



مدل داده کاوی مبتنی بر یادگیری ماشین جهت پیش بینی آلودگی هوا در کلان شهرهای ایران

دانشجوی دکترا، رشته مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه آزاد اسلامی واحد قزوین، قزوین، ایران
استادیار، دانشکده مدیریت و حسابداری، دانشگاه آزاد اسلامی واحد قزوین، قزوین، ایران
استادیار، دانشکده مدیریت و حسابداری، دانشگاه آزاد اسلامی واحد قزوین، قزوین، ایران

عباس ملکی
صادق عابدی*
علیرضا ایرج پور

چکیده مبسوط

مقدمه: در پاسخ به همه گیری کووید-۱۹، دولت ها در سراسر دنیا محدودیت های شدیدی را اعمال نموده و سناریوهای متفاوتی از کاهش انتشار آلاینده های ناشی از منابع ترافیکی را ارائه کردند. با اعمال محدودیت های ترافیکی ناشی از همه گیری کووید-۱۹ و قرنطینه، به نظر می رسد کاهش در تردد وسایل نقلیه و غلظت آلاینده ها صورت پذیرفته و به نوعی شاخص کیفیت هوا به استانداردهای کیفیت نزدیک گردد. بنابراین انتظار می رفت تغییراتی در غلظت آلاینده های منواکسید کربن (CO)، ازن (O3)، منواکسید نیتروژن (NO)، دی اکسید نیتروژن (NO2)، اکسیدهای نیتروژن (NOx)، دی اکسید گوگرد (SO2)، ذرات معلق با اندازه ۲.۵ میکرومتر یا کمتر (PM2.5) و ذرات معلق با اندازه ۱۰ میکرومتر یا کمتر (PM10) که از عوامل بیماری زا و بعضاً مرگ زودرس می باشند، مشاهده گردد.

مواد و روش ها: با استفاده از روش یادگیری ماشین، در مرحله اول تغییر غلظت آلاینده ها در دوره همه گیری کووید-۱۹ نسبت به دوره قبل از آن مورد بررسی قرار گرفت تا مشخص گردد که اعمال محدودیت های ترافیکی چه تاثیری در تغییر غلظت آلاینده ها در هر یک از کلان شهرهای مشهد و شیراز داشته است. در مرحله دوم، الگوهای پیش بین با استفاده از شبکه های عصبی پیشخور و عمیق جهت پیش بینی سطح اهمیت بهداشتی بر اساس اعمال هر یک از محدودیت های ترافیکی در هر کلان شهر ارائه گردید.

نتایج و بحث: این مطالعه نشان داد که تغییر غلظت آلاینده ها در دوران کووید-۱۹ نسبت به قبل از آن، در کلان شهرهای مشهد و شیراز متفاوت بوده است. با توجه به نتیجه حاصل شده و تفاوت در روند تغییر غلظت آلاینده ها، جهت کنترل شاخص کیفیت هوا، الگوهای پیش بین برای هر کلان شهر ارائه گردید.

نتیجه گیری: برای آلاینده ها، الگوی افزایشی و یا کاهشی یکسانی در کلان شهرهای مورد مطالعه مشاهده نشد، پس می توان گفت که تاثیر محدودیت های یکسان بر روی تغییر غلظت آلاینده ها در شهرهای مختلف، متفاوت بوده است؛ لذا اعمال محدودیت های یکسان در تمامی شهرها، لزوماً منجر به کاهش آلودگی نشده و می بایست برای هر موقعیت شهری و زیست محیطی، الگوی محدودیت های ترافیکی مختص آن موقعیت تهیه گردد. الگوی حاصل از این تحقیق را می توان به صورت روزانه، هفتگی، ماهانه و سالانه برای هر یک از دو کلان شهر مورد استفاده قرار داد و مشخص کرد که کدام محدودیت ها در کدام شهر سبب خواهد شد تا بتوانیم شاخص کیفیت هوا را کنترل نماییم.

تاریخ دریافت: ۱۴۰۳/۰۲/۰۹

تاریخ پذیرش: ۱۴۰۳/۰۳/۱۹

واژه های کلیدی: پیش بینی
آلودگی هوا، کلان شهرهای ایران،
یادگیری ماشین، شبکه عصبی،
کووید-۱۹.

نویسنده مسئول: صادق عابدی

نشانی: دانشکده مدیریت و حسابداری، دانشگاه آزاد اسلامی واحد قزوین، قزوین، ایران. تلفن: ۰۹۱۲۸۱۷۳۶۶۲ پست الکترونیکی: abedi.sadegh@gmail.com

استناد: ملکی عباس، عابدی صادق، ایرج پور علیرضا. مدل داده کاوی مبتنی بر یادگیری ماشین جهت پیش بینی آلودگی هوا در کلان شهرهای ایران. پژوهش های نوین در مهندسی محیط زیست. ۱۴۰۳؛ ۵(۲): ۹-۱.

حقوق نویسندگان محفوظ است. این مقاله با دسترسی آزاد و تحت مجوز مالکیت خلاقانه <http://creativecommons.org/licenses/by/4.0> در فصلنامه پژوهش های نوین در مهندسی محیط زیست منتشر شده است. هرگونه استفاده غیرتجاری فقط با استناد و ارجاع به اثر اصلی مجاز است.



مقدمه

با اندازه ۲/۵ میکرومتر یا کمتر ($PM_{2.5}$)، ذرات معلق با اندازه ۱۰ میکرومتر یا کمتر (PM_{10})، دی اکسید نیتروژن (NO_2) و غلظت دی اکسید گوگرد (SO_2) و اثرات مثبت را با توجه به کیفیت هوا در برخی از شهرهای هند در طول همه گیری کووید-۱۹ (گاتام ۲۰۲۰؛ چلانی و گاتام ۲۰۲۲؛ بهروانی و همکاران ۲۰۲۱؛ بهروانی و همکاران ۲۰۲۱؛ کاتام و همکاران ۲۰۲۱؛ کلورو و همکاران ۲۰۲۲) و با همبستگی شدید و منفی نشان می‌دهد.

نتیجه پژوهش بیکان ما و همکاران (۲۰۲۴) نشان داد که قرنطینه سبب کاهش آلاینده‌های $PM_{2.5}$ و NO_2 در چین، ایالات متحده، ایتالیا و آلمان شده است. فاضل رستگار و سیواکومار (۲۰۲۳) در پژوهشی تحت عنوان مطالعه موردی تأثیر آب و هوا بر آلودگی هوا در دوره همه گیری COVID-19 در آفریقای جنوبی نشان دادند که غلظت آلاینده‌های NO_2 و SO_2 کاهش یافتند و تغییر قابل توجهی در غلظت آلاینده O_3 مشاهده نشد. نتیجه پژوهش ژیاو فنگ و همکاران (۲۰۲۳) در پژوهشی که با موضوع تغییرات در آلودگی هوا و دمای سطح زمین در طول قرنطینه COVID-19 در سه شهر چین انجام دادند نشان داد که غلظت آلاینده‌های $PM_{2.5}$ ، NO_2 و CO کاهش یافتند. غلظت آلاینده‌های $PM_{2.5}$ ، NO_2 بیش از ۱۴ درصد کاهش یافتند و دمای سطح زمین در این سه شهر بیش از یک درجه سانتیگراد افزایش یافت.

مهران شایگان و مرضیه مکرم (۱۴۰۲)، آلودگی هوا در دوران کرونا و پیش از آن در کلانشهرهای تهران، اصفهان و قم مورد بررسی قرار دادند و نتایج حاصل شده نشان داد که آلودگی هوا در دوران کرونا نسبت به قبل از آن کاهش یافته است، آلودگی هوای در تهران نسبت به سایر شهرهای مورد مطالعه بیشتر است و همچنین مشخص شد که افزایش دما با میزان آلودگی ارتباط معناداری دارد.

جیم گونزالس پارودو و همکاران (۲۰۲۲) در پژوهش خود به تخمین تغییرات سطوح آلودگی هوا ناشی از کووید-۱۹ با استفاده از تکنیک داده‌کاوی پرداختند که در آن پارامترهای منواکسید کربن (NO)، NO_2 ، ازن (O_3)، $PM_{2.5}$ و PM_{10} را مورد بررسی قرار داده بودند. در این تحقیق که سال ۲۰۲۰ به پنج دوره پیش از قرنطینه، قرنطینه، تشدید، عادی و قرنطینه مجدد تقسیم شده

آلودگی هوا یکی از مهمترین عوامل خطر محیطی است که می‌تواند منجر به زیان‌های اقتصادی و سلامت قابل توجهی شود. در حال حاضر ۹۱ درصد از جمعیت جهان به ویژه در آسیا، خاور میانه و آفریقا در مناطقی با غلظت میانگین بالای ۱۰ میکروگرم بر متر مکعب آلاینده ذرات معلق با اندازه ۲.۵ میکرومتر یا کمتر ($PM_{2.5}$) زندگی می‌کنند (مک دافی و همکاران ۲۰۲۱).

