

توسعه یک چارچوب ترکیبی مبتنی بر یادگیری ماشین برای پایش و پیش‌بینی بوی نامطبوع در کانال‌های آب سطحی شهری: مطالعه موردی کانال غیاثوند تهران

عمار صفایی^{۱*}، محمد خانجانی^۲، زهرا کریمی شوکی^۳، علیرضا هادیان^۴، امیرحسین حدادی^۵، هدیه رحمانی^۶

- ۱- استادیار، دانشکده مهندسی عمران، دانشگاه صنعتی شریف
- ۲- دانشجوی کارشناسی ارشد مدیریت منابع آب، دانشکده مهندسی عمران، دانشگاه صنعتی شریف
- ۳- دانشجوی دکتری مهندسی محیط‌زیست، دانشکده مهندسی عمران، دانشگاه صنعتی شریف
- ۴- دانشجوی کارشناسی ارشد مهندسی محیط‌زیست، دانشکده مهندسی عمران، دانشگاه صنعتی شریف
- ۵- دانشجوی کارشناسی، دانشکده مهندسی عمران، دانشگاه صنعتی شریف

پست الکترونیکی نویسندگان:

- ۱- asafaie.sharif.edu
- ۲- mohammad.khanjani@sharif.edu
- ۳- z.shavaki@sharif.edu
- ۴- alireza.hadian@sharif.edu
- ۵- amir.haddadi@sharif.edu
- ۶- hediye.rahmani@sharif.edu

چکیده:

این مطالعه چارچوبی عملیاتی برای پایش، کمی‌سازی و پیش‌بینی شدت بوی نامطبوع در کانال‌های جمع‌آوری آب سطحی ارائه می‌دهد که در قالب یک پروژه پایلوت در کانال غیاثوند در تهران اجرا شد. شاخص‌های کیفیت آب، متغیرهای هواشناسی و شدت بوی نامطبوع در داده‌برداری میدانی هفتگی جمع‌آوری شدند. روابط میان متغیرهای کیفی آب و هواشناسی با شدت بو ابتدا با استفاده از تحلیل‌های همبستگی بررسی و سپس با الگوریتم رگرسیون جنگل تصادفی برای پیش‌بینی کمی شدت بو مدل‌سازی شد. نتایج نشان می‌دهد دمای آب، هدایت الکتریکی و دمای هوا همبستگی مثبت و اکسیژن محلول همبستگی منفی با شدت بو داشته و مدل توسعه داده شده عملکرد مطلوبی در پیش‌بینی شدت بو دارد ($R^2 = 0.83$ برای داده‌های آموزش و آزمون). این پژوهش نشان می‌دهد ترکیب داده‌های میدانی و روش‌های یادگیری ماشین می‌تواند ابزاری موثر برای پیش‌بینی شدت بو و توسعه سامانه‌های پایش هوشمند و مدیریت کیفیت آب در بسترهای شهری فراهم آورد.

واژگان کلیدی:

بوی نامطبوع، کانال آب سطحی شهری، کیفیت آب، الگوریتم جنگل تصادفی.

ایمیل: asafaie@sharif.edu (نویسنده مسئول مقاله)

^{۱*} عمار صفایی، استادیار دانشکده مهندسی عمران - دانشگاه صنعتی

شریف.

Development of a Hybrid Machine Learning-Based Framework for Monitoring and Predicting Odor Pollution in Urban Surface Water Canals: A Case Study of the Ghiyasvand Canal, Tehran

Ammar Safaie^۱, Mohammad Khanjani^۲, Zahra Karimi Shavaki^۳, Alireza Hadian^۴, Amirhossein Haddadi^۵,
Hediye Rahmani^۵

۱- Assistant Professor, Department of Civil Engineering, Sharif University of Technology.

۲- M.Sc. Student of Water Resources Management, Department of Civil Engineering, Sharif University of Technology.

۳- Ph.D. Student of Environmental Engineering, Department of Civil Engineering, Sharif University of Technology.

۴- M.Sc. Student of Environmental Engineering, Department of Civil Engineering, Sharif University of Technology.

۵- B.Sc. Student, Department of Civil Engineering, Sharif University of Technology.

Abstract:

Unpleasant odors emitted from urban surface-runoff canals pose a significant environmental and public health challenge in large cities such as Tehran. These odors, primarily caused by anaerobic decomposition of organic matter and the inflow of untreated wastewater, degrade water quality, generate public dissatisfaction, and reduce the environmental health and livability of urban areas. Despite their widespread impact, the spatial distribution and intensity of odor emissions in such canals are rarely monitored systematically. To fill this gap, this study introduces a practical framework for monitoring, quantifying, and predicting odor intensity in urban surface water canals. The proposed framework was implemented and evaluated in a pilot project along the Ghiyasvand Canal in Tehran. Weekly field sampling was carried out at ten critical locations over a ten-week period, during which key water-quality parameters (including pH, electrical conductivity (EC), total dissolved solids (TDS), and dissolved oxygen (DO), and water temperature) and meteorological variables (such as air temperature, wind speed, and relative humidity) were collected. In addition, odor intensity was measured with a portable Odor meter. Pearson, Spearman, and Kendall correlation analyses, along with a random forest regression model, were employed to examine and predict the relationships between physicochemical and atmospheric variables and the odor intensity. Correlation analyses indicated that water temperature, electrical conductivity, and air temperature were positively correlated with odor intensity, whereas DO showed a negative correlation, indicating its critical role in odor suppression. The developed model performed well in predicting odor intensity, achieving an accuracy of 83% for both training and testing data. This study demonstrates the potential of integrated field monitoring and machine learning approaches to support practical odor management in urban water systems, leading to improved environmental quality and public well-being. While the framework was applied to a specific case in Tehran, the results and approach are broadly applicable to similar urban settings facing odor-related challenges.

