پلدختر). مجله نیوار، ۴۸ (۱۲۶-۱۲۷)، ۱۰۹-۱۲۶. DOI: <https://doi.org/10.30467/nivar.2024.464234.1296>

رایانامه: (۲) [z.akbari91@yahoo.com](mailto:z.akbari91@yahoo.com)



بالای  $PM_{10}$  منجر به بستری شدن در بیمارستان و مواجهه با غلظت بالای  $PM_{2.5}$  به بیماریهای قلبی-عروقی منتج خواهد شد. معمولاً ذرات با قطر بیشتر از ۱۰ میکرون قادر نیستند به مسیرهای تنفسی وارد شوند یا این که در دستگاه تنفسی فوقانی جذب می‌شوند (حسنوند و ندافی، ۱۳۹۵). در خصوص آلاینده‌ی ازن، لازم است به این نکته اشاره کنیم که دو نوع ازن در جو وجود دارد: ازن خوب (لایه‌ی ازن) و ازن بد (ازن موجود در سطح زمین) (ساینفلد و پاندیس، ۱۹۹۷). در این جا، بحث ما راجع به ازن موجود در سطح زمین است که معمولاً در هوای آفتابی تولید می‌شود. حداکثر غلظت ازن در ساعات بعد از ظهر رخ می‌دهد و در کل یک آلاینده‌ی تابستانه است. دو نوع مهم از اکسیژنهای موجود در هوای آزاد  $NO$  و  $NO_2$  است که روی هم به  $NO_x$  معروفند.  $NO_2$  گازی سمی است و سمی بودن آن چندین برابر  $NO$  است (تیواری و همکاران، ۲۰۱۸).  $SO_2$  در غلظتهای زیاد بوی تند و محرکی دارد. اکسیدهای گوگرد پس از ترکیب با ذرات معلق زیانبارترین اثرات مرتبط با آلودگی هوا را ایجاد می‌کنند. گاز  $CO$  دارای زمان ماندگاری بالایی در حدود چند ماه در جو می‌باشد. معمولاً این گاز در زمستان دارای غلظت بیشتری است. منشأ این آلاینده حمل و نقل، کارخانجات ذوب فلزات، صنایع شیمیایی و زباله‌سوزی است.  $CO$  گازی سمی است و بافتهای بدن را از اکسیژن لازم بی‌بهره می‌کند (حسنوند و ندافی، ۱۳۹۵).

#### ۱-۲ طوفان گرد و غبار

یکی دیگر از منابع آلودگی هوای ریزگردها هستند. به طور کلی کشور ایران مستعد مواجهه با طوفانهای گرد و غبار است. زیرا بخش عمده‌ای از سرزمین ایران را بیابانهای خشک تشکیل داده‌اند و همچنین، در کشورهای همسایه و نزدیک ایران مانند عراق، عربستان و سوریه نیز بیابانهای فراوانی وجود دارد که مستعد خیزش گرد و غبار هستند. این گرد و غبارها از طریق بادهایی موسوم به باد شمال به ایران می‌رسند (راشکی، ۲۰۲۱). چنانچه در شکل ۱ نیز دیده می‌شود، مناطق غربی ایران مانند استان لرستان بیش از سایر مناطق ایران تحت تأثیر این طوفانها هستند. فعالیتهای کشور ترکیه در خصوص

هوا به عنوان موضوعی بسیار مهم در مدیریت شهری مطرح گردید. در سالهای اخیر، با توجه به خشکسالی‌های مکرر و اقدامات گسترده‌ی کشورهای همسایه‌ی ایران در زمینه‌ی سدسازی، موضوع ریزگردها به مسئله‌ای جدی برای مناطق مختلف کشور ایران و کشورهای همسایه بدل شده است. در حال حاضر، در بسیاری از شهرهای بزرگ، غلظت برخی از آلاینده‌های هوا بسیار بالاتر از استانداردهاست. آلودگی هوا اثرات متعددی بر سلامتی دارد و منجر به مشکلاتی مانند افزایش بروز بیماری‌های قلبی عروقی و تنفسی، دیابت و فشار خون، زوال عقل، خطر سقط جنین در بارداری، اختلالات روانی، مرگ‌های زودرس، کاهش امید به زندگی و کمتر شدن شادی شهرنشینان می‌شود. همچنین، آلودگی هوا می‌تواند بر سلامت سایر جانداران، فعالیت‌های اقتصادی و اجتماعی، کشاورزی و محیط زیست اثر بگذارد. سازمان بهداشت جهانی از مرگ پیش از موعد میلیون‌ها نفر در سال بر اثر آلودگی هوا گزارش داده است.

#### ۱-۱ آلاینده‌های هوا

هوا ترکیبی طبیعی از عناصر شیمیایی است که مهمترین آنها نیتروژن (۷۸٪) و اکسیژن (۲۱٪) است (غیاث‌الدین، ۱۳۹۶). علاوه بر این، ده‌ها ترکیب شیمیایی دیگر در هوای مناطق مسکونی به ویژه محلهای فعالیت انسان، تحت عنوان آلودگی وجود دارد که منشأ طبیعی دارند یا این که حاصل فعالیتهای انسانی است. همچنین، ممکن است از واکنش بین آلاینده‌های مختلف موجود در هوا در شرایط خاص یا تحت تأثیر اشعه خورشید به وجود بیایند. مهمترین آلاینده‌های هوا عبارتند از:

۱. ذرات معلق با قطر کمتر از ۱۰ میکرون ( $PM_{10}$ )
۲. ذرات معلق با قطر کمتر از ۲/۵ میکرون ( $PM_{2.5}$ )
۳. دی اکسید نیتروژن ( $NO_2$ )
۴. مونو اکسید نیتروژن ( $NO$ )
۵. سولفات سدیم ( $SO_2$ )
۶. مونواکسید کربن ( $CO$ )
۷. ازن ( $O_3$ )

بخش مهمی از آلاینده‌ها با بیشترین تأثیرات منفی روی سلامت عمومی ذرات معلق در هوا هستند. مواجهه با غلظت

در استانهای غربی بسیار جدی است و لازم است مطالعات و اقدامات بیشتری در خصوص جلوگیری، پیش‌بینی و کنترل آن صورت پذیرد.

سدسازی‌ها بر رودخانه‌های مشترک با کشور عراق، به طور مداوم بر خشک شدن باتلاقهای عراق و افزایش بیابانها تأثیر می‌گذارد. بنابراین، مسأله‌ی آلودگی هوا در ایران و به ویژه



شکل ۱: بادهای مؤثر در ایجاد گرد و غبار در لرستان (راشکی، ۲۰۲۱).

سری‌های زمانی استفاده کرده‌اند. چن و همکارانش (۲۰۰۶) شبکه‌های عصبی موجک خطی محلی را برای پیش‌بینی سری‌های زمانی به کار گرفته‌اند. قلی پور و همکارانش (۲۰۰۶) از مدل‌های عصبی فازی برای پیش‌بینی سری‌های زمانی آشوبناک استفاده کرده‌اند. فاروک (۲۰۱۰) از شبکه‌های عصبی برای پیش‌بینی سری زمانی کیفیت آب بهره برده است. اخیراً مدل‌های یادگیری عمیق برای پیش‌بینی سری‌های زمانی چندگامی استفاده شده‌اند (چاندر، ۲۰۲۱). تیلاب (۲۰۱۸) آثار جدید مرتبط با استفاده از شبکه‌های عصبی برای پیش‌بینی سری‌های زمانی غیرخطی را بررسی کرده است. اخیراً، احمدی و دهقاندار (۲۰۲۳) از شبکه‌های عصبی نامتوازن برای پیش‌بینی سری‌های زمانی آشوبی استفاده کرده‌اند.

چندین مطالعه در خصوص پیش‌بینی آلودگی هوا با استفاده از شبکه‌های عصبی صورت گرفته است. سالاس و همکارانش (۲۰۰۵) یک مدل NARX با استفاده از شبکه‌های عصبی برای پیش‌بینی آلاینده  $PM_{10}$  در شهر سانتیاگو طراحی کرده‌اند. کوردوا و همکارانش (۲۰۲۱) به پیش‌بینی آلودگی هوا در شهر لیما، پایتخت کشور پرو، با

### ۳-۱. پیش‌بینی آلودگی هوا

پیش‌بینی به موقع آلودگی هوا می‌تواند به پیشگیری از تأثیرات منفی آن در اکثر موارد منجر شود. در سال‌های اخیر، روشهای پیش‌بینی متفاوتی برای درک چگونگی رفتار آلاینده‌ها در هوا مطرح شده است. به طور کلی این روشها از دقت پایینی برخوردار هستند. مثالهایی از این مدلها، مدل انجمن کیفیت هوای چندمقیاسی و مدل تحقیق و پیش‌بینی آب و هوا جفت شده با شیمی هستند که به ترتیب توسط چن و همکارانش (۲۰۰۶) و سایده و همکارانش (۲۰۱۱) پیشنهاد شده‌اند. از طرف دیگر، روشهای دیگری مانند شبکه‌های عصبی مصنوعی وجود دارند که به نظر می‌رسد برای مدل کردن و پیش‌بینی کیفیت هوا مناسب‌تر هستند. این مدلها به طور گسترده‌ای برای پیش‌بینی سری‌های زمانی و داده‌های محیطی به کار می‌روند. هدف ما در این کار، مطالعه‌ی قابلیت‌های شبکه‌های عصبی در پیش‌بینی میزان آلودگی هوا در استان لرستان است.

اثربخشی شبکه‌های عصبی مصنوعی در پیش‌بینی سری‌های زمانی، در بسیاری از آثار نشان داده شده است. فرانک و همکارانش (۲۰۰۱) از شبکه‌های عصبی برای پیش‌بینی

ماشینهای یادگیری حدی نامتوازن در شناسایی سیستم های زمان-پیوسته ارائه شده است (احمدی، ۲۰۱۹). در اینجا، از یک الگوریتم یادگیری بهنگام مبتنی بر قضیه پایداری لیاپانوف زمان-پیوسته برای آموزش مدل پرسپترون چندلایه بهره می‌گیریم.