گسترش جهانی همه‌گیری Covid-19 در اوایل سال ۲۰۲۰، بزرگترین بحران بهداشتی در سال‌های اخیر بوده است حتی در مقایسه با آنفولانزای اسپانیایی اوایل قرن بیستم (سانچز لورنزو و همکاران ۲۰۲۰؛ مونولی و همکاران ۲۰۲۰). در پاسخ به مرگومیر بالای ناشی از ویروس SARS-COV-2 و گسترش سریع آن، دولت‌های سراسر جهان اقدامات و محدودیت‌های کنترل شدیدی را اعلام کردند. برخی از محدودیت‌های اجرا شده شامل محدودیت‌های رفت‌وآمد و یا مقررات فاصله‌گذاری اجتماعی بود. محدودیت‌های اعمال شده به دلیل پاندمی Covid-19، یک سناریوی را ارائه می‌دهد که در واقع تأثیر این همه‌گیری در پارامترهای آلودگی هوا می‌باشد که ناشی از کاهش ترافیک و محدودیت‌های رفت‌وآمد بوده است. (جیم گونزالس پارودو و همکاران ۲۰۲۲). انتظار می‌رفت قرنطینه اعمال شده در دوره همه‌گیری، تأثیر به‌سزایی در کاهش آلودگی هوا داشته باشد. منابع آلاینده‌های هوا در منطقه شهری عمدتاً از سوخت وسایل نقلیه، فعالیت سوختی زیست محیطی و تعلیق مجدد گرد و غبار در شهرها ایجاد می‌شود؛ پس قرنطینه می‌توانست در کاهش آمد و شد وسایل نقلیه و کاهش غلظت آلاینده‌ها تأثیرگذار بوده و به نوعی شاخص کیفیت هوا را به استانداردهای کیفیت نزدیک نماید. کشور ایران، مدت‌ها است که با مشکل آلودگی هوا مواجه است. این آلودگی که پیش‌تر و تحت تأثیر تمرکز جمعیت، منحصر به پایتخت - تهران - بود، در سال‌های پیشین در سایر شهرهای ایران نیز به شدت افزایش یافته است؛ تا جایی که در آمارهای منتشر شده از آلوده‌ترین شهرهای جهان از سوی سازمان جهانی بهداشت، نام چندین شهر از ایران در رتبه‌های نخست این فهرست به چشم می‌خورد (شاهرخ ملک زاده ۱۳۹۷).

مطالعات مختلفی در مورد تأثیر آلاینده‌های هوا و عوامل هواشناسی در طول دوره همه‌گیری کووید-۱۹ انجام شده است. برخی از مطالعات اخیر کاهش آلاینده‌های هوا مانند ذرات معلق

1. Particulate Matter
2. Nitrogen dioxide
3. Sulfur dioxide
4. Nitrogen monoxide
5. Ozone

در این مطالعه علاوه بر بررسی تغییر غلظت آلاینده‌ها در دوره همه‌گیری و مقایسه آن با دوره قبل از شروع همه‌گیری برای هر یک از کلان‌شهرها، الگوهای پیش‌بین برای هر کدام از آنها نیز به صورت جداگانه ارائه خواهد شد تا در شرایط مشابه و یا شرایطی که مسئولان مربوطه بخواهند غلظت آلاینده‌ها را کنترل کنند، از این الگوها استفاده نموده و بدین ترتیب امکان پیش‌بینی شاخص کیفیت هوا و سطح اهمیت بهداشتی با اعمال محدودیت‌های مختلف فراهم گردد. در این مطالعه از روش داده‌کاوی که یک روش مبتنی بر داده می‌باشد استفاده خواهد شد. داده‌کاوی فرآیند به خدمت گرفتن یک روش‌شناسی رایانه‌ای است که با استفاده از تکنیک‌ها و الگوریتم‌های مختلف در جستجوی دانش نهفته در داده‌هاست (کانتاردزیچ ۲۰۰۳). این فرآیند مشارکتی میان انسان و رایانه (برسون و همکاران ۲۰۰۰) در نهایت به دنبال کشف الگوها و قواعد معنادار در میان داده‌ها بوده (بری و لینف ۲۰۰۰) و پایگاه‌های داده‌ای بزرگ را به عنوان منبع دانش در نظر می‌گیرد (لاروس ۲۰۰۵). هدف از انجام این پژوهش، مقایسه غلظت آلاینده‌ها در دوره کووید-۱۹ نسبت به قبل از آن و همچنین ارائه الگوهای پیش‌بین با استفاده از یادگیری ماشین برای کلان‌شهرهای شیراز و مشهد می‌باشد. با استفاده از الگوهای پیش‌بین، سازمان‌های مربوطه می‌توانند در شرایطی مشابه همه‌گیری کووید-۱۹ و یا در صورت عدم وجود همه‌گیری و صرفاً به منظور کنترل غلظت آلاینده‌ها در مقاطعی از سال از آنها استفاده نمایند.

مواد و روش‌ها

روش کار

در این پژوهش، برای استفاده بهینه از مجموعه داده، در برخی روزها که اطلاعات ویژگی مربوط به آن روز در دسترس نبود، از روزهای قبل استفاده شده است تا فیلدهای مربوط به آن روز کامل شود. تجمیع اطلاعات چندین ایستگاه با روش میانگین انجام شده است. قابل ذکر است که برای استفاده از این مجموعه داده در شبکه‌های عصبی لازم است استاندارد سازی و نرمال سازی روی داده‌ها صورت پذیرد. در زمینه شبکه‌های عصبی، روش مقیاس‌کننده استاندارد (Standard Scaler) بخشی از مراحل پیش پردازش است و برای استانداردسازی یا عادی سازی داده‌های ورودی استفاده می‌شود. استانداردسازی یک روش معمول در یادگیری ماشینی، از جمله شبکه‌های عصبی، برای رساندن همه ویژگی‌ها به مقیاس مشابه است. در روش Standard Scaler

بود، مشاهده شد که پارامتر اکسیدهای نیتروژن (NO_x) بیش از ۵۰٪ کاهش یافت، پارامتر O_3 در تضاد با پارامتر NO_x ، ۲۳/۹٪ افزایش یافت و همچنین پارامترهای PM_{10} و $\text{PM}_{2.5}$ نیز دارای کاهش جزئی به میزان ۴،۱٪ و ۲،۳٪ بودند که نشان‌دهنده تأثیرپذیری کم پارامترهای PM_{10} و $\text{PM}_{2.5}$ از منابع ترافیک و قرنطینه می‌باشد. سینگ ساهاران و همکاران (۲۰۲۲) در پژوهشی، عوامل تغییر آلودگی هوا را در موج دوم کووید-۱۹ در دهلی هند مورد بررسی قرار دادند که در این پژوهش مشخص شد که در سال ۲۰۲۱، غلظت آلاینده‌های PM_{10} ، $\text{PM}_{2.5}$ ، آمونیاک (NH_3)، NO_x ، منواکسید کربن (CO)، بنزن و تولوئن در طول منع رفت‌وآمد شبانه در مقایسه با دوره قبل از قرنطینه به میزان ۴٪ تا ۱۶٪ کاهش یافت. یانگ و همکارانش (۲۰۲۲) در مطالعه‌ای، پیامدهای کووید-۱۹ را بر آلودگی محیط زیست و انتشار کربن بررسی نمودند که طی این مطالعه مشاهده شد که سطوح $\text{PM}_{2.5}$ ، NO_2 ، PM_{10} و CO در بیشتر مناطق جهان کاهش یافت اما سطوح SO_2 و O_3 افزایش یافت یا تغییر قابل‌توجهی نشان نداد. آدای و همکاران (۲۰۲۲) نشان دادند که قرنطینه به دلیل کووید-۱۹ منجر به کاهش سطوح آلاینده‌ها از جمله CO_2 و NO_2 همراه با سطوح $\text{PM}_{2.5}$ شده است. در مطالعه جاسپر ویجناندز و همکاران (۲۰۲۲)، بیشترین کاهش NO_2 در طول قرنطینه اولیه در چین، اروپا و هند مشاهده شد. مطالعه ژوان یو لین و همکاران (۲۰۲۲) نشان داد که کاهش قابل‌توجهی در غلظت NO_2 به میزان ۲۰٪ و ۲۵،۹٪ در طول هشدار سطح ۳ کووید-۱۹ از ژوئن تا ژوئیه ۲۰۲۱ در مقایسه با مدت مشابه در سال‌های ۲۰۱۹ و ۲۰۲۰ با سیاست دورکاری، آموزش از راه دور و کاهش جریان ترافیک رخ داده است. نتایج تحقیق گالتیبری و همکاران (۲۰۲۰) حاکی از کاهش قابل توجه سطوح NO_2 ، افزایش جزئی غلظت O_3 ، کاهش PM_{10} تا ۳۱/۵٪ در پالمو و افزایش این پارامتر در ناپولی تا ۷/۳٪ و همچنین کاهش $\text{PM}_{2.5}$ بین ۱۳٪ تا ۱۷٪ بود. مطالعه دیوید هیدالگو گارسیا و جولیان آرکو دیاز (۲۰۲۱) بیانگر کاهش غلظت SO_2 به میزان ۳۳،۵٪، NO_2 به میزان ۴۴٪، PM_{10} به میزان ۳۸،۳٪، CO به میزان ۲۶،۵٪ و افزایش غلظت O_3 به میزان ۵،۹٪ بود.

در مطالعات فوق، بعضاً نتیجه‌گیری‌ها و بررسی‌ها به صورت کلی می‌باشد و یا اینکه مطالعات صورت گرفته برای یک یا چند شهر به کل کشور تعمیم داده شده است. می‌توان گفت که در بیشتر مطالعات صورت گرفته، به توصیف و بررسی تغییر غلظت آلاینده‌ها در دوره همه‌گیری کووید-۱۹ پرداخته شده است؛ درحالی که

محدودیت‌های اعمال شده در دوره کووید-۱۹ چه تاثیری بر غلظت آلاینده‌ها داشته است. در مرحله بعد، با استفاده از شبکه های عصبی پیشخور (FNN) و عمیق با Dropout (DNN)، الگوهای پیش‌بین مطابق با محدودیت های ترافیکی اعمال شده استخراج شد.

در این مرحله، داده های مربوطه با استفاده از نرم افزار پایتون آموزش داده شده و تست گردید. سازمان حفاظت محیط زیست ایران، برای تعیین شاخص کیفیت هوا (AQI) در هر یک از شهرهای کشور از رابطه ۲ استفاده می کند. در این رابطه بیشترین مقدار ثبت شده برای پارامترهای CO، O₃، NO₂، SO₂، PM_{2.5} و PM₁₀ به عنوان شاخص کیفیت هوا مشخص و مطابق با آن، سطح اهمیت بهداشتی برای آن شهر تعیین می شود.

$$AQI = \text{Max}(CO, O_3, NO_2, SO_2, PM_{2.5}, PM_{10}) \quad (2)$$

در جدول ۱، شاخص کیفیت هوا متناسب با سطح اهمیت بهداشتی، دسته‌بندی شده است.

