

پیش‌بینی کوتاه‌مدت آلاینده ذرات معلق ناشی از جریان ترافیک با استفاده از ماشین بردار پشتیبان (مطالعه موردی: شهر تهران)

محمد رضا کاشی‌پزان قمی^۱، بهروز شیرگیر^{۲*}

۱- دانشجوی کارشناسی ارشد مهندسی راه و ترابری، دانشگاه خوارزمی

۲- استادیار دانشکده فنی و مهندسی، گروه عمران، دانشگاه خوارزمی

*shirgir@khu.ac.ir

تاریخ پذیرش: [۱۳۹۷/۰۳/۲۱]

تاریخ دریافت: [۱۳۹۶/۱۲/۰۷]

چکیده

در دهه های اخیر، افزایش تراکم جمعیت و فعالیت های اقتصادی و صنعتی در شهرهای بزرگ، افزایش حجم ترافیک و در نتیجه افزایش سطح آلودگی هوا را به دنبال داشته است. منبع اصلی آلودگی هوا در شهرهای بزرگ و در حال توسعه حمل و نقل وسایل نقلیه ای است که از سوخت های معمولی استفاده می کنند و انرژی و ترافیک سنگین در خیابان های این شهرها اغلب در مشکلاتی مانند مدیریت ترافیک و فرهنگ ترافیک کمبود وجود دارد. یکی از مسائل حائز اهمیت در شهرها و کلان شهرهایی که با معضلات و آثار زیان بار آلودگی مواجه هستند، مسئله اطلاع رسانی وضعیت آلاینده کیفیت هوا و میزان آلودگی هوای شهری به مردم است. این مهم می تواند از طریق پیش بینی های روزانه یا حتی ساعتی وضعیت آلودگی هوا میسر شود و از قرارگیری افراد جامعه در مکان های آلوده و تبعات جبران ناپذیر آن جلوگیری کند. بنابراین نیاز به پیش بینی وضعیت کیفی هوا و تخمین های کمی از غلظت آلاینده ها در پی آمدن وسایل نقلیه احساس می شود که در این پژوهش به صورت موردی به مسئله پیش بینی ساعتی غلظت آلاینده ذرات معلق (PM2.5) در منطقه ۱۱ شهرداری تهران پرداخته شده که در حدود ۸۰ درصد روزهای آلوده سال تحت اثر این آلاینده از حد سالم تجاوز کرده است. روش مورد استفاده برای پیش بینی در این پژوهش، یکی از روش های تحلیل شبکه های عصبی با نام ماشین بردار پشتیبان (SVM) است. تاکنون انواع مختلفی از شبکه های عصبی مصنوعی معرفی شده است، این روش ها عمدتاً در برنامه های کاربردی مانند طبقه بندی، خوشه بندی، شناخت الگو، مدل سازی و تقریب توابع (یا رگرسیون)، کنترل، برآورد و بهینه سازی مورد استفاده می شود. ماشین های بردار پشتیبان یک نوع شبکه عصبی ویژه ای هستند که بر خلاف سایر انواع شبکه های عصبی مانند چند لایه پرسپترون MLP و به جای کم کردن خطا، خطرهای عملیاتی را کاهش می دهد. نتایج حاصل از این تحقیق نشان داد که مدل های بردار پشتیبان در پیش بینی سهم و مشارکت ترافیک ساعتی شهری در انتشار ذرات معلق به شدت خوب عمل می کنند و پیش بینی ها به خوبی با مشاهدات هماهنگی دارند و این فرصت را فراهم می کند تا به عنوان ابزار مدیریت کیفیت هوا به کار روند.

واژگان کلیدی: پیش بینی آلودگی هوا، ماشین بردار پشتیبان، جریان ترافیک، ذرات معلق، مدیریت کیفیت هوا

۱- مقدمه

به حمل و نقل انبوه خودروهایی است که بیش از حد استاندارد، سوخت و انرژی مصرف می کنند و بار سنگین ترافیک خیابان های این شهرها، بیشتر ریشه در معضلاتی مانند ضعف مدیریت ترافیک و فرهنگ ترافیکی دارد [1]. یکی از مهم ترین مشخصه های کیفیت شبکه حمل و نقل، میزان آلودگی ناشی از آمدن و

در دهه های اخیر، افزایش تراکم جمعیت و فعالیت های اقتصادی و صنعتی در کلان شهرها، باعث افزایش حجم ترافیک و در نتیجه، بالا رفتن سطح آلودگی هوا شده است. عمده ترین منبع آلوده کننده هوا در شهرهای بزرگ در حال توسعه مربوط

تبعات جبران ناپذیر آن جلوگیری کند؛ همچنین نهادهای سیاست گذار در بخش آلودگی هوا و ترافیک می توانند در زمان مناسب اقدامات لازم را برای کاهش یا کنترل آلودگی هوا اتخاذ نمایند. بنابراین نیاز به پیش بینی وضعیت کیفی هوا و تخمین های کمی از غلظت آلاینده ها در پی آمدو شد وسایل نقلیه احساس می شود. در حالتی که پیش بینی غلظت آلاینده های هوا به صورت ساعتی در طول روز فراهم شود، می توان با دقت بالاتری نسبت به پیش بینی های روزانه تصمیمات مدیریتی را قبل از شرایط بحرانی لحاظ کرد و در عین حال محدودیت های به نسبت کمتری در تردد وسایل نقلیه ایجاد کرد که در این حالت شبکه حمل و نقل شهری از کارایی بیشتری برخوردار خواهد بود.