در این مقاله، ما به مطالعه‌ی مسئله‌ی آلودگی هوا در استان لرستان می‌پردازیم. مبنای این مطالعه داده‌های جمع‌آوری شده در ایستگاه‌های پایش آلودگی هوای استان لرستان است. با توجه به کمیت و کیفیت بهتر داده‌ها در شهرهای خرم‌آباد و پلدختر، این شهرها به عنوان مطالعات موردی در نظر می‌گیریم. با استفاده از داده‌های هواشناسی در این شهرها به محاسبه‌ی میزان همبستگی بین غلظت آلاینده‌ها و متغیرهای هواشناسی می‌پردازیم. سپس، با استفاده از شبکه عصبی چندلایه میزان آلاینده‌های هوا در این شهرها را برای ساعات آینده پیش‌بینی می‌کنیم. برای تنظیم پارامترهای این شبکه از یک الگوریتم یادگیری بهنگام پایدار بهره می‌گیریم. آلاینده‌های مدنظر ما در این پژوهش، عبارتند از: ذرات معلق با قطر کمتر از ۱۰ و ۲/۵ میکرون، دی اکسید نیتروژن، مونو اکسید نیتروژن، سولفات سدیم، مونواکسید کربن و ازن.

ادامه‌ی این مقاله به صورت زیر است: در بخش ۲، مراحل روش تحقیق را به تفصیل بیان می‌کنیم. در بخش ۳، به محاسبه‌ی میزان همبستگی آلاینده‌ها و متغیرهای هواشناسی می‌پردازیم و با استفاده از شبکه‌های عصبی، میزان آلاینده‌های هوا را برای پنج ساعت آینده پیش‌بینی می‌کنیم. در بخش ۴، در خصوص نتایج به دست آمده در این پژوهش بحث می‌کنیم. در بخش ۵، به جمع‌بندی مطالب و نتیجه‌گیری می‌پردازیم.

## ۲. روش تحقیق

در این کار، ما از روش کشف دانش از پایگاه‌های داده برای به دست آوردن اطلاعات مرتبط با آلاینده‌های هوا و پیش‌بینی آنها استفاده می‌کنیم. هدف اصلی روش کشف دانش از پایگاه‌های داده استخراج اطلاعات ناشناخته و بالقوه مفید از داده‌های خام ذخیره شده در پایگاه‌های داده است (کوردوا،

استفاده از شبکه‌های عصبی پرداخته‌اند. در ایران نیز در سالهای اخیر، تحقیقاتی در خصوص پیش‌بینی آلودگی هوا با استفاده از شبکه‌های عصبی صورت گرفته است. ملکی و همکارانش (۲۰۱۹) به پیش‌بینی آلودگی هوا در شهر اهواز با استفاده از شبکه‌های عصبی پرداخته‌اند. صادقی و خاکسار آستانه (۱۳۹۳) به پیش‌بینی کوتاه‌مدت آلودگی ذرات معلق شهر اهواز با بهره‌گیری از شبکه‌های عصبی پرداخته‌اند. سجادیان (۱۳۹۴) به پیش‌بینی آلودگی هوای ناشی از حمل و نقل شهری کلانشهر تهران با بهره‌گیری از تلفیق GIS با مدل LUR و شبکه عصبی مصنوعی پرداخته است.

قائمی و همکارانش (۱۳۹۶) پیش‌بینی آلودگی هوای شهر تهران را با استفاده از ترکیب شبکه عصبی فازی- تطبیقی و آنالیز مؤلفه اصلی مورد مطالعه قرار داده‌اند. در پژوهش دیگری، پیش‌بینی آلودگی  $PM_{10}$  با استفاده از مدل‌های پرسپترون چندلایه و زنجیره مارکف مورد توجه قرار گرفته است (زنگویی و اسداله فردی، ۱۳۹۶). اخیراً، دلاور و همکارانش (۲۰۱۹) از روشهای مبتنی بر یادگیری ماشین برای پیش‌بینی میزان آلاینده‌های  $PM_{10}$  و  $PM_{2.5}$  در شهر تهران استفاده کرده‌اند. همچنین، توسعه‌ی مدل ترکیبی بر پایه شبکه عصبی جهت مدلسازی کیفیت هوای داخلی بازار تبریز به لحاظ ذرات معلق صورت گرفته است (کفاش چرندابی و غلامی، ۱۴۰۱). بشردوست و مسگری (۱۴۰۲) از شبکه‌های عصبی کانولوشن برای پیش‌بینی میزان آلاینده‌های  $PM_{10}$  و  $PM_{2.5}$  در شهر تهران استفاده کرده‌اند. در خصوص پیش‌بینی آلودگی هوای استان لرستان، تا جایی که ما می‌دانیم تاکنون تحقیقات خاصی صورت نگرفته است.

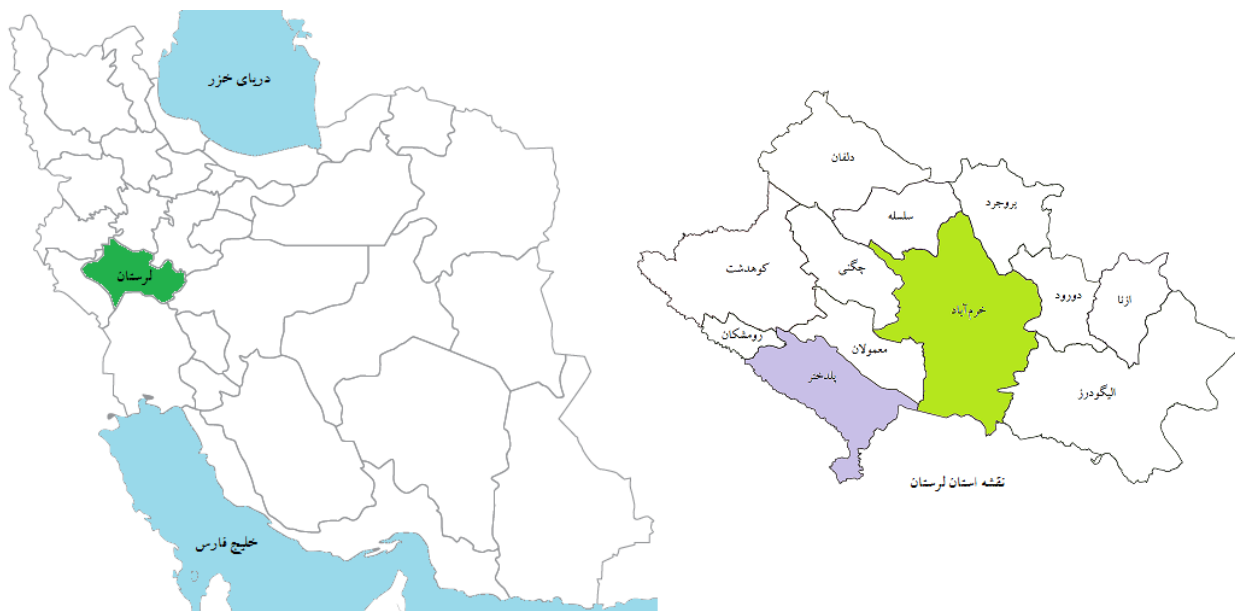
آموزش شبکه‌های عصبی در کاربردها به وسیله‌ی الگوریتمهای یادگیری صورت می‌پذیرد. بدین منظور، الگوریتم‌های فراوانی توسط محققین طراحی شده است. یکی از مباحث مهم در خصوص الگوریتمهای یادگیری، پایداری آنهاست. اخیراً، احمدی و تشنلب (۲۰۱۷) یک الگوریتم یادگیری بهنگام مبتنی بر قضیه پایداری لیاپانوف برای شبکه عصبی نامتوازن (راف) در شناسایی سیستم‌های غیرخطی ارائه کرده‌اند. همچنین، یک الگوریتم مبتنی بر لیاپانوف برای

صورت گرفته در سال ۱۴۰۰، برابر ۱۷۹۱۰۰۰ نفر است (احمدی و همکاران، ۱۴۰۰). با توجه به پدیده‌ی ریزگردها و نزدیکی این استان به مرزهای غربی و بیابانهای عراق، عربستان و سوریه، بحث آلودگی هوا مشکلی نسبتاً جدی برای این استان است. شهر خرم‌آباد با جمعیتی بالغ بر ۴۰۰۰۰۰ نفر در میان رشته کوههای زاگرس محصور گردیده است. لذا، به نظر می‌رسد آلاینده‌ها در هوای این شهر انباشته می‌شوند و با توجه به منابع متعدد برای آلاینده‌ها، موضوع آلودگی هوا در این شهر اهمیت زیادی دارد. شهر پلدختر، واقع در ناحیه جنوب غربی استان لرستان، بیش از اکثر شهرهای این استان در معرض ریزگردهاست. همچنین، با توجه به وجود چاهها و پالایشگاههای نفت در اطراف شهر پلدختر این شهر بیش از سایر شهرهای استان لرستان مستعد وجود برخی از آلاینده‌هاست (چاوشی و همکاران، ۱۳۹۰). از سویی دیگر، داده‌های جمع‌آوری شده در ایستگاههای مستقر در این شهرها (خرم‌آباد و پلدختر) از کمیت و کیفیت بهتری برخوردار هستند. بنابراین، ما این شهرها را به منظور مطالعه موردی انتخاب می‌کنیم.

۲۰۲۱). در این رویکرد، مدل‌ها می‌توانند به عنوان مثال، یک یا چند ساعت جلوتر، کیفیت هوا را پیش‌بینی کنند و در تصمیم‌گیری مدیریت شهری مؤثر واقع شوند. روش کشف دانش از پایگاه‌های داده دارای مراحل زیر است: الف) مطالعه‌ی پدیده؛ ب) مطالعه‌ی داده‌ها؛ ج) آماده‌سازی داده‌ها؛ د) مدل‌سازی؛ ه) ارزیابی. در بخش‌های زیر، ما هر مرحله از فرآیند را توضیح می‌دهیم.