اغلب از کتابخانه‌هایی مانند Scikit-learn در پایتون استفاده می شود. روش Standard Scaler ویژگی ها را با حذف میانگین و مقیاس بندی به واریانس واحد استاندارد می کند. میانگین و انحراف استاندارد هر ویژگی در داده‌های آموزشی را محاسبه می کند و سپس تبدیل را به مجموعه داده‌های آموزشی و آزمایشی اعمال می کند. فرمول تبدیل به شرح رابطه شماره ۱ است

$$X_{standardized} = X - \text{mean}(X) / \text{std}(X) \quad (1)$$

X: داده اصلی

Mean(x): میانگین x

Std(x): انحراف معیار x

در زمینه شبکه های عصبی، Standard Scaler بر روی ویژگی های ورودی قبل از تغذیه آنها به شبکه اعمال می گردد که معمولاً به عنوان یک مرحله پیش پردازش انجام می شود. در پژوهش حاضر، با استفاده از روش یادگیری ماشین و با نرم‌افزار پایتون (Python)، میانگین غلظت آلاینده‌ها در دوره کووید-۱۹ و قبل از آن با یکدیگر مقایسه گردید تا مشخص گردد که

جدول ۱- شاخص کیفیت هوا براساس استاندارد سازمان حفاظت محیط زیست ایران

شاخص کیفیت هوا	شماره کلاس	سطح اهمیت بهداشتی
۰-۵۰	۱	پاک
۵۱-۱۰۰	۲	سالم
۱۰۱-۱۵۰	۳	ناسالم برای گروه‌های حساس
۱۵۱-۲۰۰	۴	ناسالم
۲۰۱-۳۰۰	۵	بسیار ناسالم
۳۰۱-۵۰۰	۶	خطرناک

شبکه های عصبی مصنوعی (ANNs)ها که به صورت مخفف به آنها شبکه های عصبی (NNs) هم گفته می‌شود) شاخه‌ای از مدل‌های یادگیری ماشینی هستند که با استفاده از اصول سازمان‌دهی عصبی که توسط پیوندگرایی در شبکه‌های عصبی بیولوژیکی تشکیل‌دهنده مغز حیوانات کشف شده است، ساخته شده‌اند (هاردستی ۲۰۱۷). یک ANN بر اساس مجموعه‌ای از واحدها یا گره‌های متصل به نام نورون‌های مصنوعی است که نورون‌ها را در یک مغز بیولوژیکی مدل‌سازی می‌کنند. هر اتصال، مانند سیناپس های یک مغز بیولوژیکی، می تواند سیگنالی را به شبکه های عصبی مصنوعی (ANNs)ها که به صورت مخفف به آنها شبکه های عصبی (NNs) هم گفته می‌شود) شاخه‌ای از مدل‌های یادگیری ماشینی هستند که با استفاده از اصول سازمان‌دهی عصبی که توسط پیوندگرایی در شبکه‌های عصبی بیولوژیکی تشکیل‌دهنده مغز حیوانات کشف شده است، ساخته

در این تحقیق، به منظور پیش بینی سطح اهمیت بهداشتی هر یک از کلان‌شهرهای مورد مطالعه، با توجه به داده های موجود در طول دوره همه‌گیری که شامل غلظت آلاینده ها و محدودیت های ترافیکی می باشد، از محدودیت های اعمالی به عنوان ویژگی استفاده شد. به همین منظور، محدودیت های ترافیکی به عنوان هدف (Target) در نظر گرفته شد. سطح اهمیت بهداشتی شامل شش کلاس می‌باشد که عبارتند از: هوای پاک، سالم، ناسالم برای گروه های حساس، ناسالم، بسیار ناسالم و خطرناک. با استفاده از شبکه آموزش داده شده توسط شبکه‌های عصبی فوق، می‌توان داده‌های روزانه، ماهانه و سالانه را در شبکه آموزش دیده قرار داد و بدین ترتیب پیش‌بینی لازم صورت خواهد پذیرفت.

شبکه عصبی مصنوعی (ANN)

Dropout به این معنا است که در حین آموزش این نورون‌ها، از تعدادی از آن‌ها به صورت تصادفی چشم‌پوشی شود. چشم‌پوشی یعنی اینکه آن نورون‌های خاص، در مسیر رفت یا برگشت در نظر گرفته نمی‌شوند. به عبارتی، Dropout یعنی اینکه در هر مرحله از آموزش، نودهایی از شبکه، با احتمال P-1 کنار گذاشته شده و نودهای دیگری با احتمال P، حفظ می‌شوند. بنابراین یک شبکه کاهش یافته باقی می‌ماند. Dropout از over-fitting جلوگیری می‌کند.

منطقه مورد مطالعه

منطقه مورد مطالعه، کلان‌شهرهای شیراز و مشهد می‌باشند. برای این تحقیق ۷ ایستگاه تحت نظارت ثبت پارامترهای آلودگی هوا انتخاب شدند. کلان‌شهر مشهد دارای ۵ ایستگاه تحت عناوین سجاد، ساختمان، لشگر، نخریسی و ویلا و کلان‌شهر شیراز دارای ۲ ایستگاه تحت عناوین دروازه کارزون و امام حسین می‌باشد. از هر ایستگاه، ۱۵۲۰ نمونه داده جمع‌آوری شد که ۷۹۰ نمونه مربوط به دوره Covid-19 و ۷۳۰ نمونه مربوط به دوره قبل از آن می‌باشد. لازم به ذکر است که شهرهای مورد مطالعه صرفاً شهرهای شیراز و مشهد بوده و کل استان‌های فارس و خراسان رضوی را شامل نمی‌شود. شکل ۱ موقعیت این شهرها نشان می‌دهد. داده‌های مرتبط با آلاینده‌های هوا از سازمان حفاظت محیط زیست کشور ایران به صورت روزانه از تاریخ ۱۳۹۶/۱۱/۰۱ الی ۱۴۰۰/۱۲/۲۹ در کلان‌شهرهای مشهد و شیراز جمع‌آوری شد. پارامترهای آلودگی هوا شامل CO ، O_3 ، NO_2 ، NO ، NO_x ، SO_2 ، $PM_{2.5}$ ، PM_{10} می‌باشد. داده‌های مربوط به دوران قبل از کرونا از تاریخ ۹۶/۱۱/۰۱ الی ۹۸/۱۰/۳۰ و داده‌های مربوط به دوره کرونا از تاریخ ۹۸/۱۱/۰۱ الی ۱۴۰۰/۱۲/۲۹ است.

داده‌های مرتبط با محدودیت‌های ترافیکی در دوره همه‌گیری کووید-۱۹ از وزارت کشور (مرکز ستاد ملی کرونا) دریافت شد که با توجه به تکامل محدودیت‌های کووید-۱۹ به پنج دوره تقسیم می‌شوند که عبارتند از محدودیت تردد بین استانی (ITR)، حضور یک سوم کارکنان (ATE)، تمدید خودکار اجاره‌نامه مستاجران (ART)، تعطیلی همه مشاغل از ساعت ۱۸:۰۰ به غیر از مشاغل گروه یک (CAB) و محدودیت تردد شبانه (NTR).

شده‌اند (هاردستی ۲۰۱۷). یک ANN بر اساس مجموعه‌ای از واحدها یا گره‌های متصل به نام نورون‌های مصنوعی است که نورون‌ها را در یک مغز بیولوژیکی مدل‌سازی می‌کنند. هر اتصال، مانند سیناپس‌های یک مغز بیولوژیکی، می‌تواند سیگنالی را به نورون‌های دیگر منتقل کند. یک نورون مصنوعی سیگنال‌ها را دریافت می‌کند و سپس آنها را پردازش می‌کند و می‌تواند به نورون‌های متصل به آن سیگنال دهد.

شبکه عصبی پیشخور با استفاده از TensorFlow و Keras

شبکه عصبی پیشخور (FNN) یکی از دو نوع گسترده شبکه عصبی مصنوعی است که با جهت جریان اطلاعات بین لایه‌های آن مشخص می‌شود (زل ۱۹۹۴). جریان آن یک طرفه است، به این معنی که اطلاعات در مدل تنها در یک جهت (به جلو) از گره‌های ورودی، از طریق گره‌های پنهان (در صورت وجود)، به گره‌های خروجی، بدون هیچ چرخه یا حلقه، جریان می‌یابد. برخلاف شبکه‌های عصبی بازگشتی (اشمایدن‌ها ۲۰۱۵) که دارای جریان دو جهته هستند، شبکه‌های پیش‌خور مدرن با استفاده از روش پس انتشار آموزش داده می‌شوند.

شبکه عصبی عمیق با Dropout

شبکه عصبی عمیق یک شبکه عصبی مصنوعی با چندین لایه بین لایه‌های ورودی و خروجی است (اشمایدن‌ها ۲۰۱۵؛ بنگیو ۲۰۰۹). شبکه‌های عصبی عمیق نیز مانند همه شبکه‌های عصبی دیگر از اجزای یکسانی تشکیل شده‌اند: نورون‌ها، سیناپس‌ها، وزن‌ها، سوگیری‌ها و توابع (گاوریلووا ۲۰۲۰) به دلیل اینکه DNN‌های پیچیده دارای لایه‌های زیادی هستند، شبکه‌های "عمیق" نامیده می‌شوند.