Keywords: Unpleasant odor, Urban Surface Water Canals, Water Quality, Random Forest Algorithm.

سطحی، علاوه بر تخریب زیستگاه‌های آبی، تأثیرات منفی مستقیمی بر کیفیت زندگی ساکنین این مناطق دارد و می‌تواند منجر به نارضایتی شهروندان گردد.

رودخانه‌های زرجوب و گوهررود در شهر رشت [۶]، رودخانه بشار در یاسوج [۷]، رودخانه چنار راهدار در شیراز [۸] و رودخانه کشف‌رود در مشهد [۹] از جمله منابع آبی سطحی هستند که به دلیل ورود مستقیم رواناب‌های سطحی، پساب‌های شهری و فاضلاب‌های صنعتی، با افزایش بار آلودگی و در نتیجه، تغییر رنگ، کاهش کیفیت و انتشار بوی نامطبوع مواجه شده‌اند که نشان‌دهنده اختلال در اکوسیستم آبی می‌باشد. چالش‌های ناشی از کیفیت پایین آب و انتشار بوهای نامطبوع از آب‌های سطحی، محدود به ایران نیست و در سطح جهانی نیز مشاهده می‌شود. برای مثال، کانال هوم وود در ایالات متحده، رودخانه تیمز در انگلستان و دریاچه بیوا در ژاپن نمونه‌هایی هستند که به دلیل ورود مستمر آلاینده‌ها با مشکلات جدی مواجه شده‌اند [۱۰]. همچنین، آب‌های به شدت آلوده در مناطق مختلف سیاه و بدبو می‌شوند. رودخانه چونگ‌گی^۱ در کره جنوبی [۱۱]، رودخانه امشیر^۲ در آلمان [۱۲]، رودخانه سن در فرانسه [۱۳] و رودخانه دانوب در اتریش [۱۴]، کانال گوانوس در ایالات متحده آمریکا [۱۵] و رودخانه‌های آلوده شنیلنگ در چین [۱۶] نمونه‌هایی از این بحران هستند که با چالش‌های جدی‌ای مواجه بوده و هستند. اگرچه آلودگی بوی نامطبوع آب در آیین‌نامه‌های محیط‌زیستی اغلب کمتر مورد توجه قرار می‌گیرد، اما این مسئله علاوه بر اینکه می‌تواند منجر به نارضایتی و شکایات شهروندان شود، هشدار مهمی در خصوص کیفیت نامطلوب و آلودگی آب است. سنجش کیفیت آب در مناطقی که مشکل بوی نامطبوع دارند کمک می‌کند عوامل ایجاد بو شناسایی و اقدامات مقتضی جهت رفع آن و بهبود کیفیت آب صورت گیرد. برای مثال، در رودخانه‌های به‌شدت آلوده در برزیل، پژوهشگران دریافتند که غلظت‌های سولفید هیدروژن (ناشی از تجزیه گوگرد در مواد آلی) و آمونیاک (ناشی از فعالیت‌های کشاورزی، فاضلاب‌های تصفیه‌نشده و تجزیه مواد نیتروژنی) از آستلنه‌های مجاز فراتر رفته و عامل اصلی ایجاد بوی نامطبوع و شکایات گسترده عمومی بوده است [۱۷].

کیفیت نامطلوب آب در کانال‌های جمع‌آوری آب سطحی شهری یکی از چالش‌های اساسی محیط‌زیستی و مدیریت شهری است که مستقیماً بر سلامت عمومی، رفاه شهروندان و پایداری اکوسیستم‌های آبی تأثیرگذار است. کیفیت پایین آب در این کانال‌ها علاوه بر تأثیرات محیطی، استفاده ایمن از منابع آبی آنها را محدود کرده و خطرات بهداشتی را افزایش می‌دهد. این موضوع می‌تواند زندگی ساکنان مناطق اطراف این کانال‌ها را مختل کند. سازمان ملل متحد نیز با تعریف هدف توسعه پایدار ششم که بر دسترسی همگانی به آب تمیز و بهداشت تأکید دارد، به این چالش جهانی اذعان کرده است [۱]. با گسترش شهرنشینی و افزایش ورود آلاینده‌ها به این کانال‌ها، مشکلاتی نظیر کاهش کیفیت آب، تولید گازهای بدبو و انتشار آلودگی به محیط اطراف تشدید شده است. این آلاینده‌ها شامل مواد آلی، فاضلاب‌های تصفیه‌نشده، مواد شیمیایی و زباله‌های صنعتی و خانگی هستند که علاوه بر کاهش کیفیت آب، باعث تولید گازهای بدبو نظیر سولفید هیدروژن، آمونیاک و ترکیبات آلی فرار می‌شوند. انتشار بوی نامطبوع محدود به آب آلوده در رودخانه‌ها و کانال‌ها نیست و منابع مختلفی می‌تواند داشته باشد. برای مثال، واحدهای تولید کمپوست در مراکز پرورش قارچ معمولاً به دلیل تجزیه مواد آلی، موجب انتشار غلظت‌های بالای آمونیاک و سولفید هیدروژن و در نتیجه، بوی ناخوشایند می‌شوند [۲]. برخی کارخانه‌های فرآوری شیر پاستوریزه نیز، به دلیل دفع غیراصولی پسماندهای آلی، منبع تولید بوهای نامطبوع محسوب می‌شوند [۳]. فعالیت واحدهای تصفیه روغن نیز می‌تواند بوی ناخوشایندی مشابه با بوی قیر در نواحی اطراف آن ایجاد کند [۴]. همچنین، مراکز دامداری، مرغداری‌ها و گاوداری‌ها که به واسطه تجزیه فضولات دامی، آمونیاک آزاد می‌کنند نیز، از جمله منابع مهم آلودگی بو به‌شمار می‌روند [۵]. این بوهای نامطبوع، به‌ویژه در شهرهایی مانند تهران که شبکه گسترده‌ای از کانال‌های آب سطحی در مناطق پرتراکم شهری وجود دارد، می‌توانند مشکلات جدی اجتماعی و محیط‌زیستی را به همراه داشته باشند. بوی نامطبوع ناشی از فعالیت‌های بی‌هوازی و تجزیه مواد آلی در کانال‌های آب