## ۱-۲. مطالعه‌ی پدیده

در گام اول، ما پدیده‌ی آلودگی هوا در استان لرستان را مورد بحث قرار می‌دهیم. تمرکز اصلی بر پیش‌بینی آلودگی هوا به منظور پشتیبانی از تصمیم‌گیری‌های مربوط به سیاست‌های کاهش آلودگی است. بدین منظور از شبکه‌های عصبی مصنوعی برای پیش‌بینی میزان آلاینده‌های هوا در ساعات آینده استفاده می‌کنیم. استان لرستان با مساحت تقریبی ۲۸۱۵۷ کیلومترمربع در ناحیه جنوب غربی ایران واقع شده است (شکل ۲). این استان شامل ۱۱ شهرستان می‌باشد و مرکز آن شهر خرم‌آباد است. جمعیت این استان بر اساس سرشماری

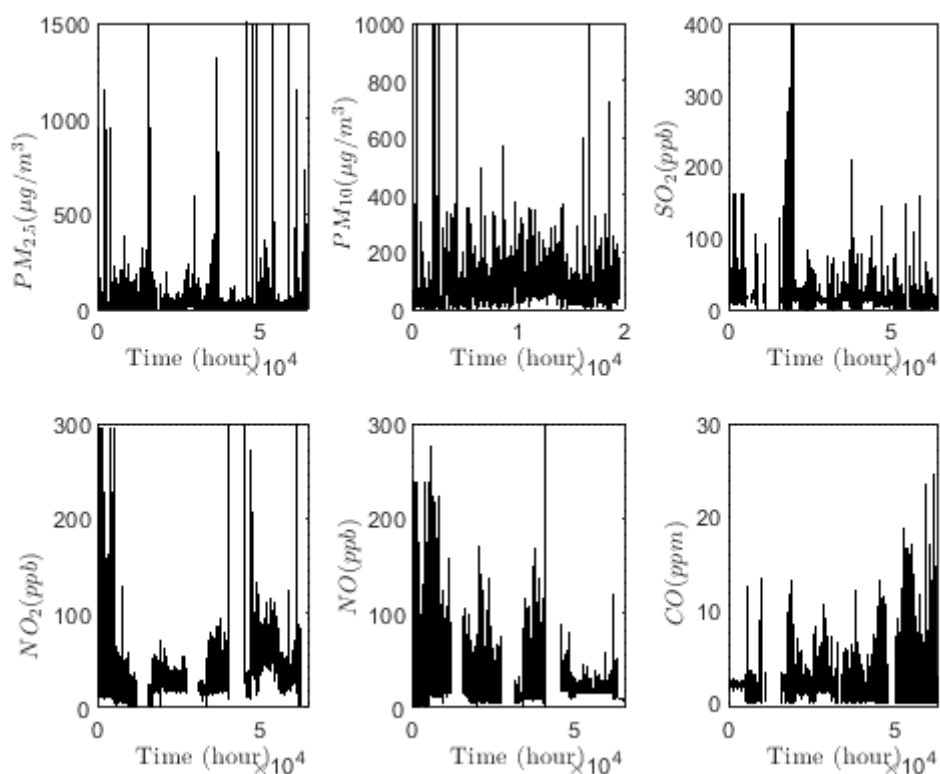


شکل ۲: نقشه استان لرستان در غرب کشور ایران

## ۲-۲. مطالعه داده‌ها

هدف این پژوهش، بررسی داده‌های مربوط به آلودگی هوا در استان لرستان، محاسبه میزان همبستگی غلظت آلاینده‌های هوا و متغیرهای هواشناسی و پیش‌بینی کوتاه‌مدت میزان آلاینده‌های هواست. با توجه به دلایل مطرح شده در بخش قبل، داده‌های جمع‌آوری شده از ایستگاههای پایش کیفیت هوا در شهرهای خرم‌آباد و پلدختر را مورد مطالعه قرار می‌دهیم. جهت بررسی میزان همبستگی غلظت آلاینده‌های هوا و متغیرهای هواشناسی، از داده‌های جمع‌آوری شده توسط ایستگاه‌های اداره هواشناسی استان لرستان در همین شهرها (خرم‌آباد و پلدختر) استفاده می‌کنیم. داده‌های آلودگی هوا به صورت ساعتی در دوره‌ی زمانی از ابتدای سال ۱۳۹۵ تا آذرماه ۱۴۰۲ جمع‌آوری شده‌اند. این داده‌ها شامل آلاینده‌های هوا به شرح زیر هستند:  $PM_{10}$ ,

$PM_{2.5}$ ,  $NO_2$ ,  $NO$ ,  $SO_2$ ,  $O_3$  و  $CO$ . متأسفانه، داده‌های از دست رفته و داده‌های نادرست زیادی وجود دارد و این موضوع کار را برای مدلسازی سخت می‌کند. به منظور غلبه بر این کاستیها، روشهایی را به کار می‌گیریم که در بخشهای بعد توضیح داده می‌شود. داده‌های خام ساعتی جمع‌آوری شده در ایستگاه پایش کیفیت هوا مستقر در دانشگاه علوم پزشکی استان لرستان مربوط به دوره زمانی از ۱ فروردین ۱۳۹۵ تا ۲۰ آذر ۱۴۰۲ در شکل ۳ نشان داده شده است. جدول ۱ برخی از مشخصات آماری این داده‌ها را نشان می‌دهد. در این جدول، میانگین داده‌ها و انحراف معیار آنها نشان داده شده است. داده‌های مربوط به هواشناسی در خرم‌آباد و پلدختر، به صورت روزانه از ابتدای سال ۱۳۹۵ تا پایان سال ۱۴۰۱ جمع‌آوری شده‌اند. این داده‌ها شامل بارش، فشار هوا، رطوبت نسبی، دما، سرعت و جهت باد می‌باشد.



شکل ۳: داده‌های خام (ساعتی) ایستگاه پایش کیفیت هوا در دانشگاه علوم پزشکی خرم‌آباد مربوط به دوره زمانی از ۱ فروردین ۱۳۹۵ تا ۲۰ آذر ۱۴۰۲

جدول ۱. مشخصات آماری داده‌های شهر خرم‌آباد مربوط به دوره زمانی از ۱ فروردین ۱۳۹۵ تا ۲۰ آذر ۱۴۰۲

متغیر	واحد	میانگین	انحراف معیار
$PM_{2.5}$	$\mu g/m^3$	29.05	61.91
$PM_{10}$	$\mu g/m^3$	85.10	76.07
$SO_2$	$ppb$	14.35	47.46
$NO_2$	$ppb$	34.96	43.08
$NO$	$ppb$	22.01	18.73
$O_3$	$ppb$	51.38	72.21
$CO$	$ppm$	1.50	1.30

### ۲-۳. آماده‌سازی داده‌ها

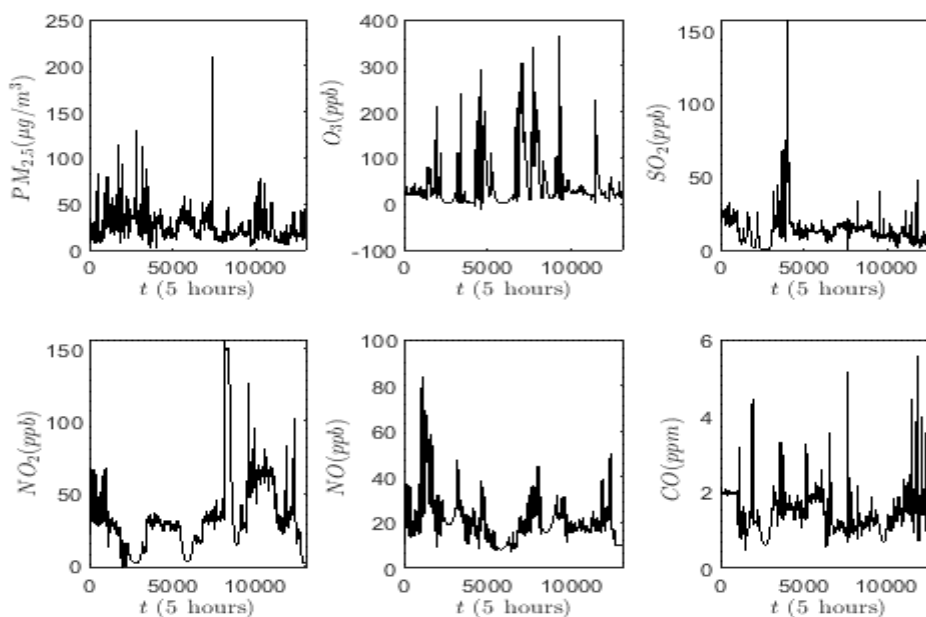
بدیهی است که این داده‌ها بدون پیش‌پردازش‌های ضروری، برای انجام پیش‌بینی مناسب نیستند. آماده‌سازی داده‌ها دارای گام‌های مختلفی است. ابتدا، ما باید داده‌های از دست رفته را تعیین تکلیف کنیم. در این جا، ما برای جایگزینی آنها با مقادیری مناسب از درون‌یابی استفاده می‌کنیم. داده‌های به دست آمده از ایستگاه‌های پایش شامل دنباله‌ای از مقادیر مشاهده شده به صورت  $\{x_t\}$  است که در زمان  $t$  ضبط شده‌اند. در اینجا داده‌های مربوط به سریهای زمانی در بازه‌های ساعتی جمع‌آوری شده‌اند. علاوه بر این، متوازن کردن برخی از داده‌ها نیز ضروری است. در اینجا، ما حدود تقریبی مربوط به هر کدام از متغیرها را مشخص و سپس، مقادیر خیلی بزرگ یا خیلی کوچک را با حدود تعیین

شده جایگزین می‌کنیم. سپس، همه‌ی داده‌ها را با استفاده از

رابطه‌ی زیر در بازه‌ی  $[0,1]$  نرمال می‌کنیم:

$$X_t = \frac{x_t - \min\{x_t\}}{\max\{x_t\} - \min\{x_t\}} \quad (1)$$

در این رابطه،  $X_t$  مقدار نرمال شده در زمان  $t$ ،  $x_t$  مقدار مشاهده شده در زمان  $t$ ،  $\min\{x_t\}$  مقدار می‌نیم و  $\max\{x_t\}$  مقدار ماکسیمم در مجموعه داده‌هاست. به منظور هموارسازی سریهای زمانی، میانگین پنج داده‌ی متوالی را به دست آورده، پیش‌بینی را برای سری زمانی حاصل شده انجام می‌دهیم. گذشته از تمام موارد مطرح شده، از یک فیلتر نیز برای دستیابی به مجموعه داده‌هایی با نمودار نسبتاً هموار استفاده می‌کنیم. شکل ۴ داده‌های آلودگی هوا را پس از فیلتر کردن با فیلتر باترورث ( $\omega_n = 0.05$ ) نشان می‌دهد.