DNN‌ها می‌توانند روابط پیچیده غیرخطی را مدل کنند. معماری‌های DNN مدل‌های ترکیبی را تولید می‌کنند که در آن شی به عنوان ترکیب لایه‌ای از موارد اولیه بیان می‌شود (اسزیدی و همکاران ۲۰۱۳). لایه‌های اضافی ترکیب ویژگی‌ها را از لایه‌های پایین‌تر امکان‌پذیر می‌کنند و به طور بالقوه داده‌های پیچیده با واحدهای کمتر را نسبت به یک شبکه کم عمق با عملکرد مشابه مدل‌سازی می‌کنند (بنگیو ۲۰۰۹).

Dropout به معنای کنار گذاشتن بخش‌هایی از یک شبکه عصبی است. در یک شبکه عصبی که شامل تعدادی نورون است،

1 Interprovincial TRaffic

2 ATendance limit of one third of Employees

3 Automatic Renewal of Tenants' leases

4 Closing of All Businesses from 18:00, except for group 1 businesses

5 Night Traffic Restrictions



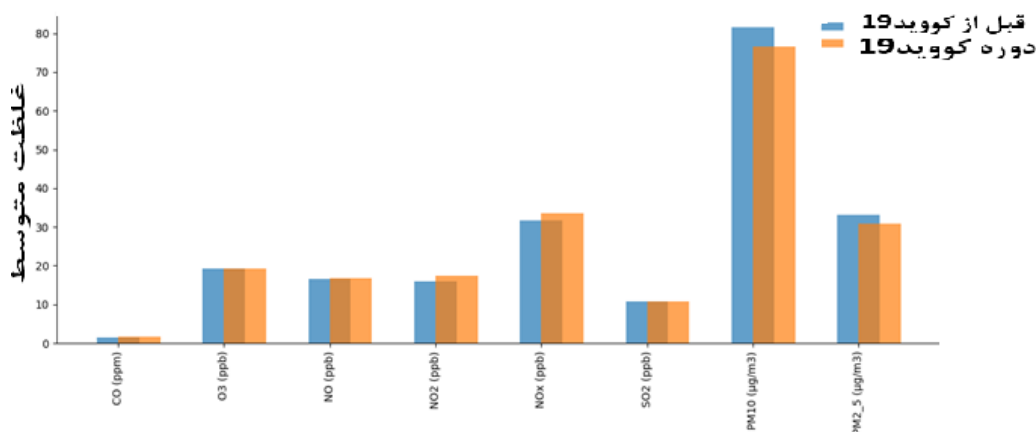
شکل ۱- موقعیت شهرهای مورد مطالعه

در این قسمت، تغییر غلظت آلاینده‌ها در دوره کووید-۱۹ نسبت به دوره قبل از همه‌گیری به تفکیک کلان‌شهرها در شکل‌های ۲ و ۳ نشان داده شده است تا مشخص گردد که اعمال محدودیت‌های ترافیکی چه تاثیری بر غلظت آلاینده‌ها داشته است. این مقایسه، بینش‌هایی را در مورد تغییرات مقادیر میانگین غلظت آلاینده‌ها در دوره‌های "قبل از همه‌گیری" و "در طول همه‌گیری" ارائه می‌دهد. جدول شماره ۲ نیز نتایج تغییر غلظت آلاینده‌ها را با سایر تحقیقات مقایسه می‌کند.

نتایج و بحث

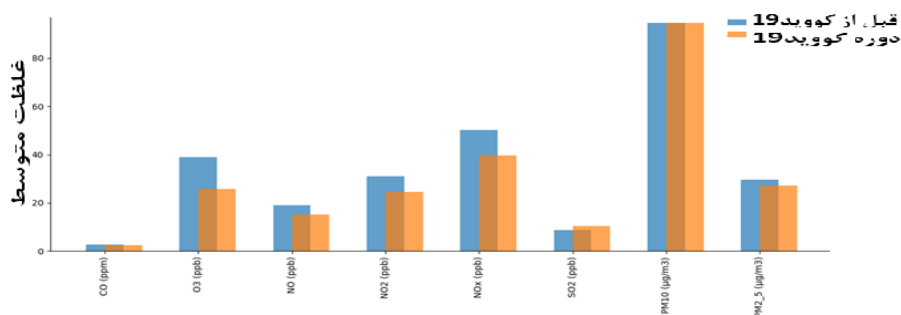
در این بخش، نتایج آزمایش و مدل‌های پیش‌بین ارائه می‌شود. آزمایشات بر روی مجموعه داده تهیه شده برای آلودگی هوای کلان‌شهرهای مشهد و شیراز انجام می‌شود.

تغییر غلظت آلاینده‌ها در دوره کووید-۱۹ نسبت به قبل از آن



ویژگی‌های آلودگی هوا

شکل ۲- مقایسه غلظت آلاینده‌های هوا (شهر مشهد)



ویژگی‌های آلودگی هوا

شکل ۳- مقایسه غلظت آلاینده‌های هوا (شهر شیراز)

۲۰۲۲ بیان نمودند که غلظت آلاینده‌های PM_{10} ، $PM_{2.5}$ ، NH_3 ، CO ، NO_x ، بنزن و تولوئن در طول منع رفت و آمد شبانه در مقایسه با دوره قبل از قرنطینه به میزان ۴٪ تا ۱۶٪ کاهش یافته است که این تغییرات بسته به اینکه محدودیت‌های مربوط به کووید-۱۹ در طول شب یا روز اعمال می‌شد و اینکه چه فعالیت‌هایی محدود بود، بسته به زمان روز، متفاوت بود. نکته قابل توجه، عدم کاهش پارامترهای SO_2 و O_3 در این دوره بود. در مقایسه با دوره قرنطینه مشابه در سال ۲۰۲۰، به جز پارامتر O_3 ، در سایر پارامترها، به میزان ۳۱٪ تا ۱۲۹٪ افزایش مشاهده شد. باتی و همکاران در سال ۲۰۲۲، بیان کردند که میانگین تغییر $PM_{2.5}$ از قبل از کووید به کووید فعال، ۱۸٪ کاهش یافته است. پس از کووید-۱۹ فقط ۲٪ کاهش یافته است. PM_{10} از قبل از کووید به کووید فعال، ۱۹٪ کاهش یافته است، اما غلظت آلاینده پس از کووید، ۲۳٪ افزایش داشته است. جدول ۲ مقایسه نتایج تحقیق حاضر با سایر تحقیقات را نشان داد.

در مقایسه با سایر تحقیقات مشابه صورت گرفته نیز می‌توان گفت که جیووانی کالتیری و همکاران در سال ۲۰۲۰، شش شهر از هشت شهر پرجمعیت ایتالیا با شرایط آب و هوایی متفاوت را مورد تجزیه و تحلیل قرار دادند که در آنها سطوح NO_2 به طور قابل توجهی در تمام مناطق شهری کاهش یافت (از ۲۴٫۹٪ در میلان به ۵۹٫۱٪ در ناپل). در مقابل، غلظت O_3 تقریباً بدون تغییر باقی ماند و حتی کمی افزایش یافت (تا ۱۳٫۷٪ در پالرمو و ۱۴٫۷٪ در رم). PM_{10} ، کاهش تا ۳۱٫۵٪ (پالرمو) و افزایش تا ۷٫۳٪ (ناپل) را نشان داد؛ در حالی که $PM_{2.5}$ کاهش ۱۳٪ الی ۱۷٪ را نشان داد. در تحقیق جیم گونزالس و همکاران در سال ۲۰۲۲، پارامتر NO_x بیش از ۵۰٪ کاهش یافت، پارامتر O_3 در تضاد با پارامتر NO_x ، NO_x ، ۲۳٫۹٪ افزایش یافت و همچنین پارامترهای PM_{10} و $PM_{2.5}$ نیز دارای کاهش جزئی به میزان ۴٫۱٪ و ۲٫۳٪ بودند که نشان دهنده تأثیرپذیری کم پارامترهای PM_{10} و $PM_{2.5}$ از منابع ترافیک و قرنطینه می‌باشد. سینگ ساهاران و همکاران در سال

جدول ۲- مقایسه نتایج تغییر غلظت آلاینده‌ها در دوره کووید-۱۹ نسبت به قبل از آن با سایر تحقیقات

آلاینده	تغییرات	کلان‌شهر	مطالعات همسو
CO	افزایش	مشهد	چانگ (۲۰۲۱)
	کاهش	شیراز	سahاران (۲۰۲۲) - یانگ (۲۰۲۲) - ژیانو فنگ (۲۰۲۳)
O ₃	ثابت	مشهد	فاضل رستگار و سیواکومار (۲۰۲۳)
	کاهش	شیراز	لین (۲۰۲۲)
SO ₂	افزایش	شیراز	چانگ (۲۰۲۱) - ساهاران (۲۰۲۲) - یانگ (۲۰۲۲)
	ثابت	مشهد	-
NO, NO ₂ , NO _x	افزایش	مشهد	-
	کاهش	شیراز	پاردو (۲۰۲۲) - کالتیری (۲۰۲۰) - ساهاران (۲۰۲۲) - یانگ (۲۰۲۲) - آدای (۲۰۲۲) - ویجناندز (۲۰۲۲) - لین (۲۰۲۲) - فاضل رستگار و سیواکومار (۲۰۲۳) - ژیانو فنگ (۲۰۲۳) - بیکان (۲۰۲۴)
PM ₁₀	ثابت	شیراز	-
	کاهش	مشهد	پاردو (۲۰۲۲) - کالتیری (۲۰۲۰) - ساهاران (۲۰۲۲) - یانگ (۲۰۲۲) - ویجناندز (۲۰۲۲)
PM _{2.5}	افزایش	-	-
	کاهش	مشهد و شیراز	پاردو (۲۰۲۲) - کالتیری (۲۰۲۰) - ساهاران (۲۰۲۲) - یانگ (۲۰۲۲) - آدای (۲۰۲۲) - ویجناندز (۲۰۲۲) - ژیانو فنگ (۲۰۲۴) - بیکان (۲۰۲۳)