^۲ Emscher

^۱ Cheonggye

در شرایطی که کیفیت آب کاهش یافته، میزان اکسیژن محلول کم باشد و دمای آب و هوا بالا باشد، تشدید می‌شوند. این شرایط ضرورت پایش منظم شدت بو و متغیرهای کیفی آب را نشان می‌دهد و بر لزوم مدیریت منابع آلاینده برای جلوگیری از بی‌هوازی شدن آب و تشدید بوی نامطبوع به‌ویژه در فصول گرم تأکید دارد.

بررسی تجربه‌های موفق و ناموفق احیای رودخانه‌ها و کانال‌های آب نامبرده حاکی از آن است که طراحی و اجرای یک برنامه پایش مستمر برای ارزیابی دوره‌ای کیفیت آب و انتشار بوی نامطبوع در رودخانه‌ها و کانال‌های شهری گامی ضروری است. این موضوع علاوه بر شناسایی به موقع ورود آلاینده‌های سمی، به ارزیابی اقدامات نظارتی و ارزیابی اثر بخشی سیاست‌ها و قوانین حفاظت از محیط‌زیست کمک زیادی می‌کند. بررسی مطالعات پیشین نشان می‌دهد که پژوهشگران از زوایای متنوعی سعی کرده‌اند به این مسئله بپردازند و برخی از روش‌ها را برای درک فرآیندهای ایجاد بو، توسعه روش‌های اندازه‌گیری دقیق بوی نامطبوع و ارائه راهکارهای مدیریت مؤثر به کار برده‌اند [۱۸-۲۰].

از چالش‌های اساسی در اندازه‌گیری بو، ماهیت چندعاملی ایجاد بو در سیستم‌های آبی، تغییرپذیری و پیچیدگی بو است که تحت تأثیر شرایط محیطی، ترکیبات بیوشیمیایی آب و حضور آلاینده‌هایی مانند سولفیدها، آمونیاک و ترکیبات آلی فرار قرار می‌گیرد. در مطالعات پیشین روش‌هایی مانند روش بویایی پویا و روش کیسه بو مثلثی برای اندازه‌گیری بوی نامطبوع مورد استفاده قرار گرفته‌اند، اما این روش‌ها با محدودیت‌هایی از جمله حسی بودن و دشواری در تشخیص بوهایی با آستانه پایین مواجه‌اند [۲۱]. با این حال، کمی‌سازی و شناسایی بو از طریق روش‌های پیشرفته تحلیلی به‌ویژه با ادغام تکنیک‌های حسی و ابزاری، فرصت‌های جدیدی برای ارزیابی‌های دقیق‌تر و قابل‌تکرار فراهم کرده است. برای پر کردن خلأ میان ادراک بویایی انسان و داده‌های عینی، روش‌هایی مانند کروماتوگرافی گازی-طیف‌سنجی جرمی-ولفکومتري، که پیش‌تر در تحلیل آب آشامیدنی و بررسی طعم غذا به کار رفته‌اند، به‌عنوان ابزارهایی امیدوارکننده برای شناسایی دقیق‌تر ترکیبات بوزا و ایجاد همبستگی‌های قوی‌تر بین شدت بو و غلظت شیمیایی پیشنهاد شده‌اند [۲۲]. با این حال، این روش‌ها اغلب با چالش‌هایی نظیر

هزینه، پیچیدگی عملیاتی و نیاز به تخصص‌های خاص روبرو هستند، که بر ضرورت توسعه جایگزین‌های مقرون‌به‌صرفه‌تر تأکید می‌کند.

در برخی از مطالعات اخیر از روش‌هایی نظیر یادگیری ماشین و یادگیری عمیق برای پیش‌بینی و مدیریت بو استفاده شده است [۲۳-۲۴]. با ایجاد همبستگی میان متغیرهای کیفیت آب مانند اکسیژن محلول، هدایت الکتریکی و نیاز اکسیژن بیوشیمیایی با شدت بو، پژوهشگران پتانسیل مدل‌های پیش‌بینی را برای بهبود مدیریت بو نشان داده‌اند. الگوریتم‌های یادگیری ماشین، از جمله جنگل تصادفی و شبکه‌های عصبی، در شناسایی نقاط بحرانی انتشار بو و پیش‌بینی تغییرات زمانی آن دقت بالایی از خود نشان داده‌اند [۲۴-۲۶]. این رویکردها به دلیل اجرایی بودن مزایای قابل توجهی داشته و بینش‌های عملیاتی ارزشمندی در اختیار برنامه‌ریزان شهری و تصمیم‌گیرندگان قرار می‌دهند. علاوه بر این، پیشرفت‌های فناورانه امکان پایش انتشار بو را از طریق ابزارهایی مانند بینی‌های الکترونیکی^۱ فراهم کرده‌اند [۲۷]. این دستگاه‌ها که به حسگرهای شیمیایی حساس به ترکیبات فرار مجهز هستند، بر اساس شاخص‌های ساده و قابل اندازه‌گیری، مانند pH^۲، نیاز شیمیایی اکسیژن، غلظت نیتروژن کل و فسفر کل عمل می‌کنند و امکان شناسایی مناطق بحرانی و پیش‌بینی شدت بو را فراهم می‌آورند. مطالعات انجام‌شده در کشور چین کارایی این فناوری در ارزیابی انتشار بو در رودخانه‌های آلوده را نشان داده و همبستگی‌های بالایی ($R^2 > 0.9$) بین مقادیر پیش‌بینی‌شده و مشاهداتی گزارش کرده‌اند [۲۷]. این یافته‌ها نشان می‌دهند که ترکیب روش‌های داده‌محور با ابزارهای سنجش پیشرفته می‌تواند چارچوبی مؤثر برای پایش مستمر و مدیریت هدفمند انتشار بو در محیط‌های آبی فراهم کند.