دشواری و عدم اطمینان مربوط به برآورد و پیش بینی سهم حجم ترافیک شهری در سطح عمومی کیفیت هوا مهم ترین عاملی است که در صورت تشخیص درست می تواند بسیار کمک کننده باشد. برای لحاظ نمودن آثار متغیر حجم ناوگان مختلف ترافیکی در روند تغییرات غلظت آلاینده ها و آلودگی هوا باید به آثار سایر متغیرهای تأثیرگذار مختلف شامل متغیرهای آب و هوایی، متغیرهای جغرافیایی و ... نیز توجه شود؛ برای نیل به این هدف، نیاز به روش های تحلیل منعطفی احساس می شود که بتواند همه این آثار را در کنار هم و به صورت همه جانبه مورد بررسی قرار دهد.

۲- مروری بر مطالعات گذشته

در مطالعات متعددی به مسئله پیش بینی سطح کیفیت هوای شهری و میزان غلظت آلاینده های انتشار یافته در جو تحت تاثیر ترافیک جاده ای پرداخته شده است [5-8]. یکی از مطالعات انجام شده در این زمینه برای پیش بینی تاثیر ترافیک جاده ای بر سطح غلظت ذرات معلق (PM_{10}) در شهر لندن بود که داده های جمع آوری شده غلظت PM_{10} از ایستگاه های نظارت کیفیت هوا را تحت تاثیر متغیرهای ترافیکی و هواشناسی مورد بررسی قرار داد. در این مطالعه، از روش مدل شبکه عصبی مصنوعی^۱ برای تخمین و برآورد تاثیر ترافیک

وسایل نقلیه است [2]. نیازهای بیش تر به زیرساخت های حمل و نقل جاده ای به عنوان یکی از نتایج توسعه و رشد اقتصادی، خود را در قالب افزایش تعداد وسایل نقلیه سنگین^۴ در سراسر جهان آشکار کرده است. این افزایش به طور ذاتی با افزایش در تراکم، سر و صدا، مصرف انرژی و انتشارات گازی و ذرات معلق و همچنین افزایش در استفاده بیش از حد زیرساخت ها که تمامی آن ها پیامدهای اجتماعی گسترده ای هستند همراه است. اگرچه هزینه های سروصدا و تعمیر و نگهداری زیرساخت ها در درجه اول، آثار محلی و موضعی در همان منطقه هستند، اما افزایش گازهای گلخانه ای ناشی از حمل و نقل یک مشکل جهانی است و آثار قابل ملاحظه ای در تغییرات آب و هوا دارد که باید با آن در سطح جهانی برخورد شود [3]. وسایل نقلیه سنگین که عموماً با موتور دیزلی کار می کنند، به دلیل احتراق ناقص گازوئیل همواره سطحی از آلودگی در گازهای خروجی از موتور این خودروها وجود دارد. آلاینده های خروجی از گازها، ذرات معلق مایع و ذرات معلق جامد است و نکته حائز اهمیت این است که میزان نشر آلاینده های ذرات معلق از خودروهای دیزلی ۱۰ برابر بیش تر از میزان همین آلاینده ها در خودروهای بنزینی است [4]. در چین کنترل انتشارات وسایل نقلیه دیزلی سنگین یک مؤلفه کلیدی مدیریت برنامه ریزی موثر کیفیت هوا است [3]. در کلان شهر تهران نیز اکثر روزهای آلوده سال (بالغ بر ۸۰ درصد روزهای آلوده) تحت تاثیر آلاینده ذرات معلق به خصوص ذرات معلق با قطر کم تر از ۲/۵ میکرون از حد سالم تجاوز کرده و شرایط کیفیت هوا را بحرانی می کند که اغلب به واسطه تردد وسایل نقلیه دیزلی (اعم از مسافری و باری) ایجاد می شود. این امر آلودگی هوای تهران را به ابعاد تهدید آمیزی رسانده است که باید برای کنترل یا مقابله با آن و حفظ سلامت افراد، اقدامات لازم به کار گرفته شود. یکی از مسائل حائز اهمیت در این زمینه مسئله اطلاع رسانی وضعیت آینده کیفیت هوا و میزان آلودگی هوای شهری به مردم و نهادهای مسئول است. این مهم می تواند از طریق پیش بینی های روزانه یا حتی ساعتی وضعیت آلودگی هوا میسر شود و از قرارگیری افراد جامعه در مکان های آلوده و

1 Artificial Neural Network (ANN)

بینی مدل می‌گردد؛ به‌ویژه که این مدل‌ها آثار حرکت ترافیک^۵ به خوبی نشان می‌دهند [7-8].

۳- روش تحقیق

۳-۱- ماشین بردار پشتیبان

شبکه‌های عصبی مصنوعی ابزارهایی مهم در میان مباحث هوش محاسباتی^۶ به حساب می‌آیند. انواع مختلفی از شبکه‌های عصبی مصنوعی معرفی شده‌اند که به‌طور عمده در کاربردهایی همچون طبقه‌بندی، خوشه‌بندی، تشخیص الگو، مدل‌سازی و تقریب توابع (یا رگرسیون)، کنترل، تخمین و بهینه‌سازی مورد استفاده قرار می‌گیرند. ماشین بردار پشتیبان^۷ نوع خاصی از شبکه‌های عصبی هستند که بر خلاف سایر انواع شبکه عصبی (مانند پرسپترون چندلایه MLP و توابع پایه شعاعی RBF) به جای کمینه کردن خطا، اقدام به کمینه کردن ریسک عملیاتی طبقه‌بندی یا مدل‌سازی می‌کند. این ابزار، بسیار قدرتمند است و در زمینه‌های مختلفی چون طبقه‌بندی، خوشه‌بندی و رگرسیون می‌تواند مورد استفاده قرار بگیرد. ماشین بردار پشتیبان (SVM) یکی از قوی‌ترین الگوریتم‌های یادگیری ماشین در حوزه‌های کاربردی متعدد و ناهمگن است که عملکرد بسیار خوبی برای گستره‌ای از برنامه‌های کاربردی مانند پردازش اطلاعات، تشخیص ارقام دست‌نویس، تشخیص چهره، مهندسی مالی، تحلیل پایگاه داده، بیوانفورماتیک و ... دارد. بیش از ده‌ها سال است که پژوهش‌های زیادی در مورد یادگیری ماشین روی ماشین‌های بردار متمرکز شده است. این مفهوم به دلیل پایه نظری قدرتمند خود که بر اساس نظریه یادگیری آماری است از عملکرد عمومی بالایی برخوردار است [9-12].