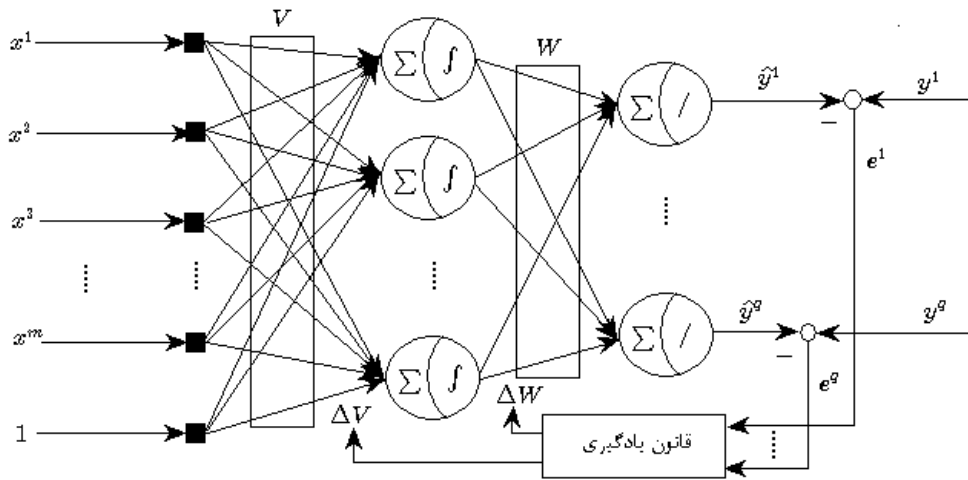


شکل ۴: داده‌های ساعتی ایستگاه پایش کیفیت هوا در دانشگاه علوم پزشکی خرم‌آباد پس از عملیات آماده‌سازی داده‌ها مربوط به دوره زمانی از ۱ فروردین ۱۳۹۵ تا ۲۰ آذر ۱۴۰۲

## ۲-۴. مدل سازی با شبکه‌های عصبی

در سالهای اخیر، شبکه‌های عصبی مصنوعی توجه زیادی را در مهندسی و علوم به خود جلب کرده‌اند. این مدل‌های غیرخطی قادر به یادگیری از داده‌ها هستند. پیش‌بینی سری‌های زمانی یکی از مهم‌ترین مسائل در بسیاری از زمینه‌ها از جمله علوم اقتصادی، علوم پزشکی، علوم اجتماعی و مهندسی می‌باشد (مونت‌گومری و همکاران، ۲۰۰۸). ماهیت فرآیندهای واقعی و اثرات عدم قطعیت‌ها، پیش‌بینی سری‌های زمانی را به موضوعی چالش‌برانگیز تبدیل کرده است. محققان سعی در حل این مشکل با استفاده از رویکردهای قدرتمند و جذاب جدیدی مانند شبکه‌های عصبی مصنوعی دارند که اخیراً در زمینه‌ی علوم کامپیوتر ارائه شده است.

در طراحی شبکه‌های عصبی مصنوعی از مغز و سیستم عصبی انسان الگوبرداری شده است (گوپتا، ۲۰۰۳). نورون‌ها واحدهای بنیادی در مغز و سیستم عصبی هستند و اطلاعات توسط آنها پردازش می‌شود. در شبکه‌های عصبی مصنوعی، یک نورون مدلی ریاضی از نورون بیولوژیکی است که شامل مجموع ورودی‌های وزنی، بایاس‌ها و توابع فعال‌ساز غیرخطی می‌باشد. نورون‌ها به یکدیگر متصل می‌شوند و پارامترها با استفاده از برخی الگوریتم‌های یادگیری به گونه‌ای تنظیم می‌شوند که بتوانند توابع غیرخطی را تقریب بزنند. اخیراً، از شبکه‌های عصبی برای حل بسیاری از مسائل غیرخطی در علوم کاربردی و مهندسی استفاده شده است. ساختار یک شبکه عصبی چندلایه در شکل ۵ به تصویر کشیده شده است.



شکل ۵: ساختار مدل پرسپترون چندلایه

در این رابطه،  $T$  تعداد مراحل زمانی است. برای  $d = 1$ ، پیش‌بینی یک گام جلوتر و برای  $d > 1$ ، پیش‌بینی چندگامی انجام می‌شود.

### ۲-۴-۲. پیش‌بینی سریهای زمانی با استفاده از

#### شبکه‌های عصبی

تغییرات زمانی در سریهای زمانی را می‌توانیم در یک مدل فضای حالت در سیستم‌های دینامیکی تعریف کنیم (احمدی و دهقاندار، ۲۰۲۳). مشاهدات سریهای زمانی  $\{x(t)\}_{t=0}^{\infty}$  را می‌توانیم به بردارهای حالت  $\{z(t)\}_{t=0}^{\infty}$  تبدیل کنیم که در آن  $z(t) \in \mathbb{R}^n$  حالت سیستم را نشان می‌دهد. دینامیک‌های این حالات با استفاده از رابطه‌ی زیر توصیف می‌شود:

### ۲-۴-۱. پیش‌بینی سری‌های زمانی

سری زمانی مجموعه‌ای از داده‌ها به صورت  $\{x_t\}_{t=0}^{\infty}$  می‌باشد که در آن،  $t$  نشان‌دهنده شاخص زمان است. می‌توانیم فرض کنیم که  $x_t$  یک تابع پیوسته با متغیر  $t$  است. در فرآیندهای طبیعی، برای دستیابی به یک مجموعه داده‌ی گسسته، از داده‌ها نمونه‌برداری می‌شود. در پیش‌بینی سری‌های زمانی با استفاده از شبکه‌های عصبی، مقادیر آینده با استفاده از مقادیر گذشته پیش‌بینی می‌شوند. در واقع سعی می‌کنیم تابع  $f$  را به گونه‌ای پیدا کنیم که

$$x_{t+d} = f(x_t, x_{t-1}, \dots, x_{t-T}) \quad (2)$$

کمترین خطای ممکن را در تقریب مقادیر واقعی داشته باشد. در این جا، ما از یک الگوریتم یادگیری بهنگام طراحی شده بر مبنای نظریه پایداری لیاپانوف استفاده می کنیم.

**قضیه ۱.** فرض کنیم مدل پرسپترون چندلایه می تواند تابع  $g$  در رابطه (۳) را تقریب بزند و از قوانین یادگیری زیر برای تنظیم پارامترهای این شبکه استفاده کنیم:

$$\dot{W}(t) = e(t)\varphi(\hat{V}(t)^T x(t))\Gamma_1^{-1} \quad (۹)$$

$$\dot{\hat{V}}(t) = \Gamma_2^{-1} \left( \varphi'(\hat{V}(t)^T x(t)) \right)^T \hat{W}(t)^T e(t)x(t)^T \quad (۱۰)$$

که در آن  $e(t)$  خطای مدل می باشد و  $\Gamma_1$  و  $\Gamma_2$  ماتریسهای قطری شامل نرخهای یادگیری هستند. فرض کنیم

$$\|e(t)\| \geq \frac{\|\zeta\|}{|\lambda_{\min}(A)|}$$

که در آن  $\lambda_{\min}(A)$  کوچکترین مقدار ویژه  $A$  و  $\zeta$  دینامیکهای مدل نشده است. در این صورت، خطای پیش بینی  $e(t)$  به صفر میل می کند.

### ۲-۵ ارزیابی مدل

برای ارزیابی قابلیت مدل به کار گرفته شده از میانگین مربعات خطا (MSE) استفاده می کنیم. فرض کنیم  $y_i$  ( $i = 1, 2, \dots, n$ ) مقادیر واقعی و  $\hat{y}_i$  ( $i = 1, 2, \dots, n$ ) پیش بینی های مدل باشند. در این صورت، داریم:

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (۱۱)$$

### ۳. اندازه گیری، مشاهده و محاسبه

در این بخش، با به کارگیری شبکه های عصبی، به پیش بینی میزان آلاینده های هوا در استان لرستان می پردازیم. ارتباط شرایط جوی با میزان آلودگی های هوا را بررسی می کنیم و تحلیل هایی را در این خصوص خواهیم داشت.

#### ۳-۱. محاسبه همبستگی داده ها

یکی از سوالات مهم در زمینه آلودگی هوا، میزان همبستگی بین آلاینده های هوا و همچنین، میزان همبستگی بین شرایط جوی و آلاینده های هواست. در واقع، می خواهیم

$$\dot{z}(t) = f(z(t)) \quad (۳)$$

که در آن  $f$  تابعی ناشناخته است. فرض کنیم  $A$  یک ماتریس هورویتز است و

$$f(z(t)) = Az(t) + g(z(t)) \quad (۴)$$

که در آن  $g$  بخش غیرخطی تابع  $f$  می باشد. فرض کنیم مدل پرسپترون چندلایه می تواند با استفاده از پارامترهای ایده آل  $W_*$  و  $V_*$  تابع  $g$  را تقریب بزند. بنابراین خواهیم داشت:

$$\dot{z}(t) = Az(t) + W_*^T \varphi(V_*^T x) \quad (۵)$$

مدل پارامتری مربوط به این رابطه را می توانیم به صورت زیر بسازیم:

$$\dot{\hat{z}}(t) = A\hat{z}(t) + \hat{W}(t)^T \varphi(\hat{V}(t)^T x) \quad (۶)$$

که در آن  $\hat{z}$ ،  $\hat{W}$  و  $\hat{V}$  به ترتیب تقریب هایی برای  $z$ ،  $W_*$  و  $V_*$  هستند. در رابطه فوق، بردار  $x$  ورودیهای شبکه عصبی را نشان می دهد که درایه های آن مقادیر سری زمانی مدنظر در زمانهای گذشته است. به منظور بکارگیری این مدل در پیش بینی سری های زمانی،  $\hat{z}(t)$  را به صورت زیر تقریب می زنیم:

$$\dot{\hat{z}}(t) = \frac{\hat{z}(t+1) - \hat{z}(t)}{dt} \quad (۷)$$

بنابراین، خواهیم داشت:

$$\hat{z}(t+1) = \hat{z}(t) + dt \left( A\hat{z}(t) + \hat{W}(t)^T \varphi(\hat{V}(t)^T x(t)) \right) \quad (۸)$$

با استفاده از یک الگوریتم یادگیری مناسب می توانیم پارامترهای شبکه عصبی را آموزش دهیم و مقادیر تقریباً بهینه برای آنها بیابیم و سپس با استفاده از شبکه عصبی آموزش دیده به پیش بینی سری زمانی پردازیم.