با وجود پیشرفت‌های اخیر در به‌کارگیری روش‌های داده‌محور و یا ابزارهای پیشرفته برای سنجش شدت بو، هنوز چارچوبی عملیاتی و قابل اجرا برای ارزیابی و پایش انتشار بوی نامطبوع در پهنه‌های آبی در محیط‌های شهری ارائه نشده است [۲۱]. همچنین ابزارهای فعلی قادر به تشخیص بو به صورت در محل بخصوص با آستانه تشخیص بو بسیار پایین نیستند. فقدان چنین چارچوبی

^۲ Potential of Hydrogen

^۱ E-nose

موجب شده است که درک درستی از پهنه‌بندی و توزیع مکانی شدت بوی نامطبوع و ارتباط آن با شاخص‌های کیفی آب به‌دست نیامده و در نتیجه امکان اقدامات مدیریتی به موقع و مؤثر فراهم نباشد.

بنابراین هدف اصلی این پروژه پژوهشی، توسعه یک چارچوب عملیاتی برای پایش و پیش‌بینی بوی نامطبوع در کانال‌های جمع‌آوری آب سطحی شهری است. در این راستا، کانال غیائوند در تهران به‌عنوان مطالعه موردی انتخاب شده است. این کانال ضمن ایفای نقش مهم در هدایت رواناب‌های شهری، به‌دلیل ورود فاضلاب‌های تصفیه‌نشده و تجمع آلاینده‌ها، به یکی از کانون‌های اصلی انتشار بوی نامطبوع در محله اختیاریه (منطقه ۳ شهرداری تهران) تبدیل شده است. به‌منظور بررسی این پدیده، داده‌های میدانی از بخش‌هایی که بر اساس گزارش‌های مردمی و مشاهدات میدانی بیشترین شدت بوی نامطبوع یا تجمع آلاینده‌ها در آن‌ها رخ می‌دهد گردآوری شد. در این نقاط شاخص بوی نامطبوع همراه با متغیرهای کیفی آب و هواشناسی اندازه‌گیری گردید و در ادامه با بهره‌گیری از روش‌های آماری و یک مدل یادگیری ماشین، چارچوبی داده‌محور برای پیش‌بینی شدت بو توسعه یافت. سوالات تحقیقاتی مهم در این پروژه شامل آن است که:

- (۱) آیا می‌توان شاخص بوی نامطبوع را به‌صورت کمی بیان کرد؟
- (۲) پهنه‌بندی بوی نامطبوع و جانمایی نقاط بحرانی در طول کلنال جمع‌آوری آب سطحی در منطقه مطالعاتی به چه صورتی هست؟
- (۳) چه عواملی بیشترین تأثیر را بر شدت بو دارند؟ (۴) آیا الگو و رابطه‌ای بین متغیرهای کیفی آب و شدت انتشار بوی نامطبوع وجود دارد؟ (۵) آیا می‌توان با اندازه‌گیری داده‌های کیفی در دسترس و قابل اندازه‌گیری در محل، شدت بو را در زمان و مکان با دقت مناسب پیش‌بینی کرد؟

نتایج این تحقیق می‌تواند مبنای توسعه سامانه‌های پایش و مدیریت بو در سایر کانال‌های شهری باشد و به بهبود کیفیت زندگی در مناطق شهری کمک کند. در نهایت، چارچوب پیشنهادی این پژوهش می‌تواند الگویی قابل تعمیم برای طراحی راهکارهای بومی در مدیریت بوی نامطبوع در مناطق مختلف شهری باشد و زمینه‌ساز سیاست‌گذاری‌های مؤثرتر در ارتقاء بهداشت محیط و رفاه عمومی گردد.