باید گفت که ماشین‌های بردار پشتیبان مدل‌های یادگیری تحت نظارتی هستند که با الگوریتم‌های یادگیری‌ای که داده را تحلیل و الگوها را شناسایی می‌کنند همکاری می‌نمایند و برای دسته‌بندی و تحلیل رگرسیون استفاده می‌شوند

جاده‌ای استفاده شد. نتایج حاصل از این بررسی نشان داد که مدل‌های شبکه عصبی مصنوعی در پیش‌بینی سهم و مشارکت ترافیک ساعتی جاده‌ای خوب عمل می‌کنند و پیش‌بینی‌ها به خوبی با مشاهدات هماهنگی دارند [5]. در پژوهش دیگری که در گوانگژوی چین انجام شد نیز از شبکه عصبی مصنوعی برای پیش‌بینی غلظت ساعتی آلاینده‌های هوا در نزدیکی یک راه شریانی استفاده گردید؛ و در آن، عوامل تاثیرگذار روی غلظت آلاینده‌ها در چهار دسته کلی شامل متغیرهای ترافیکی، غلظت پس‌زمینه، هواشناسی و جغرافیایی تقسیم‌بندی و تعریف شد و میانگین ساعتی آلودگی این عوامل و غلظت مونوکسید کربن، دی‌اکسید نیتروژن، ذرات معلق و ازن در سه محل انتخابی در نزدیکی راه شریانی با استفاده از تجهیزات نظارت اتوماتیک مربوط به وسایل نقلیه اندازه‌گیری شده است. الگوریتم مورد استفاده برای یادگیری شبکه عصبی در این پژوهش روش انتشار برگشتی^۱ بود که با استفاده از داده‌های جمع‌آوری شده آموزش داده شدند و تحت اعتبارسنجی و آزمایش قرار گرفتند و نتایج نشان داد که مدل‌های شبکه عصبی مصنوعی، می‌توانند به‌عنوان یک رویکرد موثر برای پیش‌بینی غلظت آلاینده‌های هوا در نزدیکی شریان‌های شهری در نظر گرفته شوند [6]. در مقالات دیگری که مطالعات آن در شهر دهلی‌نو کشور هند صورت گرفته، برآورد و پیش‌بینی انتشارات گازی وسایل نقلیه^۲ با استفاده از شبکه عصبی مصنوعی با روش پرسپترون چندلایه^۳ برای دو منطقه کنترل کیفیت هوا^۴، یکی تقاطع ترافیکی و دیگری جاده شریانی انجام شده است. داده‌های هواشناسی به همراه داده‌های حجم ترافیک ساعتی وسایل نقلیه و نیز غلظت‌های پس‌زمینه برای پیش‌بینی ساعتی غلظت آلاینده‌های CO و NO₂. به‌کار گرفته شدند. در این مطالعه مشخص شد که دقت پیش‌بینی مدل‌های ANN با در نظرگیری توأمان داده‌های هواشناسی و ترافیکی به‌عنوان متغیرهای ورودی افزایش می‌یابد و هرچقدر تعداد متغیرهای ورودی که دینامیک پراکندگی آلاینده را نشان می‌دهند زیاده‌تر شود، باعث بهبود عملکرد پیش-

1 Back-Propagation (BP)

2 Vehicular Exhaust Emissions (VEEs)

3 Multilayer Perceptron (MLP)

4 Air Quality Control Region (AQCR)

5 Traffic Wake Effects

6 Computational Intelligence

7 Support Vector Machines (SVM)

که به ترتیب، دسته‌بندی بردار پشتیبان¹ و رگرسیون بردار پشتیبان² نام دارند. با داشتن مجموعه‌ای از نمونه‌های آموزشی که مشخص شده است هر کدام به چه دسته‌ای از دو دسته موجود متعلق هستند، یک الگوریتم آموزشی SVM مدلی را می‌سازد که نمونه‌های جدید را به دسته اول یا دوم تخصیص بدهد. نسخه‌ای از SVM با نام رگرسیون بردار پشتیبان برای حل رگرسیون توسط ولادیمیر واپنیک، هریس دروکر، کریستوفر برگرز، لیندا کافمن و الکساندر سمولو ارائه شد. مدل تولید شده توسط این متد، تنها وابسته به زیرمجموعه‌ای از داده‌های آموزشی است؛ زیرا تابع هزینه برای ساخت مدل، هر داده آموزشی‌ای که (با توجه به یک بازه - ϵ) نزدیک به مدل پیش‌بینی باشد را در نظر نمی‌گیرد [10, 12].

اگر مدل‌ها کنترل نشوند، تمایل دارند که بیشینه برازش (و کم‌ترین خطا) را روی داده‌های آموزش³ ایجاد کنند؛ اما این همیشه مطلوب نیست و مدل را در معرض فرابرازش⁴ قرار می‌دهد. فرابرازش به وضعیتی گفته می‌شود که مدل با بزرگ شدن بیش از حد، به طور مصنوعی و تعمیم‌ناپذیر، خود را به داده آموزش برازش می‌دهد و با وجود کوچک شدن بسیار زیاد خطای آموزش، مدل روی داده آزمون⁵ ضعیف عمل می‌کند. برای جلوگیری از این پدیده، باید بزرگی (ظرفیت) مدل به دقت انتخاب یا کنترل شود. در این پژوهش از داده‌های آموزش برای برازش مدل استفاده شده و برای انجام آزمون و عدم ایجاد فرابرازش، از داده‌های آزمون که قبل از مدل‌سازی کنار گذاشته شده‌اند، استفاده شده است.