جدول ۲ نشان داد که وضعیت غلظت آلاینده‌ها در دوره کووید-۱۹ نسبت به قبل از آن در شیراز بهتر از مشهد بوده است؛ به گونه‌ای که در شهر شیراز فقط غلظت آلاینده SO_2 افزایش، آلاینده PM_{10} ثابت و غلظت سایر آلاینده‌ها کاهش یافته بود؛ اما در مشهد، غلظت آلاینده‌های PM_{10} و $PM_{2.5}$ کاهش، آلاینده‌های O_3 و SO_2 ثابت و سایر آلاینده‌ها افزایش یافتند. علت افزایش آلاینده SO_2 در شیراز می‌تواند مزوت سوزی، استفاده از ماشین‌های سنگین و همچنین صناعی باشد که سوخت آنها سبب تولید زیاد آلاینده SO_2 می‌شود. لازم به ذکر است که تقریباً همه سوخت‌های فسیلی دارای گوگرد هستند که در جریان احتراق به صورت دی‌اکسید گوگرد وارد هوا می‌شوند. گوگرد موجود در بنزین در مقایسه با سایر منابع تولید آن ناچیز است. عموماً بالاترین سطوح SO_2 نزدیک به مجتمع‌های صنعتی است، لذا علت افزایش این آلاینده در این شهر می‌تواند این موضوع باشد که مطالعه جداگانه‌ای را نیاز دارد. پس می‌توان نتیجه گرفت که بالا بودن غلظت آلاینده SO_2 لزوماً به دلیل افزایش تردد خودروهای شخصی نبوده است. نکته جالب توجه دیگر، افزایش اکسیدهای نیتروژن و مونوکسید کربن در مشهد می‌باشد. مهم‌ترین منابع انسانی انتشار NO_2 ، فرآیندهای احتراق نظیر وسایط نقلیه، سیستم‌های گرمایش و نیروگاه‌ها می‌باشد. منبع عمده آن در مناطق شهری، حمل و نقل است. در خصوص CO نیز می‌توان گفت که منابع طبیعی و انسان‌ساخت هر دو در تشکیل مونوکسیدکربن سهیم هستند. به طور کلی منابع طبیعی در تولید مونوکسیدکربن نسبت به منابع انسان‌ساخت، سهم بیشتری دارند. علی‌رغم سهم کوچک منابع انسان‌ساخت در مقایسه با انتشار جهانی مونوکسیدکربن، تأثیر آنها نباید نادیده گرفته شود؛ زیرا منابع مصنوعی در مناطق بسیار کوچکی متمرکز شده‌اند. برای مثال ۹۸٪-۹۵٪ از مونوکسیدکربن در مناطق شهری در اثر فعالیت‌های انسانی حاصل می‌شود. بدیهی است حمل و نقل بیش از سایر فعالیت‌های انسانی (کارخانجات ذوب و پردازش فلزات، صنایع شیمیایی و زباله‌سوزها) در این امر نقش دارد.

همان‌گونه که از روند افزایش و یا کاهش غلظت آلاینده‌ها مشخص شد، شهرها از یک الگوی خاص پیروی نمی‌کنند و هر شهر الگوی خاص خود را دارد و در صورت اعمال محدودیت‌های یکسان برای شهرها، تغییر غلظت آلاینده‌ها متفاوت می‌باشد. عدم پیروی شهرها از الگوی یکسان در دوره کووید-۱۹، علی‌رغم اعمال محدودیت‌های مشابه را در تحقیق گالتیری و همکاران ایشان در سال ۲۰۲۰ در ایتالیا نیز می‌توان به وضوح مشاهده نمود. نتایج تحقیق گالتیری (۲۰۲۰) نشان داد که آلاینده PM_{10} در شهر پالمو، ۳۱/۵٪ کاهش و در شهر ناپولی تا ۷٪ افزایش یافته است. با مشاهده نتایج تحقیقات پیشین مشاهده شد

که آلاینده‌ها در کشورهای مختلف، تغییرات متفاوتی را در دوره کووید-۱۹ نسبت به قبل از آن از خود نشان داده‌اند. با نتایج حاصل شده می‌توان گفت که اعمال محدودیت‌های ترافیکی در دوره همه‌گیری، تأثیر قابل توجه و محسوسی در کاهش غلظت آلاینده‌های هوا در کلان‌شهر مشهد نداشته است. البته این موضوع می‌تواند دلایل مختلفی داشته باشد که از آن جمله می‌توان به تردد با وسیله نقلیه شخصی به جای استفاده از وسایل نقلیه عمومی در اثر ترس و نگرانی از ابتلا به بیماری اشاره کرد. دلیل دیگر آن برداشته شدن برخی از محدودیت‌های ترافیکی و طرح‌های ترافیک به دلیل فراهم نمودن امکان استفاده از وسایل نقلیه شخصی توسط افراد در راستای رعایت فاصله‌گذاری اجتماعی می‌باشد. یکی از دلایل مهم دیگر که جای تامل داشته و می‌تواند در تحقیقات آتی نیز مورد توجه قرار گیرد، تأثیر سایر عوامل ایجاد آلودگی هوا شامل صنایع، کارخانجات، پالایشگاه‌ها، نیروگاه‌ها، وارونگی دما، مصارف خانگی، زباله‌سوزی و موارد مشابه می‌باشد که ممکن است درصد بالایی از سهم آلودگی را شامل شده و به عبارتی در مقایسه با تأثیر آلودگی ناشی از تردد خودروها سهم قابل توجهی داشته باشد. نکته مهم دیگری که می‌توان به آن اشاره نمود، موقعیت جغرافیایی دو کلان‌شهر مشهد و شیراز می‌باشد که می‌تواند به عنوان یک عامل تأثیرگذار در میزان غلظت آلاینده‌ها عمل نماید.

در تحقیقی که احمدی و همکاران ایشان (۱۳۹۸) با موضوع بررسی تغییرات مکانی و زمانی انتشار آلاینده‌های CO و NO_x ناشی از وسایل نقلیه موتوری در شهر رشت انجام دادند، مشاهده شد که سهم اصلی انتشار CO در ترافیک رشت را خودروهای سواری دارا می‌باشند؛ به گونه‌ای که حدود ۸۸٪ آلاینده CO توسط خودروهای سواری تولید می‌شود. در مورد NO_x نیز وضعیت مشابه CO می‌باشد، با این تفاوت که حدود ۵۵٪ آلاینده NO_x توسط سواری‌ها تولید می‌شود. افزایش این آلاینده در کلان‌شهر مشهد می‌تواند موید این موضوع باشد که علی‌رغم اعمال محدودیت‌های تردد، تردد با خودروی شخصی در دوره کووید-۱۹ به دلایلی که در بالا عنوان گردید افزایش یافته است.

الگوهای پیش بین با استفاده از شبکه‌های عصبی

اساس این تحقیق، بررسی آلاینده‌های هر شهر به صورت جداگانه و عدم تعمیم آنها به یکدیگر می‌باشد؛ به گونه‌ای که برای هر یک از کلان‌شهرها، الگوی جداگانه‌ای در نظر گرفته شود. بدین منظور، شبکه‌های عصبی پیشخور و عمیق در نظر گرفته می‌شود و بر همین اساس اقدام به کدنویسی برای هر شهر شده و سپس نمودارهای مربوطه استخراج و تحلیل می‌گردد. مدل‌های به دست آمده می‌توانند در پیش‌بینی کیفیت هوا بر اساس ویژگی‌های در نظر گرفته شده مورد استفاده قرار گرفته و مسئولان مربوطه را در

لیبل که دارای ۶ کلاس سطح اهمیت بهداشتی به صورت زیر است، در داده‌ها نمایش داده می‌شود.
 کلاس ۱: پاک کلاس ۲: سالم کلاس ۳: ناسالم برای گروه‌های حساس کلاس ۴: ناسالم کلاس ۵: بسیار ناسالم کلاس ۶: خطرناک
 مشخصات ساختارها در شبکه‌های عصبی استفاده شده نیز در جدول ۳ عنوان شده است:

تصمیم‌گیری‌های مرتبط با آلودگی هوا یاری رسانند. محدودیت‌های ترافیکی تعریف شده عبارتند از محدودیت تردد بین استانی (ITR)، حضور یک سوم کارکنان (ATE)، تمدید خودکار اجاره نامه مستاجران (ART)، تعطیلی همه مشاغل از ساعت ۱۸ به غیر از مشاغل گروه یک (CAB) و محدودیت تردد شبانه (NTR). اگر این محدودیت‌ها (ویژگی‌ها) وجود داشته باشند مقدار ۱ و در غیر اینصورت مقدار ۰ می‌گیرند. هدف، طبقه‌بندی نمونه‌ها به کلاس‌های مختلف است که توسط ستون

جدول ۳- مشخصات ساختار شبکه‌های عصبی

ساختار	نام شبکه	مدل شبکه عصبی
FNN	شبکه عصبی پیشخور با استفاده از TensorFlow و Keras	لایه ورودی با ۶۴ نورون و عملکرد فعال سازی ReLU. لایه پنهان با ۳۲ نورون و عملکرد فعال سازی ReLU. لایه خروجی با ۶ نورون (برای ۶ کلاس) و عملکرد فعال سازی softmax.
DNN	شبکه عصبی عمیق با Dropout	لایه ورودی: لایه متراکم با ۱۲۸ نورون (واحد). بعد ورودی با تعداد ویژگی‌های مجموعه داده ما تعیین می‌شود (input_dim=X_train.shape[1]). تابع فعال سازی واحد خطی اصلاح شده (ReLU) است. لایه حذفی: لایه‌های Dropout با تنظیم تصادفی کسری از واحدهای ورودی بر روی ۰ در طول تمرین، به جلوگیری از برآزش بیش از حد کمک می‌کنند. در این حالت، پس از اولین لایه متراکم، یک لایه حذف با نرخ افت ۰.۵ (۵۰٪) اضافه می‌شود. لایه پنهان: لایه متراکم با ۶۴ نورون. تابع فعال سازی ReLU است. لایه حذفی: بعد از لایه دوم Dense یک لایه دیگر با نرخ افت ۰.۵ اضافه می‌شود. لایه پنهان: لایه متراکم با ۳۲ نورون. تابع فعال سازی ReLU است.