۲- روش شناسی

۲-۱- معرفی محدوده مطالعاتی

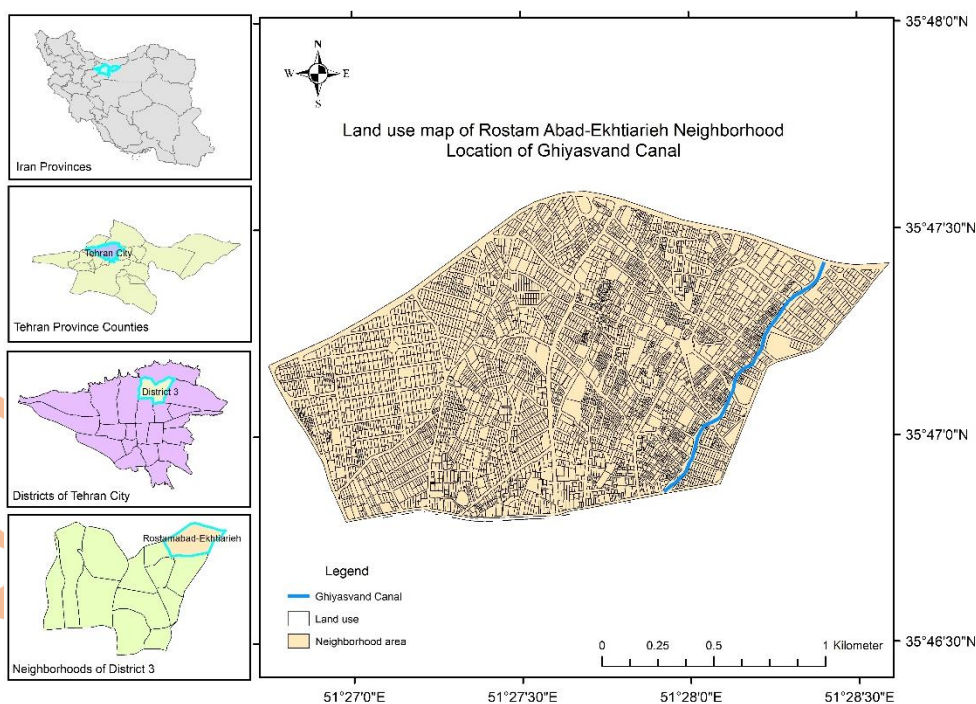
کانال مورد مطالعه در این پژوهش، در ناحیه ۳ منطقه ۳ شهرداری تهران و در محله رستم آباد- اختیاریه واقع شده است که موقعیت مکانی آن در شکل (۱) نمایش داده شده است. منطقه ۳ شهرداری تهران یکی از مناطق شهری تهران است که در پهنه شمال شرقی شهر تهران واقع شده است. این منطقه با دارا بودن ۳۰/۲۰۸ کیلومتر مربع مساحت، از لحاظ وسعت در مناطق ۲۲ گانه تهران در رتبه دهم شهر تهران قرار دارد و ۴/۳ درصد از کل وسعت شهر تهران را شامل می‌شود [۲۸]. کانال غیائوند در منطقه ۱ از تقاطع خیابان باهنر و شهید آقایی آغاز شده و در ادامه‌ی کانال جمشیدیه قرار گرفته است. کانال مذکور کاملاً در محدوده شهری قرار گرفته و آب‌های سطحی درون شهری را جمع‌آوری کرده و به پایین‌دست انتقال می‌دهد. مساحت حوضه تحت پوشش این کانال با احتساب حوضه‌های برون شهری ۲۹/۵۵ کیلومتر مربع می‌باشد. این کانال در راستای شمالی-جنوبی پس از تقاطع با بزرگراه شهید صدر وارد منطقه ۳ شده و مسیر آن در راستای شمالی-جنوبی و به موازات خیابان‌های غیائوند، خاش و هیرمند ادامه می‌یابد. ارتفاع متوسط در این منطقه حدوداً ۱۴۰۰ متر، جهت شیب شمال به جنوب و به طور متوسط حدود ۷/۵ درصد می‌باشد [۲۹]. با توجه به موقعیت شهری و الزامات شهرسازی، ساختار کانال به‌صورت متناوب شامل بخش‌های روباز، مدفون و سرپوشیده طراحی شده است. طول کل کانال غیائوند برابر با ۲۷۸۲ متر بوده که شامل ۱۵۰۲ متر بخش روباز، ۱۸۵ متر بخش مدفون و ۱۰۹۵ متر بخش سرپوشیده می‌باشد [۲۹]. مشاهدات میدانی حاکی از آن است که در سال‌های اخیر، مجدداً بخش قابل‌توجهی از این کانال به‌دلیل قرارگیری در بافت متراکم شهری، به‌صورت سرپوشیده درآمده است. با این حال، کانال غیائوند در حدفصل زیر پل بزرگراه شهید صدر تا خیابان شهید کلاهدوز هم‌چنان به صورت روباز است که در این مطالعه مورد بررسی قرار گرفته است. مقطع کانال در این محدوده به شکل مستطیلی با عرض ۳ متر و ارتفاع دیواره بین ۲ تا ۳/۵ متر است. جنس دیواره کانال با طول مسیر سنگی با ملات ماسه سیمان و جنس کف آن سیمانی با ضریب زبری مانینگ ۰/۰۱۵ می‌باشد. ظرفیت مقطع کانال نیز از

تعمیرگاه‌ها، ایستگاه‌های تاکسی و پارکینگ‌ها حامل آلاینده‌هایی نظیر روغن، سوخت، ذرات معلق و ضایعات آلی هستند که مستقیماً وارد جریان آب می‌شوند. فضاهای سبز موجود در حاشیه کانال (از جمله بوستان مازیار) با ورود کودها و سموم شیمیایی موجب افزایش بار آلودگی و کاهش کیفیت آب کانال می‌شوند. دفع نامناسب زباله‌های جامد در کاربری‌های تجاری مانند سوپرمارکت‌ها و مشاورین املاک نیز، از دیگر منابع آلاینده در کانال مورد مطالعه به‌شمار می‌روند.

۲-۲- داده‌برداری میدانی

در این پروژه پژوهشی، ۱۰ نقطه بحرانی با توجه به شدت بو مشاهده شده در کانال انتخاب شدند و به صورت هفتگی از نقاط بحرانی در طول ۱۰ هفته در بازه زمانی شهریور تا آذر ۱۴۰۳ نمونه‌برداری شد. برای اندازه‌گیری شدت بوی نامطبوع در این مطالعه از دستگاه سنجنده بو با مدل OMX-SRM استفاده شده است. این

۰/۳۷ تا ۰/۹۹ متر مکعب بر ثانیه در طول کانال تغییر می‌کند [۲۹]. مطابق با نمونه‌برداری‌های انجام شده در محدوده زمانی مورد مطالعه، دمای آب به طور میانگین از ۲۲/۷ درجه سانتیگراد در اواخر شهریور به تدریج به ۱۵/۷ درجه سانتیگراد در اواخر پاییز کاهش یافت. همچنین، میانگین تراز آب در کانال در این بازه زمانی حدود ۲۰ سانتیمتر، متوسط سرعت آب حدود ۰/۶ متر بر ثانیه و متوسط دبی حدود ۰/۳۶ متر مکعب بر ثانیه تخمین زده شده است. در مورد حجم و موقعیت دقیق ورودی‌های فاضلاب به کانال گزارش دقیقی وجود ندارد، اما بررسی‌های میدانی در مورد کاربری‌های اطراف کانال و مصاحبه با مسئولان و شهروندان حاکی از آن است که منازل مسکونی، مدارس، رستوران‌های موجود در حاشیه آن و همچنین پارکینگ حمل زباله شهرداری در مجاورت میدان خاش با تولید فاضلاب‌های حاوی مواد آلی، شوینده‌ها و چربی‌ها در صورت دفع نامناسب، موجب افزایش بار آلی و آلودگی میکروبی کانال می‌شوند. رواناب‌های سطحی ناشی از بارش در خیابان‌های منتهی به کانال،