۳-۲- تعریف و تعیین متغیرهای مدل

با توجه به موضوع مورد بررسی در این پژوهش و مطالب ذکر شده در پیشینه پژوهش، هدف نهایی، پیش‌بینی ساعتی تغییرات در میزان غلظت آلاینده ذرات معلق در اثر حجم تردد ناوگان جریان ترافیک و تعدد ترکیب ناوگان است؛ از آنجا که مقیاس بررسی‌های این پژوهش از نوع ساعتی است، متغیر وابسته مدل‌سازی همان مقدار غلظت ساعتی آلاینده ذرات معلق خواهد بود که به‌عنوان متغیر خروجی مدل تعریف می‌شود. بخشی از متغیرهای مستقل مدل‌سازی که در دسته متغیرهای ورودی اصلی مدل قرار می‌گیرند، شامل دو پارامتر هستند که یکی پارامتر حجم تردد ساعتی ناوگان ترافیک شهری و دیگری پارامتر غلظت پس‌زمینه ساعتی است که متغیر خروجی مدل، حاصل از این دو متغیر ورودی اصلی خواهد بود. به‌عبارت دیگر، به دلیل ساعتی بودن مقیاس این پژوهش، می‌توان گفت که غلظت ساعت حاضر از تجمیع غلظت موجود در یک

ایده SVM مبتنی بر مفاهیم حاشیه و توابع هسته است. هدف آن این است که ابرصفحه‌ای را پیدا کند که مثال‌هایی از نمونه را به بهترین شکل به چندین دسته تقسیم کند. در این روش، داده‌های دو دسته توسط ابرصفحه‌ای به طور کامل از هم جدا می‌شوند و فاصله‌ی نزدیک‌ترین نقاط دو دسته تا ابرصفحه بیشینه می‌شود. این رویکرد علاوه بر اینکه منجر به پاسخ یکتا می‌شود، کارایی بهتر مدل را هم در پی دارد. یافتن این ابرصفحه منجر به حل کردن یک مساله بهینه‌سازی محدب طبق رابطه (۱) شده که از نظر محاسباتی به‌صرفه است.

ایده SVM مبتنی بر مفاهیم حاشیه و توابع هسته است. هدف آن این است که ابرصفحه‌ای را پیدا کند که مثال‌هایی از نمونه را به بهترین شکل به چندین دسته تقسیم کند. در این روش، داده‌های دو دسته توسط ابرصفحه‌ای به طور کامل از هم جدا می‌شوند و فاصله‌ی نزدیک‌ترین نقاط دو دسته تا ابرصفحه بیشینه می‌شود. این رویکرد علاوه بر اینکه منجر به پاسخ یکتا می‌شود، کارایی بهتر مدل را هم در پی دارد. یافتن این ابرصفحه منجر به حل کردن یک مساله بهینه‌سازی محدب طبق رابطه (۱) شده که از نظر محاسباتی به‌صرفه است.

$$\min_{W, \xi} \frac{1}{2} \|W\|^2 + C \sum_{i=1}^n \xi_i \quad (1)$$

با شرایط:

$$y_i [w' \phi(x_i) + b] \geq 1 - \xi_i \\ \xi_i \geq 0, \quad i \in I = \{1, 2, \dots, n\}$$

که در آن x_i مثال بردار ویژگی (متغیرهای مستقل) و متعلق به فضای R^p است (p تعداد ویژگی‌های هر نمونه)، y_i متغیر وابسته، w بردار ضرایب، b مقدار ثابت (بایاس)، ϕ تابع کرنل (هسته)، و n تعداد مثال‌های آموزش است. دو پارامتر

3 Training Data
4 Overfitting
5 Testing Data

1 Support Vector Clustering (SVC)
2 Support Vector Regression (SVR)

مسافری هستند که باید این دو طبقه کاربری نیز از هم جدا شوند. دسته‌بندی دیگر، تفکیک وسایل نقلیه دیزلی مسافری از باری است. در نهایت ۳ نوع دسته‌بندی برای حجم‌های ساعتی تردد ناوگان مختلف با نام‌های وسایل نقلیه بنزینی شخصی^۱، وسایل نقلیه دیزلی مسافری^۲ و وسایل نقلیه دیزلی سنگین^۳ مدنظر قرار می‌گیرد.

پارامترهای هواشناسی مانند دما، رطوبت هوا و سرعت و جهت وزش باد، نزولات جوی (شامل بارش باران و برف) نیز می‌توانند روی میزان غلظت آلاینده‌های هوا به‌طور ثانویه اثرگذار باشند و میزان غلظت آلاینده‌های تولید شده از منابع انتشار را دست‌خوش تغییر کنند که باید اثر آن‌ها برای مدل‌سازی غلظت آلاینده‌ها در کنار منابع انتشار دیده شود. بنابراین پارامترهای هواشناسی مدنظر مدل‌سازی از سایت www.wunderground.com برای محدوده زمانی و مکانی پژوهش حاضر دریافت شد. در جدول (۱) لیستی از مشخصات متغیرها در مدل‌سازی غلظت آلاینده ذرات معلق و اجزای تشکیل دهنده هر کدام ارائه شده است.