### ۲-۴-۳. الگوریتم یادگیری پایدار بهنگام

پس از تعیین ساختار شبکه عصبی، استفاده از یک الگوریتم مناسب برای تنظیم پارامترهای آن اهمیت فراوانی دارد. هدف ما از بکارگیری الگوریتم یادگیری تعیین مقادیر ایده آل برای پارامترهای شبکه است به نحوی که مدل طراحی شده

که در آن  $x$  و  $y$  دو متغیر هستند،  $Cov(x, y)$  کوواریانس دو متغیر است و  $\sigma_x$  و  $\sigma_y$  انحراف معیار متغیرها هستند. در این کار، ما از ضریب همبستگی پیرسون برای محاسبه‌ی میزان همبستگی بین متغیرهای هواشناسی و آلودگی هوا بهره می‌گیریم.

### ۳-۲. میزان همبستگی آلاینده‌ها با یکدیگر

میزان همبستگی آلاینده‌های هوا با یکدیگر در جدول ۲ آمده است. مطابق این جدول، همبستگی مثبت زیادی بین  $PM_{10}$  و  $PM_{2.5}$  وجود دارد (0.92). همچنین، همبستگی مثبت معناداری بین  $NO$  و  $NO_2$  وجود دارد (0.98). بین آلاینده‌ی  $O_3$  و سایر آلاینده‌ها یک همبستگی منفی موجود است. لازم به ذکر است که با توجه به میزان بالای داده‌های از دست رفته و نامناسب، پیدا کردن بازه‌ی بزرگی از زمان که در آن همه‌ی آلاینده‌ها داده‌های درست داشته باشند، مشکل است. لذا، مقادیر جدول ۲ با استفاده از داده‌های مربوط به بازه‌ی نسبتاً کوچکی از زمان تهیه شده‌اند.

بدانیم آیا کنترل یکی از متغیرهای جوی یا آلاینده‌ها می‌تواند روی میزان سایر آلاینده‌ها تأثیر بگذارد؟ یا این که آیا یک متغیر جوی می‌تواند به عنوان یک کنترل کننده برای همه یا برخی از آلاینده‌ها عمل کند؟ در این زمینه، مطالعاتی صورت گرفته ولی، همچنان، سوالات مهمی در این زمینه به خوبی پاسخ داده نشده است. اخیراً، لیو و همکارانش (۲۰۲۰) به مطالعه‌ی ارتباط بین شرایط جوی و میزان آلاینده‌ها در کشور چین پرداخته‌اند. در مطالعه‌ی دیگری، همبستگی غلظت آلاینده‌های جوی با کمیت‌های هواشناسی در بندرعباس بررسی شده است (سی‌سی‌پور و همکاران، ۱۴۰۳).

معمولاً از ضریب همبستگی پیرسون برای تعیین میزان همبستگی خطی بین متغیرهای آماری استفاده می‌شود (ماتو و همکارانش، ۲۰۲۱). فرمول ضریب همبستگی پیرسون به صورت زیر است:

$$\rho_{xy} = \frac{Cov(x, y)}{\sigma_x \cdot \sigma_y} \quad (۱۲)$$

جدول ۲. میزان همبستگی آلاینده‌ها به یکدیگر

متغیر	$PM_{2.5}$	$PM_{10}$	$SO_2$	$NO_2$	$NO$	$O_3$	$CO$
$PM_{2.5}$	1	-	-	-	-	-	-
$PM_{10}$	-	1	-	-	-	-	-
$SO_2$	-	-	1	-	-	-	-
$NO_2$	-	-	-	1	-	-	-
$NO$	-	-	-	-	1	-	-
$O_3$	-	-	-	-	-	1	-
$CO$	-	-	-	-	-	-	1

دلایل آلودگی هوا در استان لرستان محسوب می‌شود. معمولاً در میان متغیرهای هواشناسی، دما و سرعت باد به طور کلی به عنوان دو عامل اصلی موثر در نظر گرفته می‌شوند. جابجایی منطقه‌ای آلاینده‌ها با باد راه مهمی برای ایجاد آلودگی است. بدون شک شرایط هواشناسی تأثیر مهمی در شکل‌گیری آلودگی هوا و تغییر غلظت آلاینده‌ها دارد. بررسی سیستماتیک رابطه بین آلاینده‌های هوا و شرایط هواشناسی هنوز به مقدار کافی صورت نگرفته است.

### ۳-۳. میزان همبستگی داده‌های جوی و

#### آلاینده‌ها

در سالهای اخیر، فعالیت‌های شدید انسانی و شرایط نامطلوب هواشناسی مانند پدیده ریزگردها کیفیت هوا را بدتر کرده و به سلامت انسان آسیب می‌رساند و منجر به کمتر شدن شادی شهرنشینان، کاهش پتانسیل انرژی خورشیدی و خسارات اقتصادی هنگفت شود (لیو و همکاران، ۲۰۲۱). علاوه بر تأثیر شدید فعالیت‌های انسانی، شرایط آب و هوایی نیز یکی از

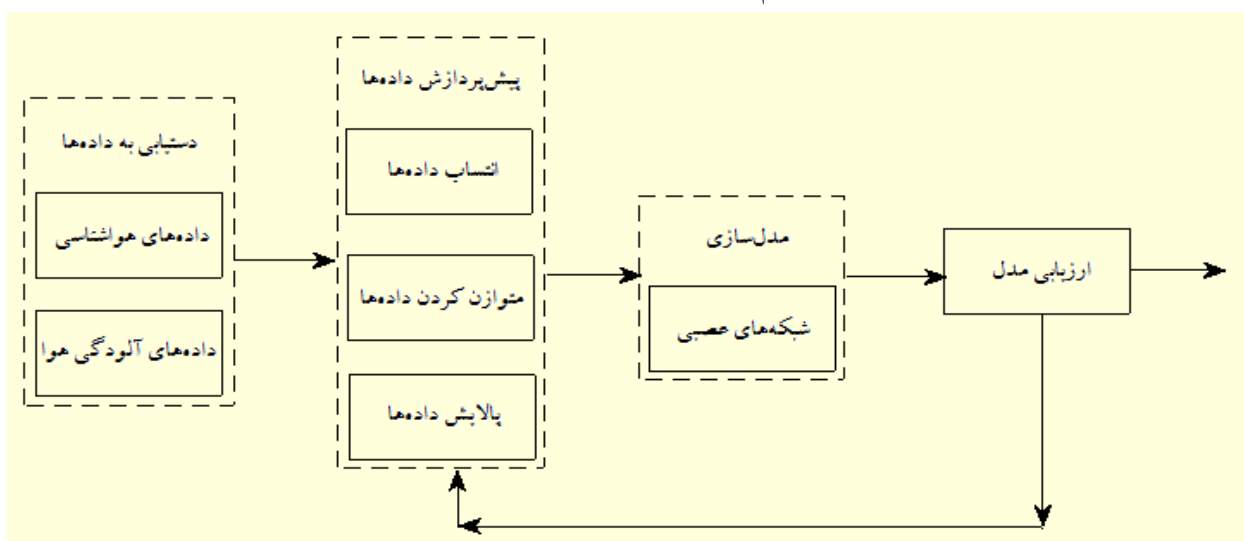
جدول ۳. میزان همبستگی آلاینده‌ها به متغیرهای هواشناسی

متغیر	WS	T	RH	RA	QFE
$PM_{2.5}$	0.12	0.00	0.09	-0.03	0.23
$PM_{10}$	0.02	0.02	-0.02	-0.05	0.18
$SO_2$	0.08	0.64	-0.59	-0.18	-0.66
$NO_2$	-0.14	-0.57	0.64	0.21	0.69
NO	-0.13	-0.58	0.63	0.20	0.68
$O_3$	0.15	0.55	-0.65	-0.21	-0.69
CO	-0.14	-0.56	0.63	0.24	0.69