شکل ۱: موقعیت مکانی کانال جمع‌آوری آب سطحی غیاثوند در شهر تهران

از آب کانال برداشت می‌شد تا متغیرهای کیفی آب در محل مورد سنجش قرار گیرند. برای سنجش متغیرهای غلظت کل جامدات

دستگاه با مقایسه غلظت بو در هوا با هوای تصفیه‌شده، شدت نسبی بو را به صورت عددی تعیین می‌کند. سپس نمونه‌ای به حجم ۲ لیتر

محلول در آب (TDS)، هدایت الکتریکی، دمای آب، pH و شوری آب از دستگاه مولتی پارامتر مدل Apera PC400S استفاده شد. برای اندازه گیری اکسیژن محلول در آب نیز از دستگاه اکسیژن سنج Hanna HI98193 استفاده شد. همچنین داده های ثبت شده هواشناسی نظیر رطوبت هوا، سرعت باد و دمای هوا در ایستگاه سینوپتیک تهران- شمیران به عنوان نزدیک ترین ایستگاه جمع آوری شدند.

۲-۳- مدل سازی و پیش بینی بوی نامطبوع

همانطور که اشاره شد، یکی از اهداف اصلی از این پژوهش یافتن الگو و رابطه بین متغیرهای کیفی آب و شدت انتشار بوی نامطبوع می باشد. از آنجائیکه که رابطه تجربی مشخصی بین متغیرهای معمول کیفی آب (اعم از دمای آب، اکسیژن محلول، هدایت الکتریکی، متغیرهای آب و هوایی و ...) و شدت بوی نامطبوع وجود نداشته است، بنابراین استفاده از مدلسازی های آماری به طور خاص یادگیری ماشین که مبتنی بر داده ها می باشد، انتخاب مناسبی برای این هدف می باشد. یکی از پرکاربردترین روش ها در یادگیری ماشین الگوریتم جنگل تصادفی می باشد که در این پژوهش مورد استفاده قرار گرفته است. الگوریتم جنگل تصادفی^۱ یکی از روش های یادگیری ماشین مبتنی بر یادگیری گروهی^۲ است که از ترکیب چندین درخت تصمیم گیری^۳ برای بهبود دقت پیش بینی و جلوگیری از بیش برآزش^۴ استفاده می کند. این الگوریتم برای مسائل طبقه بندی^۵ و رگرسیون^۶ به کار می رود و توسط لئو بریمن^۷ در سال ۲۰۰۱ معرفی شده است [۳۰]. مطالعات پیشین در مسائل مهندسی آب و محیط زیست نشان داده اند که الگوریتم جنگل تصادفی در پیش بینی کیفیت آب در مقایسه با سایر روش های یادگیری ماشین عملکرد بسیار قابل قبولی دارد. تحقیقات گذشته نشان می دهد این مدل نه تنها توانسته است متغیرهای مختلفی مانند نیتروژن کل، فسفر، کدورت و شوری را با دقت بالا پیش بینی کند، بلکه حتی با داده های محدود نیز نتایج قابل اتکایی ارائه داده و به عنوان ابزاری مؤثر در مدیریت منابع آبی به کار رفته

است [۳۱-۳۵]. مدل جنگل تصادفی یک مدل از زیرشاخه های روش های مبتنی بر درخت است قادر است بدون نیاز به تنظیمات پیچیده پارامترها، روابط غیرخطی میان چند متغیر مستقل و متغیر هدف را به طور همزمان مدلسازی و کمی سازی کند. در برخی از مطالعات پیشین، الگوریتم جنگل تصادفی با موفقیت برای مدل سازی و مدیریت بوی نامطبوع نیز به کار گرفته شده است. برای مثال، در یک پژوهش در مراکز تاسیسات بازگردانی منابع آب، از این الگوریتم برای پیش بینی غلظت ترکیباتی مانند H_2S و تعیین دوز بهینه مواد کنترل کننده بو استفاده شده است [۳۶]. همچنین در مطالعه ای دیگر، مدل جنگل تصادفی برای پیش بینی شکایات ناشی از بوی نامطبوع بر اساس داده های حسگرهای H_2S و اطلاعات هواشناسی به کار رفته و دقت قابل قبولی در هشدار زودهنگام بوی نامطبوع ارائه داده است [۳۷]. این شواهد نشان می دهند که جنگل تصادفی می تواند ابزار مؤثری در تحلیل و پیش بینی بوی نامطبوع در سامانه های آبی باشد.

در این مطالعه برای ساخت مدل پیش بینی بوی نامطبوع مراحل مختلفی مورد نیاز است که به طور خلاصه نمودار جریان روند انجام مدلسازی در شکل (۲) آمده است. مرحله اول داده برداری میدانی است که در بخش ۲-۲ توضیح داده شد. در مرحله بعد لازم است داده های جمع آوری شده پیش پردازش شوند تا از نبود داده های پرت اطمینان حاصل شود و از ورود این نوع داده ها در مجموعه داده قبل از آموزش مدل ممانعت به عمل آید. سپس در مرحله تحلیل داده های اکتشافی، به منظور تحلیل روابط میان شدت بوی نامطبوع (متغیر هدف) و متغیرهای کیفی و اقلیمی (ویژگی ها)، از روش های تحلیل اکتشافی داده ها و به ویژه از ماتریس های همبستگی استفاده شد. هدف از این تحلیل، شناسایی الگوها، درک نوع و شدت وابستگی میان متغیرها، و کاهش ابعاد مسئله از طریق حذف ویژگی های کم اهمیت برای توسعه مدلی دقیق تر بود. برای سنجش نوع همبستگی میان متغیرها، سه روش آماری مورد استفاده قرار گرفت: ضریب همبستگی پیرسون^۸ برای