شکل ۱. محدوده مکانی کلی پژوهش (شهر تهران) نشان‌گرها: ایستگاه‌های کنترل کیفیت هوا نواحی آبی‌رنگ: مناطق طرح زوج‌وفرد نواحی قرمز رنگ: مناطق طرح ترافیک

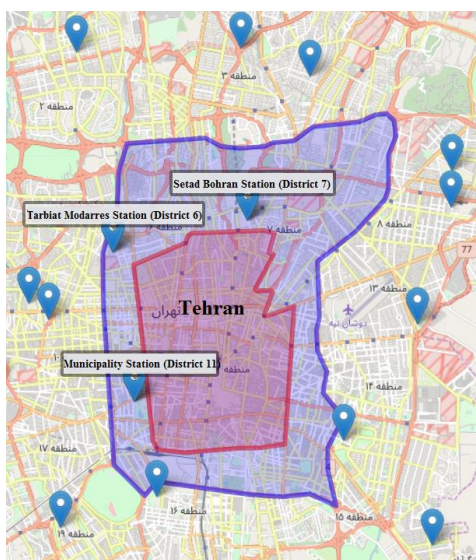


Fig. 1. Overall research area (Teahran City) Markers: Air Quality Control Stations Blue areas: Even and Odd Plan areas Red areas: Traffic Plan areas

- 1 Personal Gasoline Vehicles (PGV)
- 2 Passenger Diesel Vehicles (PDV)
- 3 Heavy Diesel Vehicles (HDV)

ساعت قبل و غلظت ایجاد شده در اثر تردد ترافیک شهری در ساعت فعلی حاصل می‌شود. بنابراین عنوان غلظت پس‌زمینه ساعتی در این پژوهش برای غلظت ساعت قبل در نظر گرفته شده است که یکی از متغیرهای ورودی اصلی مطالعه است. متغیرهای مستقل دیگری نیز تحت عنوان متغیرهای تعدیل‌گر برای انجام مدل‌سازی در نظر گرفته می‌شوند که بخش دوم متغیرهای مستقل ورودی مدل را تشکیل می‌دهند و آن‌ها همان پارامترهای هواشناسی هستند؛ این پارامترها منبع اصلی تولید آلاینده‌ها نیستند و تنها اثر ثانویه‌ای روی میزان غلظت آلاینده‌های تولید شده هوا دارند و در دسته متغیرهای ورودی فرعی مدل جای می‌گیرند.

در این پژوهش اطلاعات غلظت ساعتی آلاینده ذرات معلق با قطر کم‌تر از دو و نیم میکرون ($PM_{2.5}$) که توسط ایستگاه‌های کنترل کیفیت هوا در سراسر شهر تهران اندازه‌گیری می‌شوند، برای محدوده مکانی و زمانی پژوهش (که در ادامه اشاره خواهد شد) از شرکت کنترل کیفیت هوای تهران دریافت شده است. غلظت پس‌زمینه ساعتی نیز با ایجاد تأخیر یک ساعته در غلظت‌ها، غلظت ساعت قبل آلاینده به‌عنوان غلظت پس‌زمینه ساعتی آن حاصل شد که توضیحات لازم در رابطه با آن در انتهای پاراگراف اول بخش ۳-۲ ارائه شده است.

برای دستیابی به حجم تردد ساعتی خودروها در محدوده مکانی و زمانی پژوهش نیز، حجم شبانه‌روزی وسایل نقلیه برای محدوده مکانی و زمانی پژوهش که توسط دوربین‌های شرکت کنترل ترافیک در شهر تهران وظیفه ثبت تردد خودروها را دارند، به تفکیک نوع ناوگان عبوری، دریافت شد. با توجه به تعدد ناوگان، برای کاهش ابعاد و نیز پیچیدگی مدل، یک ساده‌سازی روی دسته‌بندی ناوگان انجام می‌گیرد. برای این منظور باید به نوع سوخت مصرفی در کنار کاربری توجه ویژه داشت؛ چرا که سطح انتشار آلاینده‌های مدنظر از خودروهای بنزینی به‌طور قابل‌توجهی پایین‌تر از خودروهای دیزلی بوده و یکی از دسته‌بندی‌ها، تفکیک وسایل نقلیه بنزینی از دیزلی است. به‌علاوه، وسایل نقلیه دیزلی دارای دو طبقه کاربری متفاوت باری و

جدول ۱. متغیرهای به کار رفته در مدل سازی

Parameters	Symbol	Definition	Details	Unit
background hourly concentration of pollutant	PM _{2.5} -BG	Concentration 1 Past Hour of PM _{2.5}	-	μg/m ³
	V _{PGV}	Hourly Traffic Volume of PGV (Personal Gasoline Vehicles)	Personal Car - Taxi - Pickup - Ambulance	veh/hr
hourly traffic volume of Fleet	V _{PDV}	Hourly Traffic Volume of PDV (Passenger Diesel Vehicles)	Mini Bus - Bus	veh/hr
	V _{HDV}	Hourly Traffic Volume of HDV (Heavy Diesel Vehicles)	Trucklet - Truck - Trailer - Semi Trailer	veh/hr
	T	Air Temperature	all weather conditions included atmospheric	°F
Meteorological variable (per hour)	RH	Relative Humidity Percentage	precipitation, wind, fog, etc; except mist, dust, blowing sand, etc;	%
	WS	Wind Speed		Km/hr
	Sin(WD)	Sinus of Wind Direction	the effect of East-West of wind blowing	Rad
	Cos(WD)	Cosine of Wind Direction	the effect of North-South of wind blowing	Rad
	RF	Rainfall	severity of rainfall (Ordinal)	-

Table 1. Defined parameters for modeling

شود که ایستگاه کنترل کیفیت شهرداری منطقه ۱۱ برای مطالعه انتخاب می شود. که در خیابان قزوین واقع شده است. در شکل (۲) موقعیت دوربین تردد شمار با شماره ۱۳۳۵۷ و ایستگاه کنترل کیفیت هوای منطقه ۱۱ شهرداری تهران نشان داده شده است.