مطلقاً، حداکثر خطای مطلق و ضریب همبستگی^۴ استفاده شد. هر چقدر مقدار خطا در شبکه کمتر باشد نشان از آن است که شبکه به خوبی آموزش دیده است و داده‌های تست را با درصد خطای کمتری می‌تواند پیش‌بینی کند.

شایان ذکر است به منظور به دست آوردن میزان خطای مدل با توجه به داده‌های ورودی داده شده به مدل جهت آموزش و براساس معیارهای مختلف محاسبه خطا، از معیارهای میانگین مربعات خطا (MSE) میانگین مربعات خطای نرمال شده (NMSE)، میانگین خطای مطلق (MAE)، حداقل خطای

4Minimum Absolute Error
 5Maximum Absolute Error
 6Correlation Coefficient

1Mean Squared Error
 2Normalized Mean Squared Error
 3Mean Absolute Error

داد. دقت بالای تست نشان داد که مدل روی داده‌های دیده نشده عملکرد خوبی داشته است. نتایج ساختارها در شبکه‌های عصبی استفاده شده در جدول ۴ ارائه شده است که دارای معیارهای عملکردی برای مدل‌های مختلف است که به نوعی به پیش‌بینی یا تخمین مربوط می‌شود.

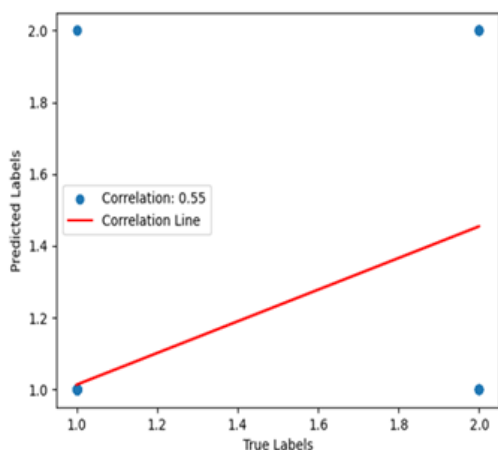
در این مرحله به بررسی داده‌های هر یک از کلان‌شهرها با استفاده از پایتون پرداخته شده است. دقت تست شبکه عصبی آموزش دیده بر روی مجموعه داده شهر مشهد تقریباً ۹۸٪ و داده شهر شیراز ۹۵٪ بود. این معیار نسبت نمونه‌های به‌درستی طبقه‌بندی شده در مجموعه آزمایشی را نشان

جدول ۴- نتایج معیارهای عملکردی برای مدل‌های شبکه عصبی مصنوعی در شهرهای مشهد و شیراز

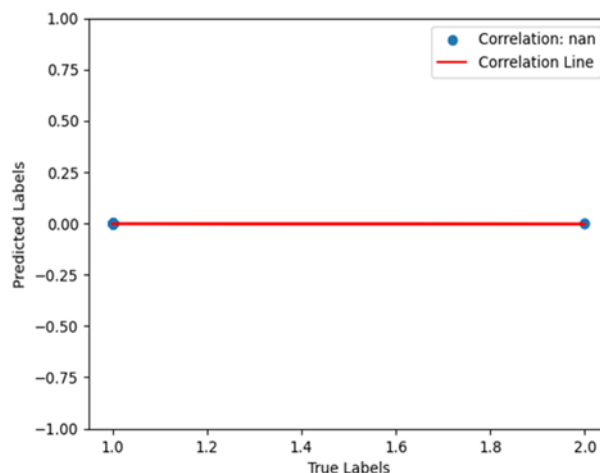
شهر	ساختار	میانگین مربعات خطا	میانگین مربعات خطای نرمال شده	میانگین خطای مطلق	مینیمم خطای مطلق	ماکزیمم خطای مطلق	ضریب همبستگی
مشهد	FNN	۰/۰۱۲۴۵	۰/۰۱۲۴۵	۰/۰۱۲۵۳	۰	۱	۰
	DNN	۰/۰۱۲۲۱	۰/۰۱۲۲۱	۰/۰۱۱۸۸	۰	۱	۰
شیراز	FNN	۰/۰۵۰۶۳	۰/۰۵۰۶۳	۰/۰۴۸۳۶	۱	۲	۰/۵۵۳۰۲
	DNN	۰/۰۵۱۲۹	۰/۰۵۱۲۹	۰/۰۴۹۰۱	۱	۲	۰/۵۴۵۳۷

مقادیر ارائه شده برای هر مدل (DNN و FNN) نشان داد که آنها عملکرد مشابهی در اکثر معیارها، با تغییرات جزئی داشته‌اند. مدل‌ها عموماً دارای میانگین مربعات خطاهای پایین، میانگین مربعات خطاهای نرمال شده و میانگین خطاهای مطلق هستند. حداقل خطای مطلق ۰ بود که نشان داد مواردی وجود دارد که پیش‌بینی‌ها کاملاً با مقادیر واقعی مطابقت داشته‌اند.

مقدار ضریب همبستگی نشان دهنده عدم همبستگی و یا همبستگی خیلی ضعیف بین مقادیر پیش‌بینی شده و واقعی برای این مدل می‌باشد. نمودار پراکندگی با خط همبستگی به ما بینشی در مورد رابطه بین برچسب‌های واقعی (y_{true}) و برچسب‌های پیش‌بینی شده (y_{pred_labels}) داد. ضریب همبستگی یک معیار عددی برای قدرت و جهت این رابطه است و نمودار پراکندگی به صورت بصری این رابطه را نشان می‌دهد. شکل‌های شماره ۴ و ۵، به ترتیب نمودار پراکندگی با خط همبستگی برای داده‌های شهر مشهد و شیراز را نشان می‌دهد.



شکل ۵- نمودار پراکندگی با خط همبستگی (برای داده‌های شهر شیراز)



شکل ۴- نمودار پراکندگی با خط همبستگی (برای داده‌های شهر مشهد)

است. در شهر مشهد، همبستگی ضعیف و یا عدم همبستگی مشاهده شد، در حالی که در شهر شیراز، شاهد یک ضریب همبستگی مثبت و برابر با $0/55$ بودیم که نشان داد پیش‌بینی‌های مدل با افزایش مقادیر واقعی تمایل به افزایش داشتند. در مرحله بعد، مدل شبکه عصبی از پیش آموزش دیده از فایل ذخیره شده بارگیری گردید. این فایل باید شامل معماری و وزن مدل آموزش دیده باشد. داده‌های تست جدید با استفاده از مقیاس‌کننده مشابهی که در طول آموزش شبکه عصبی استفاده شده بود، استاندارد شد. این موضوع تضمین کرد که ویژگی‌ها به طور مداوم تغییر می‌کنند. مدل از پیش آموزش داده شده برای پیش‌بینی داده‌های آزمون جدید مقیاس شده استفاده شد. سپس داده‌های جدول ۵ را به عنوان مجموعه داده ورودی به برنامه در نظر گرفتیم.

هر نقطه در نمودار پراکندگی نشان دهنده یک جفت برچسب درست و پیش‌بینی شده برای یک نمونه خاص در مجموعه آزمایشی بود. هر نقطه در نمودار مربوط به یک نمونه در مجموعه آزمایشی است و موقعیت آن نشان داد که پیش‌بینی‌های مدل چقدر با برچسب‌های واقعی همسو بودند. اگر خط همبستگی دارای شیب مثبت باشد، همبستگی مثبت را نشان می‌دهد، به این معنی که با افزایش مقادیر واقعی، مقادیر پیش‌بینی شده نیز تمایل به افزایش دارند. مقدار عددی ضریب همبستگی، قدرت و جهت رابطه خطی بین برچسب‌های واقعی و پیش‌بینی شده را نشان می‌دهد. ضریب همبستگی نزدیک به ۱، یک همبستگی خطی مثبت قوی را نشان می‌دهد، در حالی که مقدار نزدیک به -1 نشان دهنده یک همبستگی خطی منفی قوی است. مقدار نزدیک به ۰ نشان دهنده یک همبستگی خطی ضعیف یا بدون همبستگی

جدول ۵- مجموعه داده ورودی جهت پیش‌بینی سطح اهمیت بهداشتی

روزهای هفته	ITR	ATE	ART	CAB	NTR
روز اول	۱	۱	۰	۰	۰
روز دوم	۱	۱	۰	۰	۰
روز سوم	۱	۱	۰	۰	۰
روز چهارم	۱	۱	۰	۱	۰
روز پنجم	۱	۱	۰	۰	۰
روز ششم	۱	۰	۰	۰	۰
روز هفتم	۱	۰	۰	۰	۰

به عنوان ورودی به برنامه در نظر گرفتیم. برنامه با استفاده از مدل از پیش آموزش دیده، مقادیر هدف که همان کلاس‌های ۶ گانه مورد نظرمان درباره هوا بود را پیش‌بینی کرد. قابل ذکر است که داده‌ها برای یک هفته در نظر گرفته شده بودند و این داده‌ها می‌توانستند به صورت روزانه، هفتگی، ماهانه یا سالانه در نظر گرفته شوند. نتیجه خروجی برنامه برای پیش‌بینی مقادیر هدف در جدول ۶ آورده شده است.