^۵ Classification

^۶ Regression

^۷ Leo Breiman

^۸ Pearson

^۱ Random Forest

^۲ Ensemble Learning

^۳ Decision Trees

^۴ Overfitting

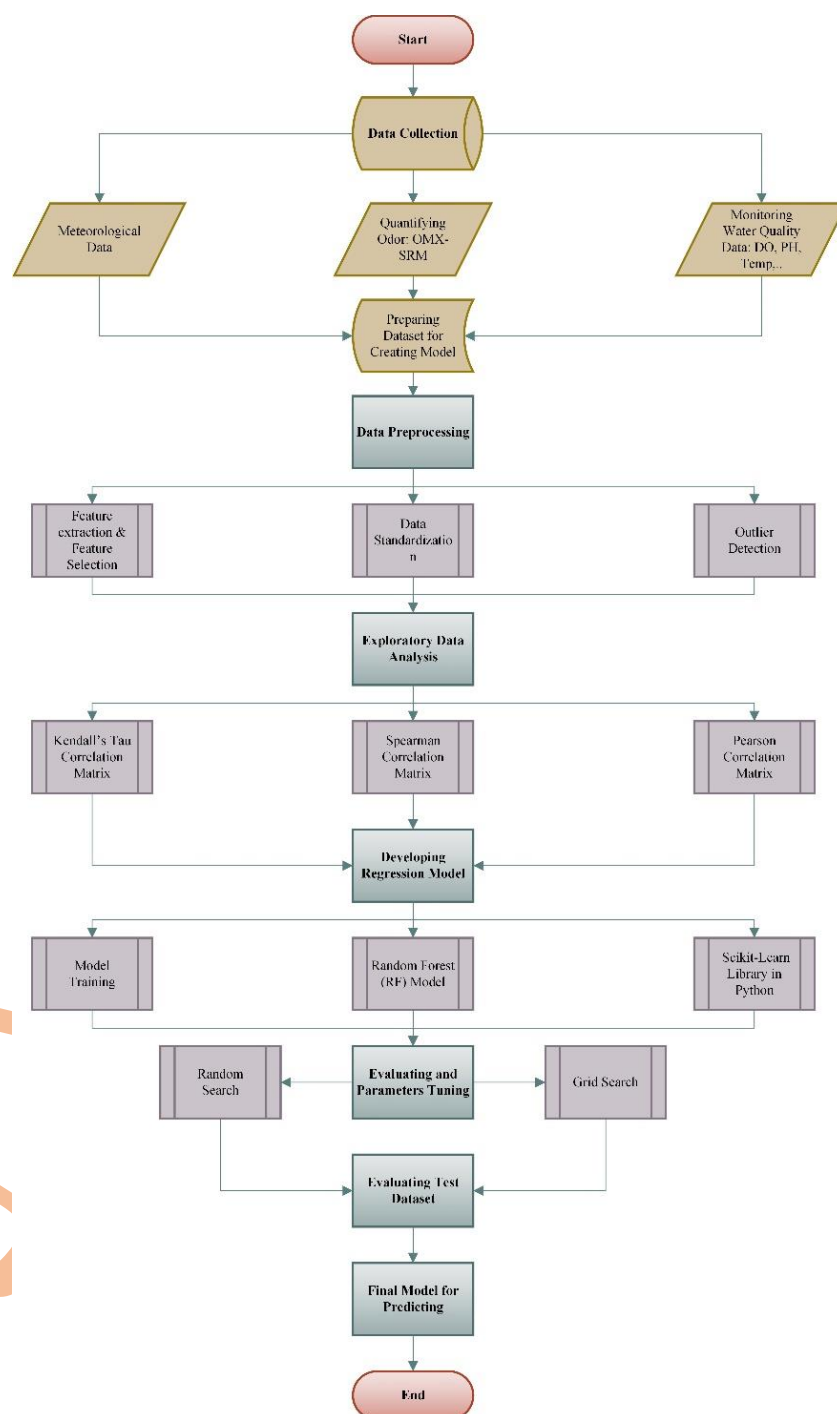
ارزیابی دقیق و چندبعدی روابط بین متغیرهای مؤثر را فراهم ساخت و مبنای انتخاب ترکیب بهینه ویژگی‌ها در فرآیند مدل‌سازی قرار گرفت.

بررسی روابط خطی، ضریب همبستگی اسپیرمن^۱ برای ارزیابی روابط غیرخطی با داده‌های رتبه‌ای، و همچنین ضریب کندال^۲ که به‌ویژه برای داده‌هایی با حجم نمونه کم عملکرد بهتری دارد. استفاده از این سه شاخص در قالب ماتریس‌های همبستگی، امکان

داده‌های
ارزیابی
نشانده

^۲ Kendall's tau

^۱ Spearman



شکل ۲: نمودار جریان مراحل اجرایی مدل‌سازی پیش‌بینی بوی نامطبوع

آموزش نیز به منظور بهینه‌سازی عملکرد مدل، به دو زیرمجموعه شامل داده‌های آموزش (۸۵٪) و داده‌های اعتبارسنجی (۱۵٪) تفکیک شد. مدل رگرسیون جنگل تصادفی بر روی داده‌های