شکل ۲. محدوده مکانی نهایی پژوهش (خیابان قزوین، شهر تهران)

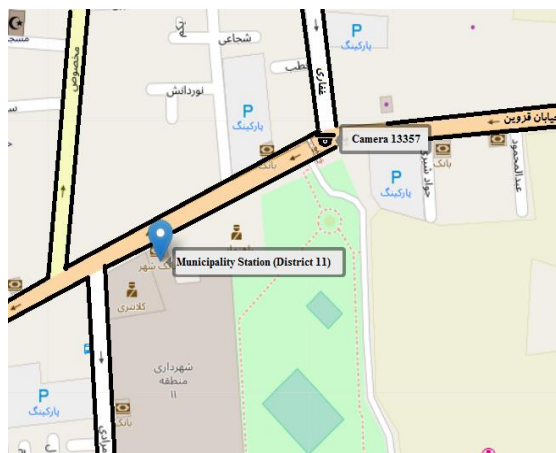


Fig. 2. Final research area (Qazvin Street, Teahran City)

۳-۳- تعیین محدوده مورد مطالعه

محدوده مورد مطالعه شامل محدوده مکانی و زمانی است که باید به آنها پرداخته شود. در وهله اول، محدوده مکانی کلی درون مناطق طرح زوج و فرد مناطق شهرداری تهران مطابق شکل (۱) انتخاب می شود. چرا که این دو منطقه از جمله مناطق شلوغ و پر رفت و آمد شهر تهران هستند که البته با اعمال اقدامات و سیاست های محدودیت در تردد وسایل نقلیه تا حدودی از میزان ازدحام ترافیکی و در نتیجه میزان آلودگی هوای آنها کاسته شده است.

با توجه به محدودیت در اخذ داده های هواشناسی و فقدان آنها برای ایستگاه های هواشناسی داخل مناطق طرح زوج و فرد تهران و تنها دسترسی به داده های ایستگاه هواشناسی فرودگاه مهرآباد، باید از بین ایستگاه های کنترل کیفیت هوا ایستگاهی که فاصله نسبتاً نزدیکی تا ایستگاه هواشناسی مذکور دارد برای بررسی های این پژوهش انتخاب

۴۲۲ داده ساعتی مجزا است که نشان دهنده ظرفیت خوبی برای مدل است.

باتوجه نتایج بدست آمده روشن است که مقادیر پیش-بینی شده غلظت ذرات معلق توسط مدل ساخته شده با ماشین بردار پشتیبان در مرحله آزمون، دارای هماهنگی و نزدیکی قابل قبول به مقادیر اندازه‌گیری شده غلظت‌ها است که بخشی از نتایج در جدول (۲) ارائه شده است.

۵- نتیجه‌گیری

پیش‌بینی ساعتی آلودگی هوا در شهرها نسبت به پیش‌بینی‌های روزانه به این دلیل که امکان اطلاع‌رسانی عمومی دقیق‌تر به شهروندان و اتخاذ سیاست‌های راهبردی در حوزه کنترل جریان ترافیک ناوگان وسایل نقلیه از سوی نهادهای مسئول در زمینه کنترل آلودگی هوا را فراهم می‌آورد از اهمیت بالاتری برخوردار است؛ به‌ویژه اینکه در حالت پیش‌بینی ساعتی، اعمال محدودیت‌های ترافیکی جهت کنترل یا کاهش آلودگی هوا موضعی بوده و محدودیت‌های کم‌تری به جریان ترافیک عبوری وارد می‌کند که در نتیجه آن، کارایی شبکه حمل‌ونقل شهری از سطح بالاتری برخوردار خواهد بود. برای دستیابی به این نوع پیش‌بینی باید اطلاعات ساعتی حجم ناوگان ترافیک عبوری که منبع اصلی انتشار آلاینده‌ها هستند، در کنار اطلاعات ساعتی شرایط آب‌وهوایی که غلظت آلاینده‌ها را دستخوش تغییرات می‌کند به‌صورت یک‌جا در نظر گرفته شوند.

امروزه پیش‌بینی‌هایی برای وضعیت ساعتی آب‌وهوایی مناطق سرتاسر جهان انجام می‌شود که وضعیت حداقل ۲۴ ساعت آینده را تخمین می‌زنند و تا حد قابل قبولی مورد تایید هستند. از طرفی با توجه به تقاضای سفر معابر مختلف در ایام سال و آماربرداریه‌ها نیز، قابلیت پیش‌بینی و تخمین ترافیک ساعتی عبوری از معابر فراهم آمده است.

داده‌ها متعلق به مدت زمان ۴۷۰ روز و محدوده زمانی نیز روزهای مابین دو تاریخ ۱۳۹۵/۰۱/۰۱ الی ۱۳۹۶/۰۴/۱۵ است. در این بازه زمانی، برخی از روزها و ساعات به‌دلیل فقدان داده‌ها اعم از پارامتر ترافیکی، غلظتی و هواشناسی، حذف و محدوده زمانی نهایی محدود به ساعاتی شد که هر سه دسته داده مذکور با یکدیگر هم‌پوشانی داشته باشند.

۴- مدل‌سازی و تحلیل نتایج

در این پژوهش برای توسعه مدل SVM از تابع پایه شعاعی RBF برای تابع هسته استفاده شد. برای بهبود عملکرد^۱ مدل SVM نیاز به تنظیم^۲ آن با توسعه شبکه‌ای از پارامترها (شامل دو پارامتر خطای مدل γ و C به همراه پارامتر تابع هسته γ) است. برای جستجوی پارامترها از روش اجرای موازی^۴ استفاده شده که نسبت به روش اجرای متوالی از سرعت و کارایی بالاتری برخوردار است. در کاربرد روش اجرای موازی، جستجوی پارامترهای SVM شبکه از طریق اعتبارسنجی متقابل^۵ 10^5 بار تکرار استفاده شده است. لازم به ذکر است که برای بهینه‌سازی پارامترها ابتدا مقادیر آن‌ها در بازه بزرگی انتخاب و پس از یافتن تقریبی محدوده جواب‌ها، بازه کوچک‌تری برای تدقیق مقادیر اتخاذ شد.