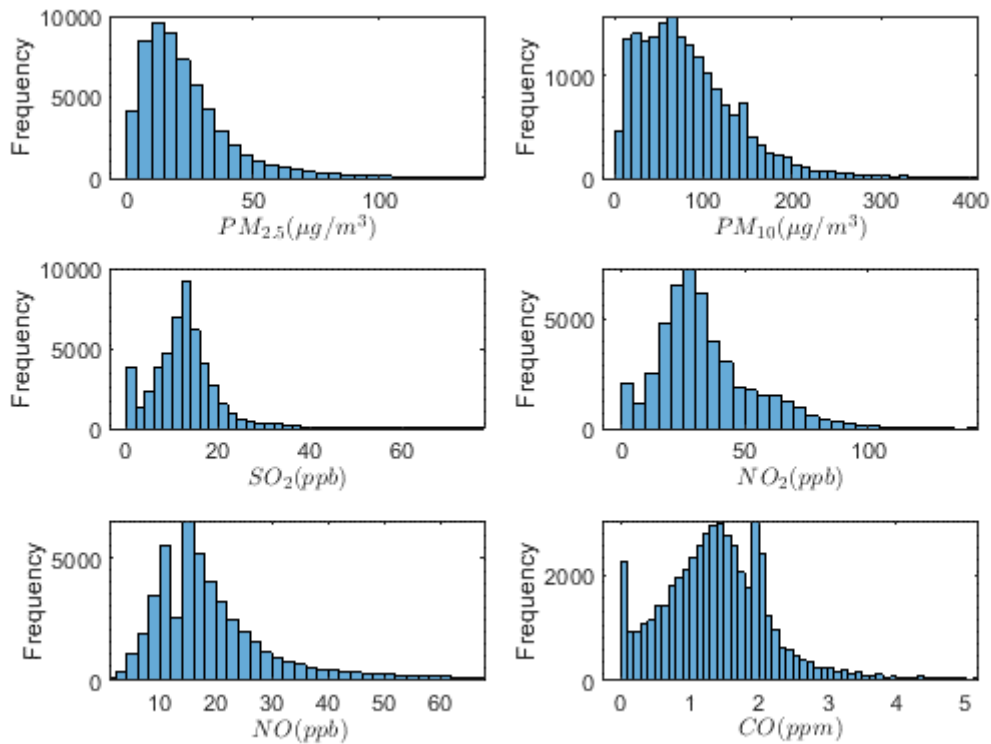
### ۳-۴. پیش‌بینی میزان آلاینده‌های هوا

در این بخش، به پیش‌بینی میزان آلاینده‌های هوا در استان لرستان با بهره‌گیری از شبکه‌های عصبی می‌پردازیم. در اینجا، ابرپارامترهای مدل عصبی و الگوریتم یادگیری مربوط به آن را به صورت تجربی و با سعی و خطا تعیین کرده‌ایم به طوری خطای مدل تا حد ممکن کم باشد. نمودار بلوکی مربوط به فرآیند پیش‌بینی آلودگی هوا در شکل ۶ آمده است. این نمودار بر مبنای روش کشف اطلاعات از داده‌ها که در بخش قبل توضیح دادیم، طراحی شده است. شکل‌های ۷ و ۸ به ترتیب نمودارهای بافت‌نگار مربوط به آلاینده‌ها در شهرهای خرم‌آباد و پلدختر را نشان می‌دهند. چنانچه در این شکل‌ها نیز دیده می‌شود، میزان آلودگیها در خرم‌آباد بسیار بیشتر از پلدختر است. این موضوع را می‌توانیم به حمل و نقل و مصرف سوخت بیشتر در خرم‌آباد مرتبط بدانیم.

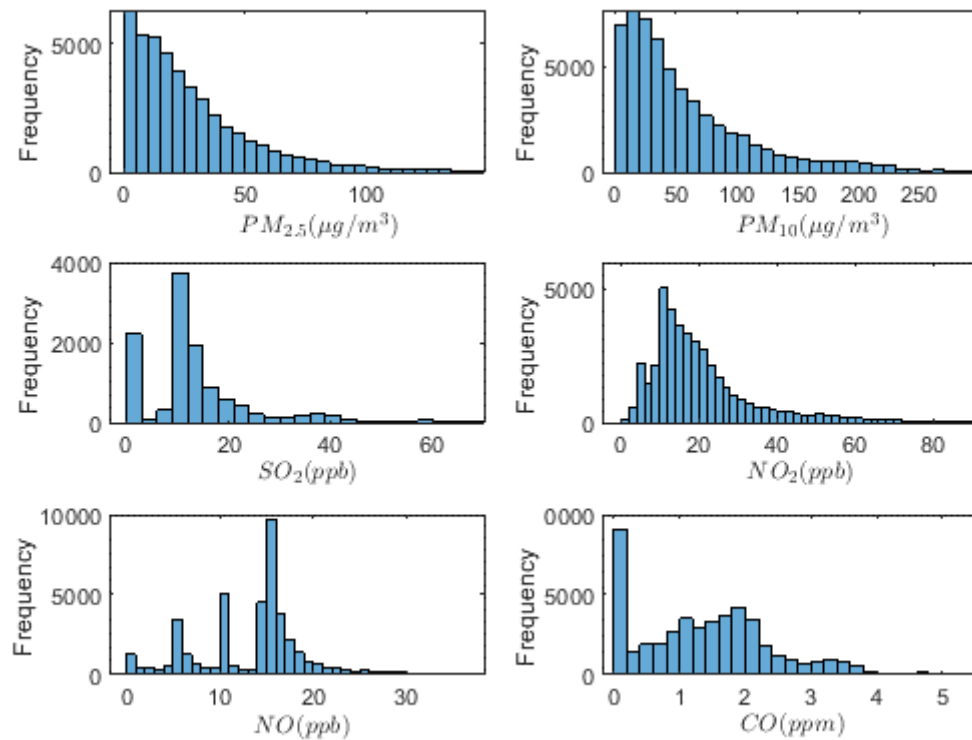
در این مطالعه، داده‌های هواشناسی مربوط به متغیرهای دما ( $T$ )، سرعت باد ( $WS$ )، رطوبت نسبی ( $RH$ )، بارش ( $RA$ ) و فشار جوی ( $QFE$ )، از ابتدای سال ۱۳۹۵ تا آخر سال ۱۴۰۱ مورد استفاده قرار گرفت. جدول ۳ میزان همبستگی میان متغیرهای هواشناسی و آلاینده‌ها را نشان می‌دهد. چنانچه در این جدول دیده می‌شود، همبستگی معنی‌داری بین میزان آلاینده‌ها با متغیرهای هواشناسی وجود دارد. به عنوان نمونه، همزمان با افزایش دما، میزان آلاینده‌ی ازن نیز افزایش می‌یابد، ولی در زمان افزایش رطوبت، میزان آلاینده‌ی ازن کاهش می‌یابد. در خصوص  $NO_2$  و  $CO$  وضعیت کاملاً برعکس است و با افزایش دما، میزان آلاینده‌ی آنها کاهش می‌یابد، ولی در زمان افزایش رطوبت نسبی، میزان آنها افزایش می‌یابد. در خصوص آلاینده‌های  $NO$  و  $NO_2$ ، همبستگی معنی‌دار مثبت آنها را با رطوبت نسبی هوا و همبستگی معنی‌دار منفی آنها را با دما مشاهده می‌کنیم.



شکل ۶: فرآیند پیش‌بینی آلودگی هوا



شکل ۷: نمودار بافت‌نگار میزان غلظت آلاینده‌های هوا در خرم‌آباد مربوط به دوره زمانی از ۱ فروردین ۱۳۹۵ تا ۲۰ آذر ۱۴۰۲



شکل ۸: نمودار بافت‌نگار میزان غلظت آلاینده‌های هوا در پلدختر مربوط به دوره زمانی از ۱ فروردین ۱۳۹۵ تا ۲۰ آذر ۱۴۰۲

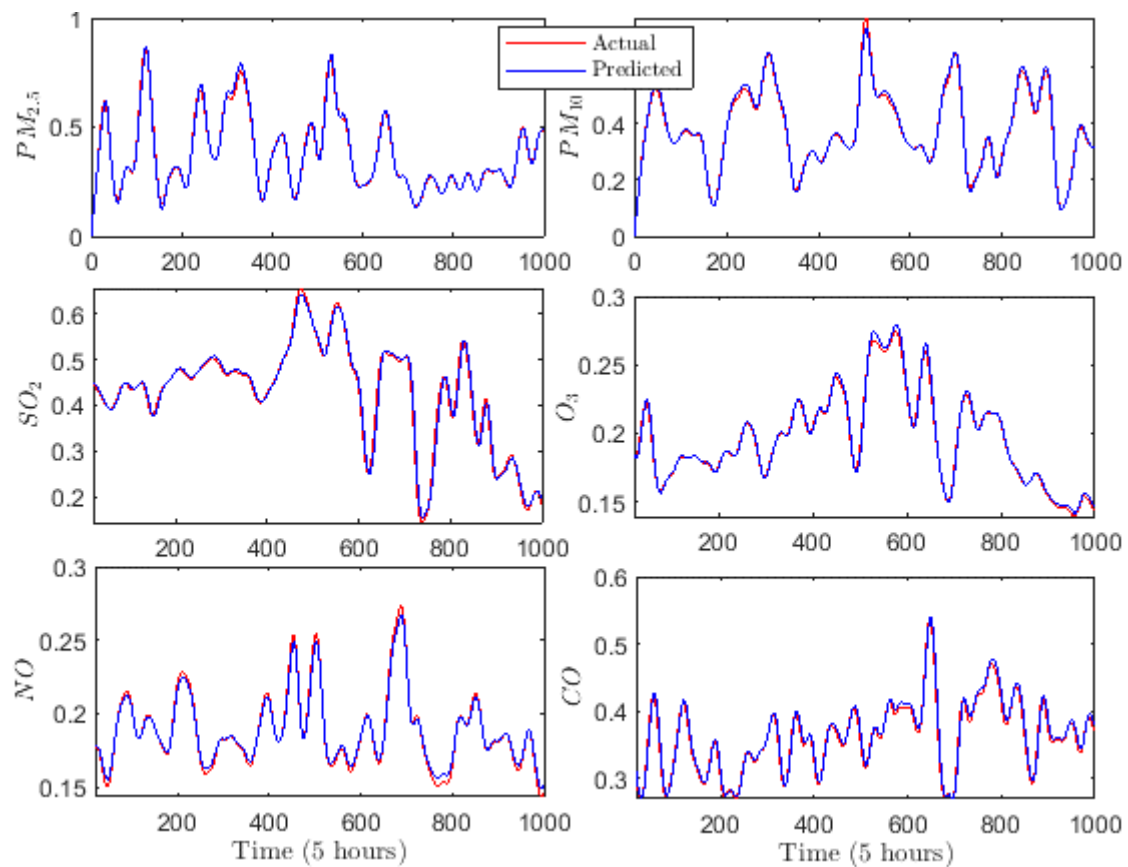
مقادیر آغازین برای پارامترهای شبکه برای هر کدام از آلاینده‌ها در جدول ۴ آمده است. بردار ورودی شبکه عصبی به صورت زیر انتخاب شده است:

$$x = [Z_k, Z_{k-3}, Z_{k-6}, Z_{k-9}, Z_{k-12}, Z_{k-15}, 1]^T \quad (14)$$

در نورون‌های پنهان شبکه عصبی چندلایه از تابع فعال‌ساز تانژانت هیپربولیک استفاده کرده‌ایم. نتایج عددی مربوط به پیش‌بینی‌های صورت گرفته با استفاده از شبکه عصبی چندلایه در جدول ۴ آمده است.