ستون اول مربوط به ویژگی محدودیت تردد بین استانی، ستون دوم مربوط به ویژگی حضور یک سوم کارکنان، ستون سوم مربوط به ویژگی تمدید خودکار اجازه نامه مستاجران، ستون چهارم مربوط به ویژگی تعطیلی همه مشاغل از ساعت ۱۸ به غیر از مشاغل گروه یک و در نهایت ستون پنجم مربوط به ویژگی محدودیت تردد شبانه بود. در این حالت اگر هر ویژگی‌ای در روز مختص به خودش وجود داشت، مقدار عددی ۱ و در غیر اینصورت ۰ در نظر گرفته شد. پس از آماده‌سازی مجموعه داده، این مجموعه داده را

جدول ۶- پیش‌بینی کلاس و سطح اهمیت بهداشتی مطابق با مجموعه داده ورودی شهرهای مشهد و شیراز

روز	مشهد		شیراز	
	کلاس پیش‌بینی شده	سطح اهمیت بهداشتی	کلاس پیش‌بینی شده	سطح اهمیت بهداشتی
اول	۲	سالم	۲	سالم
دوم	۲	سالم	۲	سالم
سوم	۲	سالم	۲	سالم
چهارم	۲	سالم	۳	ناسالم برای گروه‌های حساس
پنجم	۲	سالم	۲	سالم
ششم	۲	سالم	۲	سالم
هفتم	۲	سالم	۲	سالم

نتیجه‌گیری و پیشنهادها

با توجه به اعمال محدودیت‌های ترافیکی صورت گرفته در دوره همه‌گیری کووید-۱۹، تصمیم بر آن شد تا برای هر یک از کلان‌شهرهای شیراز و مشهد الگویی استخراج شود تا با استفاده از آموزش داده‌های جمع‌آوری شده، بتوان شاخص کیفیت هوا را با اعمال محدودیت‌های مختلف در روزها، هفته‌ها، ماه‌ها و حتی سالهای متفاوت پیش‌بینی کرد. به همین منظور با استفاده از شبکه‌های عصبی، اقدام به کدنویسی شد تا با توجه به داده‌های جمع‌آوری شده برای هر شهر، الگوی هر شهر استخراج گردد. جهت تست مدل نیز، هفت روز در نظر گرفته شد و ملاحظه شد که با اعمال محدودیت‌های یکسان در دو کلان‌شهر، شاخص کیفیت هوا می‌تواند در آنها متفاوت باشد. با توجه به توضیحات فوق می‌توان گفت مطالعاتی که به بررسی داده‌های یک یا چند شهر پرداخته و یا نتایج چند شهر را با هم ادغام نموده و آن را به کل کشور و یا حتی جهان تعمیم می‌دهند، ممکن است دارای خطا بوده و تصمیم‌گیران حوزه زیست محیطی را دچار اشتباه نمایند. مطالعه حاضر نتیجه تغییرات غلظت آلاینده‌های هوا در اثر محدودیت‌های ترافیکی اعمال شده در کلان‌شهرهای مشهد و شیراز در طول دوره همه‌گیری کووید-۱۹ را نیز نشان داد.

در مقایسه غلظت آلاینده‌ها در دوره کووید-۱۹ نسبت به قبل از آن در شهر شیراز، مشاهده شد که بیشترین کاهش مربوط به O_3 و اکسیدهای نیتروژن به ترتیب به میزان ۳۴ درصد و ۲۱ درصد می‌باشد؛ این درحالی است که در شهر مشهد، غلظت آلاینده O_3 تغییر محسوسی نداشته و غلظت اکسیدهای نیتروژن نیز به طور متوسط به میزان ۷ درصد افزایش یافته است. تغییر غلظت آلاینده $PM_{2.5}$ در هر دو شهر تقریباً نزدیک به هم می‌باشد به گونه‌ای که غلظت آن در شهر شیراز، ۸ درصد و در شهر مشهد ۷ درصد کاهش یافته است. غلظت آلاینده CO در شهر شیراز، ۱۱ درصد کاهش و در شهر مشهد، ۹ درصد افزایش یافته است. نکته جالب توجه در این تحقیق، عدم تغییر محسوس آلاینده PM_{10} و همچنین افزایش غلظت آلاینده SO_2 به میزان ۱۸ درصد در شهر شیراز می‌باشد و این در حالی است که در شهر مشهد، غلظت آلاینده PM_{10} به میزان ۶ درصد کاهش یافته و غلظت آلاینده

SO_2 نیز تغییر محسوسی نداشته است. عدم تغییر غلظت آلاینده PM_{10} در شهر شیراز که دارای محدودیت تردد خودرو در دوره همه‌گیری بوده است، می‌تواند مصرف بالای سوخت‌های فسیلی (به ویژه سوخت‌های سرب‌دار)، وجود کارخانه سیمان، تخریب پوش گیاهی و تغییر کاربری مراتع و زمین‌های کشاورزی رها شده در اطراف شهر شیراز باشد.

با توجه به نتایج حاصل شده از مقایسه غلظت آلاینده‌ها در دو کلان‌شهر مشهد و شیراز، الگوی افزایشی و یا کاهشی یکسانی در این دو شهر دیده نمی‌شود، بنابراین می‌توان نتیجه گرفت که تاثیر محدودیت‌ها در شهرهای مختلف، بر روی غلظت آلاینده‌ها متفاوت می‌باشد. لذا اعمال محدودیت‌های یکسان در تمامی شهرها، لزوماً منجر به کاهش آلودگی نشده و می‌بایست برای برای هر موقعیت شهری و زیست‌محیطی، الگوی محدودیت‌های ترافیکی مختص آن موقعیت تهیه گردد. از تحقیق حاضر و همچنین سایر تحقیقات مشابه صورت گرفته می‌توان نتیجه گرفت که نمی‌توان نتایج تحقیق یک یا چند شهر و یا کشور را به سایر مناطق تعمیم داد.

به منظور تهیه و تدوین سند راهبردی و مدیریت بحران آلودگی، می‌توان از الگوهای پیش بین استخراج شده از تحقیق حاضر بهره برد. الگوی حاصل از این تحقیق را می‌توان به صورت روزانه، هفتگی، ماهانه و سالانه برای هر یک از دو کلان‌شهر مورد استفاده قرار داد و نتیجه گرفت که کدام محدودیت‌ها در کدام شهر سبب خواهد شد تا بتوان شاخص کیفیت هوا را کنترل کرد. پیشنهاد می‌گردد که تعداد ایستگاه‌های پایش آلودگی هوا در هر شهر به صورت استاندارد در نظر گرفته شود تا خطای الگوهای پیش بین استخراج شده کاهش یابد و سازمان‌های ذیربط از تعمیم الگوها به شهرهای دیگر خودداری نموده و برای هر شهر، مطالعه جداگانه‌ای انجام داده و الگوهای مربوطه را استخراج و بکار گیرند. همچنین جهت تاثیرپذیری غلظت آلاینده‌ها از منابع انتشار مختلف، پیشنهاد می‌گردد درصد سهم منابع انتشار آلودگی در تولید و افزایش غلظت هر یک از آلاینده‌ها مشخص گردد که در این مورد نیز با توجه به متفاوت بودن تراکم جمعیت در شهرهای مختلف، صنعتی بودن شهر، شرایط جغرافیایی و ...، نتایج به دست آمده می‌تواند در نوع خود جالب باشد.

References

1. Ahmadi Arkami A, Kamkar A, Aghajani Z. Investigating the spatial and temporal changes of CO and NOX emissions caused by motor vehicles using MOVES software: a case study of Beheshti Boulevard and Modares in Rasht. Iranian Journal

of Health and Environment. 1398. 12th volume, 2nd issue. pages 203-216. <https://ijhe.tums.ac.ir/article-1-6212-fa.html> [In Persian]