پس از یافتن مقادیر مناسب پارامترها، برای توسعه مدل SVM در مرحله اول از داده‌های آموزش برای انجام فرآیند آموزش استفاده شد و در گام بعد برای درستی آزمایشی مدل و برآورد دقت پیش‌بینی آن، داده‌های آزمون به کار گرفته شدند و نتایج آن در شکل (۳) نشان داده شده است که نمایان‌گر دو نمودار است؛ یکی نمودار دنباله‌ی مقادیر مشاهده شده غلظت ساعتی آلاینده $PM_{2.5}$ و دیگری نمودار دنباله‌ای مقادیر پیش‌بینی شده غلظت ساعتی همین آلاینده توسط مدل ساخته شده با ماشین بردار پشتیبان است. لازم به ذکر است که مدل در مرحله آموزش و آزمون شامل ۱۶۹۰ و

- 1 Performance
- 2 Tunning
- 3 Expand Grid
- 4 Parallel Execution
- 5 Cross Validation

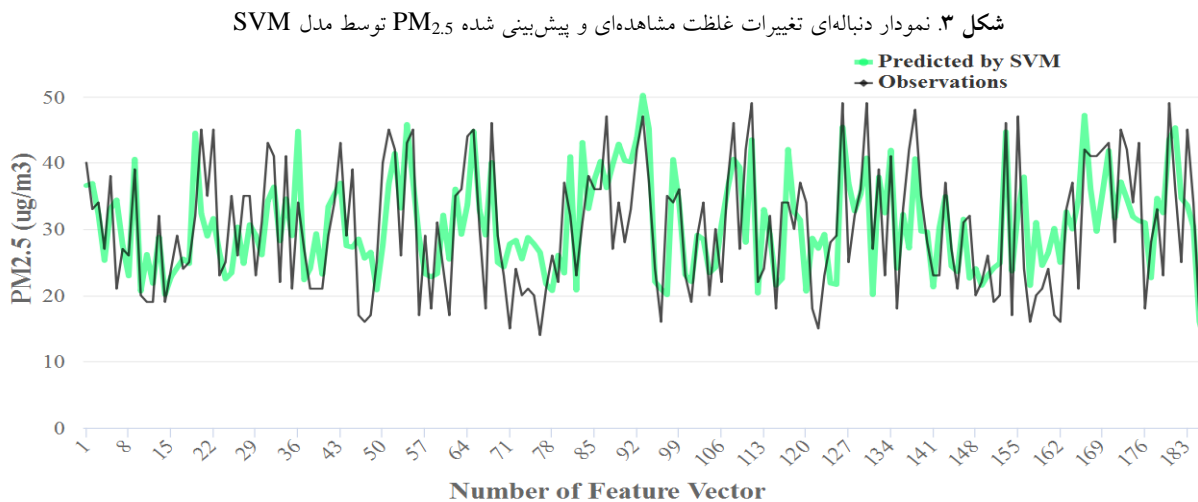


Fig. 3. Sequence plot of predicted vs observed $PM_{2.5}$ by SVM model

جدول ۲. برخی از نتایج مقادیر غلظت مشاهده شده و پیش‌بینی شده $PM_{2.5}$ توسط مدل SVM

Predicted Value	Observed Value	Predicted Value	Observed Value	Predicted Value	Observed Value
25.5	24	40.5	39	36.6	40
24.8	25	20.6	20	36.9	33
44.5	32	26.2	19	32.2	34
32.3	45	21.8	19	25.3	27
28.9	35	28.8	32	33.1	38
31.6	45	20.1	19	34.4	21
26.4	23	22.8	24	27.6	27
22.5	25	24.2	29	22.9	26

Table 2. Some of result of predicted vs observed $PM_{2.5}$ by SVM model

تا اثربخشی یک سناریوی مدیریت کیفیت هوا در رابطه با فناوری ناوگان ترافیک، ترکیب ناوگان ترافیکی یا حجم آن را مورد آزمایش قرار دهد.

۶- پیشنهادها

- برای انجام پژوهش در آینده، موضوعات زیر پیشنهاد می‌شوند:
 - پیش‌بینی ساعتی و یا روزانه غلظت آلاینده‌ها با استفاده از مدل‌های سری‌زمانی مانند ARMA، ARIMA و ...،
 - شبیه‌سازی آثار معاینه فنی احجام ترافیکی در شبکه معابر روی میزان انتشار آلاینده‌ها با استفاده از نرم‌افزارهای شبیه‌سازی مانند AIMSUN، CORFLO و VISSIM که میزان انتشارات ترافیک شبیه‌سازی شده را در خروجی خود دارند،
 - مدل‌سازی آثار تردد ناوگان مختلف روی غلظت آلاینده‌های معیار هوا.