در شکل ۹، داده‌های پنج ساعته (نرمال شده) مربوط به آلاینده‌های هوا و پیش‌بینی یک گام جلوتر مربوط به آنها (۵ ساعت آینده) با استفاده از شبکه عصبی چندلایه برای ۱۰۰۰ داده آمده است. در این جا، برای آموزش شبکه عصبی چندلایه، از ۲۵۰۰ تا ۹۰۰۰ داده استفاده شده و ابرپارامترهای مدل به صورت زیر انتخاب شده است:

$$n = 85, \Gamma_1 = \Gamma_2 = 0.01, A = [-50] \quad (13)$$



شکل ۹: مقادیر نرمال شده مربوط به آلاینده‌های هوا و پیش‌بینی یک گام جلوتر مربوط به آنها (پنج ساعت آینده) در آزمایش مدل پرسپترون چندلایه.

جدول ۴. خطاهای آموزش و آزمایش شبکه عصبی چندلایه در پیش‌بینی میزان آلاینده‌های هوا

متغیر	$dt$	بازه‌ی آغازین	داده‌های آموزش	خطای آموزش (نرمال شده)	خطای آزمایش (نرمال شده)
$PM_{2.5}$	0.01	[-3, 3]	9000	3.84e-4	2.35e-4
$PM_{10}$	0.03	[-1.5, 1.5]	7500	6.98e-4	6.62e-5
$SO_2$	0.01	[-2, 2]	2500	0.0013	2.86e-4
$NO_2$	0.03	[-2, 2]	2000	0.0034	1.22e-4
$NO$	0.03	[-2, 2]	9000	0.0026	2.47e-5
$O_3$	0.03	[-2, 2]	2000	4.35e-4	2.10e-5
$CO$	0.02	[-2, 2]	9000	6.44e-4	2.23e-4

#### ۴. بحث

در این کار، داده‌های آلودگی هوا، جمع‌آوری شده توسط ایستگاه‌های پایش آلودگی هوای سازمان محیط زیست استان لرستان مستقر در دانشگاه علوم پزشکی خرم‌آباد و شهر پلدختر، از ابتدای سال ۱۳۹۵ تا آذر ماه ۱۴۰۲ را در راستای تحلیل وضعیت آلودگی هوای این شهرها مورد استفاده قرار دادیم و با بهره‌گیری از داده‌های ایستگاه‌های هواشناسی سازمان هواشناسی استان لرستان در خرم‌آباد و پلدختر، به بررسی ارتباط شرایط جوی با میزان آلاینده‌های هوا پرداختیم. علاوه بر این، با به‌کارگیری شبکه عصبی چندلایه با یادگیری بهنگام پایدار، به پیش‌بینی کوتاه مدت غلظت آلاینده‌ها در ساعات آینده پرداختیم.

جدول ۲ میزان همبستگی غلظت آلاینده‌ها با یکدیگر را نشان می‌دهد. برای دستیابی به این نتایج، متوسط روزانه‌ی غلظت آلاینده‌ها مورد استفاده قرار گرفته است. مطابق این جدول، همبستگی‌های معنی‌داری بین انواع آلودگی‌های هوا در لرستان وجود دارد. همبستگی مثبت بالایی بین غلظت آلاینده‌های  $PM_{10}$  و  $PM_{2.5}$  وجود دارد که می‌تواند نشانه‌ای از این باشد که در غالب اوقات ذرات معلق ریز و درشت از منابع یکسانی منتشر می‌شوند. همبستگی مثبت معنی‌داری بین غلظت آلاینده‌های  $SO_2$ ،  $NO$  و  $NO_2$  وجود دارد. غلظت آلاینده‌ی  $O_3$  با سایر آلاینده‌ها همبستگی منفی دارد. لازم به ذکر است که با توجه به میزان بالای داده‌های از دست رفته و نامناسب، پیدا کردن بازه‌ی بزرگی از زمان که در آن همه‌ی آلاینده‌ها داده‌های درست داشته باشند، مشکل بود. لذا، مقادیر جدول ۲ با استفاده از داده‌های مربوط به بازه‌ی نسبتاً کوچکی از زمان تهیه شده‌اند.

جدول ۳ میزان همبستگی غلظت آلاینده‌ها با متغیرهای هواشناسی را نشان می‌دهد. برای دستیابی به این نتایج، متوسط روزانه‌ی غلظت آلاینده‌ها در ایستگاه علوم پزشکی خرم‌آباد و کمیت‌های هواشناسی روزانه مربوط به متوسط ایستگاه‌های هواشناسی مستقر در این شهر مورد استفاده قرار گرفته است. مطابق نتایج مطرح شده در جدول ۳، میان برخی از متغیرهای هواشناسی و غلظت آلاینده‌ها همبستگی نسبتاً معنی‌داری از نوع مثبت و منفی مشاهده می‌شود. آلاینده‌ی  $SO_2$  با دمای هوا همبستگی مثبت معنی‌دار دارد و با رطوبت نسبی و فشار هوا همبستگی منفی معنی‌دار دارد. غلظت آلاینده‌های  $PM_{10}$  و  $PM_{2.5}$  همبستگی مثبت اندکی با فشار جوی دارند، ولی همبستگی معنی‌داری بین غلظت این آلاینده‌ها و سایر متغیرهای هواشناسی وجود ندارد. غلظت آلاینده‌های  $CO$ ،  $NO$  و  $NO_2$  با دمای هوا همبستگی منفی معنی‌دار و با رطوبت نسبی و فشار جوی همبستگی مثبت معنی‌دار دارند. بنابراین، این آلاینده‌ها در روزهای سرد غلظت بیشتری دارند. همچنین، غلظت آلاینده‌ی  $O_3$  با دمای هوا همبستگی معنی‌دار مثبت و با رطوبت نسبی و فشار جوی همبستگی منفی معنی‌دار دارد. لذا، با توجه به این نتایج، بحث تابستانه بودن  $O_3$  تأیید می‌گردد. قابل توجه است که سرعت باد بر خلاف انتظار ما همبستگی معناداری با هیچ‌کدام از آلاینده‌ها ندارد. این موضوع می‌تواند به بحث محصور بودن شهر خرم‌آباد توسط کوه‌ها و گردش نامناسب هوا در این شهر مرتبط باشد.

با توجه به این مباحث، می‌توان نتیجه گرفت که شرایط جوی به شدت در میزان آلودگی‌ها مؤثر هستند. گاهی اوقات، افزایش شدید دما منجر به تبدیل شدن برخی از آلاینده‌ها به

پیش‌بینی سریهای زمانی مربوط به آلودگی‌های هوا را نشان می‌دهد. با توجه به منحنی‌های واقعی و تقریبی در شکل ۹، می‌توان نتیجه گرفت که مشکل اصلی مدل در پیش‌بینی‌ها مربوط به نقاط اوج و حوضیض در نمودارهاست. این نقاط اوج و حوضیض مربوط به زمانهایی است که تغییرات ناگهانی در غلظت آلودگیها رخ می‌دهد. این تغییرات می‌تواند در اثر تغییرات ناگهانی در وضعیت جوی مانند بارندگی، وزش باد و ... باشد.

### ۵. نتیجه‌گیری

آلودگی هوا موضوع پر اهمیتی در زمینه‌های سلامت، مدیریت جوامع شهری و مسائل زیست محیطی است. در این پژوهش، این موضوع را بر مبنای داده‌های جمع‌آوری شده در ایستگاه‌های پایش کیفیت هوای سازمان محیط زیست استان لرستان در شهرهای خرم‌آباد و پلدختر بررسی کردیم. با بهره‌گیری از داده‌های جمع‌آوری شده در ایستگاه‌های سازمان هواشناسی این شهرها، میزان همبستگی بین غلظت آلاینده‌های هوا و متغیرهای هواشناسی را به دست آوردیم و مورد بحث قرار دادیم. همچنین، به پیش‌بینی میزان غلظت آلاینده‌ها در پنج ساعت آینده با بهره‌گیری از شبکه عصبی چندلایه با یادگیری بهنگام پایدار پرداختیم. نتایج این پیش‌بینی‌ها نشان داد که شبکه عصبی چندلایه با یادگیری بهنگام پایدار قابلیت زیادی در پیش‌بینی غلظت آلاینده‌های هوا دارد. با توجه به مباحث صورت گرفته، برخی از آلاینده‌ها شدیداً تحت تأثیر متغیرهای هواشناسی مانند دمای هوا، رطوبت نسبی و فشار جوی هستند. بر همین اساس، در کارهای بعدی، قصد داریم از مدل‌های عصبی دیگر مانند LSTM و یادگیری عمیق در پیش‌بینی میزان آلاینده‌ها با لحاظ کردن تأثیر متغیرهای هواشناسی، بهره بگیریم. همچنین، با توجه به بازه‌ای بودن غالب داده‌های آلودگی هوا و هواشناسی، موضوع بکارگیری شبکه‌های عصبی نامتوازن (راف) در این زمینه مفید خواهد بود.

انواع خطرناکتری از آلاینده‌ها می‌شود و گاهی افزایش سرعت باد منجر به پراکنده شدن و یا دور شدن آلاینده‌ها از جو شهر می‌شود. بنابراین، می‌توانیم تأیید کنیم که برخی از متغیرهای هواشناسی در نقش کنترل‌کننده برای برخی از آلاینده‌ها ظاهر می‌شوند. بهر حال، به طور کلی نمی‌توان گفت رشد یک متغیر هواشناسی مثل رطوبت نسبی لزوماً منجر به کاهش آلودگیهای هوا خواهد شد.