2. Bengio Y. Learning Deep Architectures for AI (PDF). Foundations and Trends in Machine Learning. 2009. 2 (1): 1–127. CiteSeerX 10.1.1.701.9550. S2CID 207178999. Archived from the original (PDF) on 4 March 2016. Retrieved 3 September 2015. doi: 10.1561/2200000006
3. Berry M, Linoff G. Mastering Data Mining, John Wiley, Hoboken, NJ. 2000. doi.org/10.1108/imds.2000.100.5.245.2
4. Berson A, Smith S, Thearling K. Building Data Mining Applications for CRM, McGraw Hill. 2000. <https://search.worldcat.org/title/building-data-mining-applications-for-crm/oclc/47008776>
5. Bhatti U.A, Zeeshan Z, Nizamani M.M, Bazai S, Yu Z, Yuan.L. Assessing the change of ambient air quality patterns in Jiangsu Province of China pre-to post-COVID-19. 2022. Chemosphere, Vol.288, 132569. doi: 10.1016/j.chemosphere.2021.132569
6. Bherwani H, Gautam S, Gupta A. Qualitative and quantitative analyses of impact of COVID-19 on sustainable development goals (SDGs) in Indian subcontinent with a focus on air quality. 2021. Int. J. Environ. Sci. Technol. 18 (4) (2021) 1019–1028. doi.org/10.1007/s13762-020-03122-z
7. Bherwani H, Kumar S, Musugu K, Nair M, Gautam S, Gupta A, Ho C-H, Anshul A, Kumar R. Assessment and valuation of health impacts of fine particulate matter during COVID-19 lockdown: a comprehensive study of tropical and sub tropical countries, Environ. Sci. 2021. Pollut. Res. 28 (32) (2021) 44522–44537. doi.org/10.1007/s11356-021-13813-w
8. Chang H, Chad D, Feng A. COVID-19 prevention, air pollution and transportation patterns in the absence of a lockdown. 2021. Journal of Environmental Management 298 (2021) 113522. doi.org/10.1016/j.jenvman.2021.113522
9. Chelani A, Gautam S. Lockdown during COVID-19 pandemic: a case study from Indian cities shows insignificant effects on persistent property of urban air quality, Geosci. Front. 2022. 13 (6) (2022), 101284. doi.org/10.1016/j.gsf.2021.101284
10. Fazel-Rastegar F, Sivakmar V. A case study of weather impact on air pollution during the COVID-19 pandemic in South Africa. Scientific African. 2023. Volume 22, November 2023, e01914. doi.org/10.1016/j.sciaf.2023.e01914
11. Gautam S, Samuel C, Gautam A, Kumar S. Strong link between coronavirus count and bad air: a case study of India. 2021. Environ. Dev. 2021. Sustain. 23 (11) (2021) 16632–16645. doi.org/10.1007/s10668-021-01366-4
12. Gautam S. The influence of COVID-19 on air quality in India: a boon or inutile, Bull. Environ. 2020. Contam. Toxicol. 104 (6) (2020) 724–726. doi.org/10.1007/s00128-020-02877-y
13. GavriloVA Y. A Guide to Deep Learning and Neural Networks. archived from the original on 2020-11-02, retrieved 2020-11-16. 2020. <https://serokell.io/blog/deep-learning-and-neural-network-guide>
14. Gonzalez-pardo J, Ceballos-Santos S, Manzanas R, Santibanez M, Fernandez-Olmo I. Estimating change in air pollutant levels due to COVID-19 lockdown measures based on a business-as-usual prediction scenario using data mining models: A case-study for urban traffic sites in Spain. 2022. Sci. Total Environ. 823, 153786. doi.org/10.1016/j.scitotenv.2022.153786
15. Gualtieri G, Brilli L, Carotenuto F, Vagnoli C, Zaldei A, Gioli B. Quantifying road traffic impact on air quality in urban areas: A COVID-19 –induced lockdown analysis in Italy. 2020. Environmental pollution 267(2020) 115682. doi.org/10.1016/j.envpol.2020.115682
16. Hardesty L. "Explained: Neural networks". MIT News Office. 2017. Retrieved 2 June 2022. <https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414>
17. Hidalgo García D, Diaz J.A, Impacts of the COVID-19 confinement on air quality, the Land Surface Temperature and the urban heat island in eight cities of Andalusia (Spain). 2021. Remote Sensing Applications: Society and Environment. 25. doi.org/10.1016/j.rsase.2021.100667
18. Kantardzic M. Data Mining: Concepts, Models, Methods, and Algorithms. 2003. John Wiley. <https://ieeexplore.ieee.org/book/5265979>
19. Kolluru S.S.R, Nagendra S.M, Patra A.K, Gautam S, Alshetty V.D, Kumar P. Did unprecedented air pollution levels cause spike in Delhi's COVID cases during second wave? 2022. in: Stochastic Environmental Research and Risk Assessment. pp. 1–16. doi.org/10.1007/s00477-022-02308-w
20. Larose D, Larose C. Discovering Knowledge in Data. 2014. John Wiley. doi:10.1002/9781118874059
21. Lin G. Y, Chen W. Y, Chieh S. H, & Yang Y. T. Chang impact analysis of level 3 COVID-19 alert on air pollution indicators using artificial neural network. 2022. Ecological informatics. 69, 101674. doi.org/10.1016/j.ecoinf.2022.101674
22. Malekzadeh Sh. Investigation and assessment of air pollution in cities and the need to implement laws to reduce pollution, the fourth national conference on urban planning, architecture, construction and environment. 1397. Shirvan. <https://civilica.com/doc/774175> [In Persian]
23. McDuffie E, Martin R, Yin H, Brauer M. Global Burden of Disease from Major Air Pollution Sources (GBD MAPS): A Global

- Approach. Research Report 210. 2021. Boston, MA:Health Effects Institute. PMID: 36148817
<https://pubmed.ncbi.nlm.nih.gov/36148817>
24. Munnoli P M, Nabapure S, Yeshavanth G. Post-COVID-19 precautions based on lessons learned from past pandemics: a review. 2020. *J. Public Health* 1–9. doi.org/10.1007/s10389-020-01371-3
 25. Sanchez-Lorenzo A, Vaquero-Martínez J, Calbó J, Wild M, Santurtún A, Lopez-Bustins J, Vaquero J, Folini D, Antón M. Did anomalous atmospheric circulation favor the spread of COVID-19 in Europe? 2020. *Environ. Res.* 194, 110626. doi.org/10.1016/j.envres.2020.110626
 26. Schmidhuber J. Deep Learning in Neural Networks: An Overview". *Neural Networks*. 2015. 61: 85–117. arXiv:1404.7828. PMID 25462637. S2CID 11715509. doi:10.1016/j.neunet.2014.09.003
 27. Shaygan M, Mokarram M. Investigating Air Pollution during the Corona Era and before that in the Metropolises of Tehran, Isfahan and Qom. 1402. doi.org/ 10.48308/GISJ.2023.103607 [In Persian]
 28. Singh Saharan U, Kumar R, Tripathy P, Sateesh M, Grag J, Kumar Sharma S, Kumar Mandal T. Drivers of air pollution variability during second wave of COVID-19 in Delhi, India. 2022. *Urban climate* 41 (2022) 101059. doi.org/10.1016/j.uclim.2021.101059
 29. Szegedy C, Toshev A, Erhan D. Deep neural networks for object detection. 2013. *Advances in Neural Information Processing Systems*: 2553–2561. Archived from the original on 2017-06-29. Retrieved 2017-06-13.
https://proceedings.neurips.cc/paper_files/paper/2013/file/f7cade80b7cc92b991cf4d2806d6bd78-Paper.pdf
 30. Uday U, Bethineedi L. D, Hasanain M, Ghazi B. K, Nadeem A, Patel P, Khalid Z. Effect of COVID-19 on air pollution related illnesses in India. 2022. *Annals of medicine and surgery* 78, 103871. doi.org/10.1016/j.amsu.2022.103871
 31. Wijnands J. S, Nice K.A, Seneviratne S, Thompson J, Stevenson M. The impact of the COVID-19 pandemic on air pollution: A global assessment using machine learning techniques. *Atmospheric pollution research*. 2022. doi.org/10.1016/j.apr.2022.101438
 32. Yang M, Chen L, Msigwa G, Tang K.H.D, Yap p-s. Implications of COVID-19 on global environmental pollution and carbon emissions with strategies for sustainability in the COVID-19 era. 2022. *Sci. Total Environ*, Vol.809, 151657. doi.org/10.1016/j.scitotenv.2021.151657
 33. Yiqun M, Federica N, Anne M, Robert D, Patrick L, Annette P, Massimo S, Susanne B, Kai C. Air pollution changes due to COVID-19 lockdowns and attributable mortality changes in four countries. 2024. doi.org/10.1016/j.envint.2024.108668
 34. Zell A. *Simulation Neuronaler Netze [Simulation of Neural Networks]* (in German) (1st ed.). 1994. Addison-Wesley. p. 73. ISBN 3-89319-554-8. <https://www.amazon.com/Simulation-Neuronaler-Netze-German-Andreas/dp/3486243500>
 35. Zihao F, Xuhong W, Jiabin Y, Ying Z, Mengqianxi Y. Changes in air pollution, land surface temperature, and urban heat islands during the COVID-19 lockdown in three Chinese urban agglomerations. 2023. *Science of the Total Environment* 892 (2023) 164496. doi.org/10.1016/j.scitotenv.2023.164496



Data Mining Model Based on Machine Learning to Predict Air Pollution in The Metropolises of Iran

Abbas Maleki PhD Student, Department of Industrial Management, Qazvin Branch, Islamic Azad University, Qazvin, Iran
Sadegh Abedi * Assistant Professor, Department of Industrial Management, Qazvin Branch, Islamic Azad University, Qazvin, Iran
Alireza Irajpour Assistant Professor, Department of Industrial Management, Qazvin Branch, Islamic Azad University, Qazvin, Iran

Extended Abstract

Received: 28 April 2024

Accepted: 29 May 2024

Keywords: Covid-19, Iranian Metropolises, Machine Learning, Neural Network, Prediction Of Air Pollution.

Introduction: In response to the Covid-19 pandemic, governments around the world have imposed severe restrictions and presented different scenarios of reducing emissions from traffic sources. With the application of traffic restrictions due to the covid-19 epidemic and quarantine, it seemed that there was a reduction in the traffic of vehicles and the concentration of pollutants and the air quality index approached the quality standards. Therefore, it was expected to see changes in the concentration of CO, O₃, NO, NO₂, NO_x, SO₂, PM_{2.5} and PM₁₀ pollutants, which are pathogenic factors and sometimes premature death.

Materials and Methods: Using the data mining method, in the first stage, the change in the concentration of pollutants in the period of the Covid-19 epidemic compared to the period before it is investigated in order to determine what effect the application of traffic restrictions has on the change in the concentration of pollutants in each from the metropolises of Tehran and Shiraz. In the second stage, predictive models are presented using feedforward and deep neural networks to predict the level of health importance based on the application of each of the traffic restrictions in each metropolis.

Results and Discussion: This study shows that the change in the concentration of pollutants during the covid-19 era is different compared to before, in the cities of Tehran and Shiraz; In such a way that the concentration of most pollutants increased in Tehran metropolis and decreased in Shiraz metropolis. According to the result obtained and the difference in the process of changing the concentration of pollutants, in order to control the air quality index, predictive models were presented for each metropolis.

Conclusion: For pollutants, the same increasing or decreasing pattern is not seen in the studied metropolises, so it can be said that the effect of the same restrictions on changing the concentration of pollutants is different in different cities; Therefore, applying the same restrictions in all cities does not necessarily lead to a reduction in pollution, and for each urban and environmental situation, a model of traffic restrictions specific to that situation should be prepared.

Corresponding author: Sadegh Abedi

Address: Department of Industrial Management, Qazvin Branch, Islamic Azad University, Qazvin, Iran. **Tel:** +989128173662

Email: abedi.sadegh@gmail.com

Citation: Maleki A, Abedi S, Irajpour A. Data Mining Model Based on Machine Learning to Predict Air Pollution in The Metropolises of Iran, *New Researches in Environmental Engineering*, 2024; 2(5): 39-53.



© 2024, This article published in Journal of New Researches in Environmental Engineering (JNREE) as an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>). Non-commercial use, distribution and reproduction of this article is permitted in any medium, provided the original work is properly cited.