با در دست داشتن این دو دسته اطلاعات و روش تحلیل و پیش‌بینی مناسب می‌توان به تخمین‌های مناسبی از وضعیت کیفیت هوا و غلظت آلاینده‌ها در ساعات آینده روز رسید. بنابراین نیاز به روش‌های تحلیل منعطفی نظیر روش-های هوش مصنوعی احساس می‌شود تا بتواند همه این آثار را در کنار هم و به صورت همه‌جانبه مورد بررسی قرار دهد. یکی از این روش‌ها، ماشین بردار پشتیبان است که به اختصار SVM نامیده می‌شود. نتایج حاصل از این تحقیق نشان داد که مدل‌های SVM در پیش‌بینی ساعتی سهم و مشارکت ترافیک شهری در انتشار ذرات معلق به شدت خوب عمل می‌کنند و پیش‌بینی غلظت‌ها به خوبی با مشاهدات هماهنگی دارند. نتایج تحلیل اهمیت متغیر برای مدل‌های SVM این فرصت را فراهم می‌کند تا به عنوان ابزار مدیریت کیفیت هوا به کار روند که در آن حساسیت مدل‌ها نسبت به تغییرات در انتشار می‌تواند مورد استفاده قرار گیرد

- [6] Cai M., Yin Y. & Xie M. 2009 Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach. *Transportation Research Part D*, 14, 32-41.
- [7] Nagendra S. M. S. & Khare M. 2004 Artificial neural network based line source models for vehicular exhaust emission predictions of an urban roadway. *Transportation Research Part D*, 9(3), 199-208.
- [8] Nagendra S. M. S. & Khare M. 2006 Artificial neural network approach for modelling nitrogen dioxide dispersion from vehicular exhaust emissions. *Ecological Modelling*, 190, 99-115.
- [9] Vapnik V. N. 1995 *The nature of statistical learning theory*, in: *Information Science and Statistics*, Second Edition, Springer-verlag, New York.
- [10] Cortes C. & Vapnik V. N. 1995 Support vector networks. *Machine Learning*, 20, 273-297.
- [11] Burges C. 1998 A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2, 121-167.
- [12] Schlkopf C. & Smola A. 2001 *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, MA, USA, MIT Press Cambridge.
- [13] Djemai S., Brahmī B. & Bibi M. O. 2016 A primal-dual method for SVM training. *Neurocomputing*, 211, 34-40.

References

۷- مراجع

- [1] Berari K. , Nazapoor A. & Solat G. 2011 Modeling and estimating pollutants due to urban traffic to increase efficiency and reduce energy dissipation. *The First International Conference on Modern Approaches to Energy Conservation, Conferences in Iran, Tehran*. (In Persian)
- [2] Poorzahedi H. , Zarei H. & Kermanshah M. 2002 Evaluation of urban traffic management policy in critical situation of air pollution in Shiraz. *Scientific Journal of Urban Management*, 3(11 & 12), 54-61. (In Persian)
- [3] Paulikakos L. D., Heutschi K. & Soltic P. 2013 Heavy duty vehicles: Impact on the environment and the path to green operation. *Journal of Environmental Science & Policy*, 33, 154-161.
- [4] Naderi M. D. & Rahro Mostaghim M. 2013 The study of the effects of gasoline and diesel quality on pollutant emissions from passenger cars and heavy vehicles, Air Quality Control Company of Tehran Municipality, Tehran. (In Persian)
- [5] Suleiman A., Tight M. R. & Quinn A. D. 2016 Assessment and prediction of the impact of road transport on ambient concentrations of particulate matter PM10. *Transportation Research Part D*, 49, 301-312.

Short-term Prediction of Particulate Matter Related to Traffic Flow Using Support Vector Machine (Case Study: Tehran City)

M. Kashipazan¹, B. Shirgir^{2*}

1- M.Sc. Student, Civil Engineering Department, Faculty of Engineering, Kharazmi University, Tehran, Iran

2- Assist. Prof., Civil Engineering Department, Faculty of Engineering, Kharazmi University, Tehran, Iran

* Shirgir@khu.ac.ir

Abstract

In recent decades, increasing population density and economic and industrial activities in metropolitan cities has increased traffic volumes and, consequently, increased levels of air pollution. The major source of air pollution in major developing cities is the massive transport of vehicles that use more than standard fuel and energy, and heavy traffic in the streets of these cities is often rooted in problems such as there is a lack of traffic management and traffic culture. One of the important issues in cities and metropolises that face pollution problems and harmful effects is the issue of informing about the future status of air quality and the amount of urban air pollution to the people. This can be achieved through daily or even hourly forecasts of air pollution and preventing people from being exposed to contaminated areas and their irreversible consequences. Therefore, the need to predict the quality of the air and the quantitative estimates of the concentration of pollutants in the aftermath of the equipment makes it felt that in this study, the problem of the prediction hourly concentration of particulate matter (PM_{2.5}) in the district 11 municipalities of Tehran have exceeded 80% of the contaminated days under the influence of this pollutant. The difficulty and uncertainty associated with estimating and predicting the share of road traffic volume at the general level of air quality is the most important factor that can, if properly diagnosed, be very helpful. In order to take into account the effects of varying the volume of different traffic fleets in the process of changes in the concentration of pollutants and air pollution, it is necessary to pay attention to the effects of other influential variables including hydrological variables, geographical variables, etc. To achieve this, The methods of analytic analysis seem to be able to examine all of these effects together and in an omnipresent manner. The method used to predict this study is one of the methods for analyzing neural networks called Support Vector Machine (SVM). Artificial neural networks are important tools in the field of computational intelligence. Different types of artificial neural networks have been introduced, mainly in applications such as classification, clustering, pattern recognition, modeling and approximation of functions (or regression), control, estimation and optimization of the case Are used. Support Vector Machines (SVM) are a special type of neural network that, unlike other types of neural networks (such as multi-layer perceptron MLP and radial base functions of the RBF), instead of minimizing the error, minimize the operational risk of classification or modeling. Slowly This tool is very powerful and can be used in various fields such as classification, clustering and regression. The results of this study showed that SVM models work well in predicting the contribution and time share of urban traffic in propagation of particulate matter, and predictions are well-coordinated with observations. It provides the opportunity to be used as an air quality management tool. Variable significance analysis results for SVM models provide this opportunity to be used as a tool for air quality management, in which the sensitivity of models to variations in emissions can be used to evaluate the effectiveness of a The air quality management scenario will test traffic fleet technology, combine the traffic fleet or its volume.

Keywords: Air Pollution Prediction, Support Vector Machines, Traffic Flow, Particulate Matter, Air Quality Management