نمودارهای هیستوگرام مربوط به داده‌های آلودگی هوا در شهرهای خرم‌آباد و پلدختر به ترتیب در شکل‌های ۷ و ۸ نشان داده شده است. میانگین غلظت آلاینده‌های  $PM_{10}$ ،  $SO_2$ ،  $CO$ ،  $O_3$  و  $NO$  در ایستگاه دانشگاه علوم پزشکی خرم‌آباد به ترتیب برابر است با  $۸۵/۱۰$ ،  $۱۴/۳۵$ ،  $۱/۵$ ،  $۵۱/۳۷$ ،  $۲۲/۰۱$  و  $۳۴/۹۶$  و میانگین غلظت این آلاینده‌ها در پلدختر به ترتیب برابر است با  $۶۴/۸۵$ ،  $۱۸/۲۷$ ،  $۱/۴۷$ ،  $۳۴/۵۸$ ،  $۱۴/۳۴$  و  $۲۱/۴۵$ . با توجه به شکل‌های ۷ و ۸ و میانگین‌های محاسبه شده برای داده‌های شهرهای خرم‌آباد و پلدختر، میزان آلودگی هوای شهر خرم‌آباد در همه‌ی موارد به غیر از  $SO_2$ ، بیشتر از پلدختر است. دلیل بالاتر بودن غلظت آلاینده‌ها در خرم‌آباد می‌تواند با محصور بودن شهر خرم‌آباد با کوه‌ها و نیز جمعیت و گستردگی بیشتر آن نسبت به پلدختر مرتبط باشد. در خصوص بیشتر بودن غلظت آلاینده  $SO_2$  در پلدختر، می‌توان به بحث وجود چاه‌ها و پالایشگاه‌های نفت در اطراف شهر پلدختر اشاره کرد که در انتشار این آلاینده نقش چشم‌گیری دارند (چاوشی و همکاران، ۱۳۹۰).

نتایج بیان شده در شکل ۹ و جدول ۴ نشان می‌دهد که شبکه عصبی چندلایه قابلیت خوبی در پیش‌بینی غلظت آلاینده‌ها دارد. لازم به ذکر است که این نتایج برای داده‌های نرمال شده به دست آمده‌اند. خطای آموزش در جدول ۴ مربوط به یک دوره‌ی آموزش با الگوریتم یادگیری بهنگام پایدار مطرح شده در روابط (۹) و (۱۰) است. میزان پایین خطاهای نرمال شده در آموزش نشان‌دهنده‌ی قدرت الگوریتم یادگیری به کار گرفته شده است. خطای اندک مربوط به آزمایش مدل پیشنهاد شده در پیش‌بینی انواع آلاینده‌ها در شکل ۹ و جدول ۴، توانمندی شبکه عصبی چندلایه را در

## تشکر و قدردانی

نویسندگان این مقاله مراتب قدردانی خود را از همکاری صمیمانه‌ی مدیران و کارکنان سازمانهای هواشناسی و محیط زیست استان لرستان در خصوص داده‌های مورد استفاده در این مقاله، ابراز می‌کنند.

## منابع

۱. احمدی، غ.ر.، کردوانی، پ.، محمدی، ح. (۱۴۰۰). پهنه‌بندی اقلیم کشاورزی محصول ذرت در استان لرستان. *جغرافیای طبیعی*، ۱۳: ۲۱-۳۸.
۲. بشردوست، ع.، مسگری، م. س. (۱۴۰۳). مدل سازی مکانی ذرات معلق هوا ( $PM_{2.5}$  و  $PM_{10}$ ) در شهر تهران با استفاده از شبکه عصبی کانولوشن، *نیوار*، ۴۸(۱۲۴): ۳۱-۴۹.
۳. چاوشی، ب.، مسعودی‌نژاد، م.، ادیب‌زاده، ا. (۱۳۹۰). ارزیابی میزان انتشار و ضریب انتشار (فاکتور انتشار) گاز دی‌کسیدگوگرد از خروجی‌های پالایشگاه نفت تهران. *سلامت و محیط زیست*، ۴(۲): ۲۴۴-۲۳۳.
۴. حسونند، م.ص.، ندافی، ک. (۱۳۹۵). آلاینده‌های هوا، در: *آلودگی هوا و سلامت: مروری بر مطالعات و تحقیقات جهانی در باره‌ی اثرات آلودگی هوا بر سلامت انسان*، مؤسسه نشر شهر، تهران.
۵. زنگوئی، ح.، اسداله‌فردی، غ.، (۱۳۹۶). پیش‌بینی آلودگی  $PM_{10}$  هوای شهر مشهد با استفاده از شبکه‌های عصبی مصنوعی MLP و مدل زنجیره مارکف، *تحقیقات کاربردی علوم جغرافیایی (علوم جغرافیایی)*، ۱۷(۴۷): ۳۹-۵۹.
۶. سجادیان، ن. (۱۳۹۴). پیش‌بینی آلودگی هوای ناشی از حمل و نقل شهری کلانشهر تهران با بهره‌گیری از تلفیق GIS با مدل LUR و شبکه عصبی مصنوعی. *اطلاعات جغرافیایی*، ۲۴(۹۵): ۱۰۷-۱۲۰.
۷. سی‌سی‌پور، م.، رهنما، م.، صحت‌کاشانی، س.، خدام، ن.، نوری، ف.، قهرمانی، س. (۱۴۰۳). مطالعه همبستگی
۸. صادقی، ح.، خاکسار آستانه، س. (۱۳۹۳). پیش‌بینی کوتاه‌مدت آلودگی ذرات معلق شهر اهواز با کمک شبکه‌های عصبی، *پژوهش های محیط زیست*، ۵(۹): ۱۷۷-۱۸۶.
۹. غیاث‌الدین، م. (۱۳۹۶). *آلودگی هوا، منابع، اثرات و کنترل*، انتشارات دانشگاه تهران، تهران.
۱۰. قائمی، ز.، طالعی، م.، فرنقی، م.، جوادی، ق. (۱۳۹۶). پیش‌بینی آلودگی هوای شهر تهران با استفاده از ترکیب شبکه عصبی فازی- تطبیقی و آنالیز مؤلفه اصلی، *نشریه سنجش از دور و GIS/ایران*، ۳(۹): ۴۵-۷۰.
۱۱. کفاش چرندابی، ن.، غلامی، ا. (۱۴۰۱). توسعه مدل ترکیبی بر پایه شبکه عصبی جهت مدل‌سازی کیفیت هوای داخلی بازار تبریز به لحاظ ذرات معلق، *محیط شناسی*، ۴۵(۳): ۴۹۹-۵۱۲.
12. Chandra, R., Goyal, S., and Gupta, R., 2021. Evaluation of deep learning models for multi-step ahead time series prediction. *IEEE Access*, 9:83105–83123, 2021.
13. Cordova C.H., Portocarrero M.N.L., Salas R., et al., 2021. Air quality assessment and pollution forecasting using artificial neural networks in Metropolitan Lima-Peru. *Sci Rep*. 11:24232.
14. Chen J. et al., 2014. Seasonal modeling of  $PM_{2.5}$  in California's San Joaquin valley. *Atmos. Environ.* 92, 182–190.
15. Saide P. et al., 2011. Forecasting urban  $PM_{10}$  and  $PM_{2.5}$  pollution episodes in very stable nocturnal conditions and complex terrain using WRF-Chem CO tracer model. *Atmos. Environ.* 45, 2769–2780.
16. Ahmadi G., 2019. Stable rough extreme learning machines for the identification of uncertain continuous-time nonlinear systems. *Control and Optimization in Applied Mathematics*, 4:83–101.
17. Ahmadi G. and Dehghandar M., 2023. Chaotic Time Series Prediction Using Rough-Neural Networks, *Mathematics Interdisciplinary Research*, 8 (2) 71- 92.
18. Ahmadi G. and Teshnehlab M., 2017. Designing and implementation of stable sinusoidal rough-neural identifier. *IEEE*

- optimization and evaluation. *Sustainable Cities and Society*, 65:102567.
26. Salas R. and Bustos A., 2005. Constructing a narx model for the prediction of the  $PM_{10}$  air pollutant concentration. In *Encuentro Chileno de Computación, Jornada Chilena de Ciencias de la Computación*, Valdivia, Chile. Nov. 7-12.
  27. Kantz H. and Schreiber T., 2004. *Nonlinear Time Series Analysis*, Cambridge University Press.
  28. Delavar M.R., Gholami A., Shiran G.R., Rashidi Y., Nakhaeizadeh G.R., Fedra K., and Hatefi Afshar S., 2019. A Novel Method for Improving Air Pollution Prediction Based on Machine Learning Approaches: A Case Study Applied to the Capital City of Tehran. *ISPRS Int. J. Geo-Inf.* 8, 99.
  29. Tealab A., 2018. Time series forecasting using artificial neural networks methodologies: A systematic review. *Future Computing and Informatics Journal*, 3(2):334-340.
  30. Tiwary, A., Williams, I., and J., Colls, 2018. *Air Pollution: Measurement, Modelling and Mitigation*, 4th ed., CRC Press.
  31. Seinfeld, J.H., and Pandis, S.N., 1997, *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*, Wiley-VCH: New York.
  19. Montgomery D.C., Jennings C.L., and Kulahci M., 2008. *Introduction to time series analysis and forecasting*. John Wiley & Sons, Hoboken, New Jersey, USA.
  20. Faruk, D.O., 2010, A hybrid neural network and arima model for water quality time series prediction. *Engineering Applications of Artificial Intelligence*, 23:586-594.
  21. Frank, R.J., Davey, N., and Hunt, S.P., 2001. Time series prediction and neural networks. *J. Intell. Robot. Syst.*, 31:91-103.
  22. Gholipour, A., Araabi, B.N., and Lucas, C., 2006. Predicting chaotic time series using neural and neurofuzzy models: A comparative study. *Neural Process Lett.*, 24:217-239.
  23. Gupta, M.M., Jin, L., and Homma, N., 2003. *Static and Dynamic Neural Networks: From Fundamentals to Advanced Theory*. John Wiley & Sons, New Jersey, USA.
  24. Maleki H., Sorooshian A., Goudarzi G., Baboli Z., Birgani Y.T., and Rahmati, M., 2019. Air pollution prediction by using an artificial neural network model. *Clean Technol. Environ. Policy* 21, 1341-1352.
  25. Mao W., Wang W., Jiao L., Zhao S., and Liu A., 2021. Modeling air quality prediction using a deep learning approach: Method