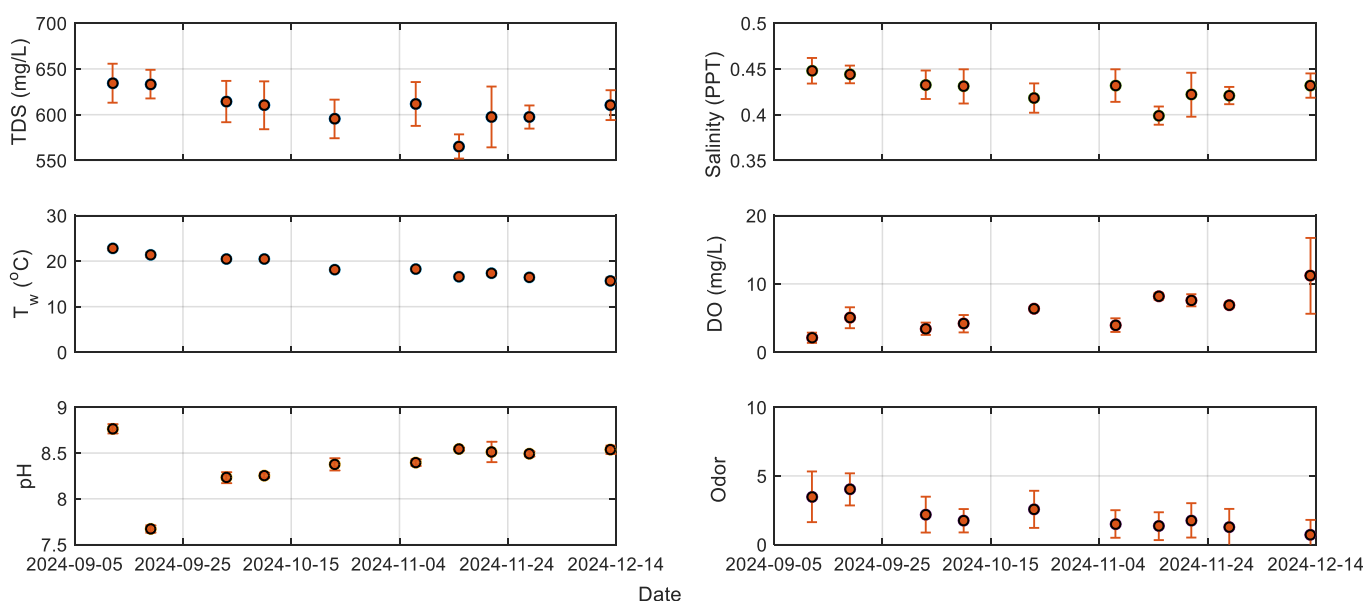
برای ساخت مدل پیش‌بینی شدت بوی نامطبوع، پس از انجام مراحل پیش‌پردازش و تحلیل ویژگی‌ها، مجموعه داده نهایی به دو بخش آموزش (۸۰٪) و آزمون (۲۰٪) تقسیم شد. مجموعه داده

ارائه شده است. سری زمانی میانگین روزانه پارامترهای کیفی آب در کانال غیاثوند نیز در شکل (۳) نمایش داده شده است. در طول داده‌برداری تغییرات هدایت الکتریکی و غلظت کل جامدات محلول به ترتیب بین ۷۷۴ تا ۸۵۴ $\mu\text{S}/\text{cm}$ و ۵۵۰ تا ۶۰۶ mg/L بوده است. نتایج ماتریس‌های همبستگی که در شکل (۳) ارائه شده‌اند، نشان می‌دهد بین هدایت الکتریکی و شدت بو، همبستگی مثبت و متوسط وجود دارد (ضریب پیرسون و اسپیرمن حدود ۰.۳۶ و ضریب کندال حدود ۰.۲۴). نتایج همبستگی بین غلظت کل مواد جامد محلول و شوری نیز مشابه با هدایت الکتریکی است. به عبارت دیگر، با افزایش مقدار هدایت الکتریکی و غلظت کل مواد جامد محلول و شوری تمایل به افزایش شدت بو وجود دارد. این یافته‌ها با مطالعات قبلی همخوانی دارد، جایی که همبستگی مثبت بین پارامترهای کیفی آب مانند هدایت الکتریکی و انتشار ترکیبات بویایی مانند H_2S و NH_3 در

آموزش، آموزش داده شد [۳۸] و عملکرد آن به‌صورت هم‌زمان با استفاده از داده‌های اعتبارسنجی مورد ارزیابی قرار گرفت. در نهایت، برای ارزیابی دقت مدل نهایی، از داده‌های آزمون که در فرایند آموزش استفاده نشده بودند، بهره گرفته شد. سپس به‌منظور بهینه‌سازی عملکرد مدل و دستیابی به بهترین ترکیب پارامترها، ابتدا از روش جست‌وجوی تصادفی^۱ [۳۸] برای شناسایی بازه‌های مناسب پارامترهای مدل استفاده شد. در مرحله بعد، از روش جست‌وجوی شبکه‌ای^۲ [۳۸] برای یافتن دقیق‌ترین ترکیب از پارامترها استفاده شد تا مدل نهایی با بیشترین دقت و کارایی بهینه‌سازی گردد.

۳- نتایج و بحث

خلاصه‌ای از متغیرهای کیفی آب و متغیرهای هواشناسی که در دوره پایش در کانال غیاثوند اندازه‌گیری شده است در جدول (۱)



شکل ۳: سری زمانی میانگین روزانه پارامترهای کیفی آب در کانال غیاثوند

بلکه بین یکدیگر نیز همبستگی بسیار بالایی دارند. این موضوع به این دلیل است که دستگاه سنجنده تنها یکی از این متغیرها را مستقیماً اندازه‌گیری کرده و بقیه از طریق روابط معین محاسبه

رودخانه‌های شهری آلوده گزارش شده است [۱۷]. همانطور که نتایج نشان می‌دهد، متغیرهای هدایت الکتریکی، کل جامدات محلول و شوری نه تنها با متغیر پیش‌بینی همبستگی مشابهی دارند

^۲ Grid Search

^۱ Random Search

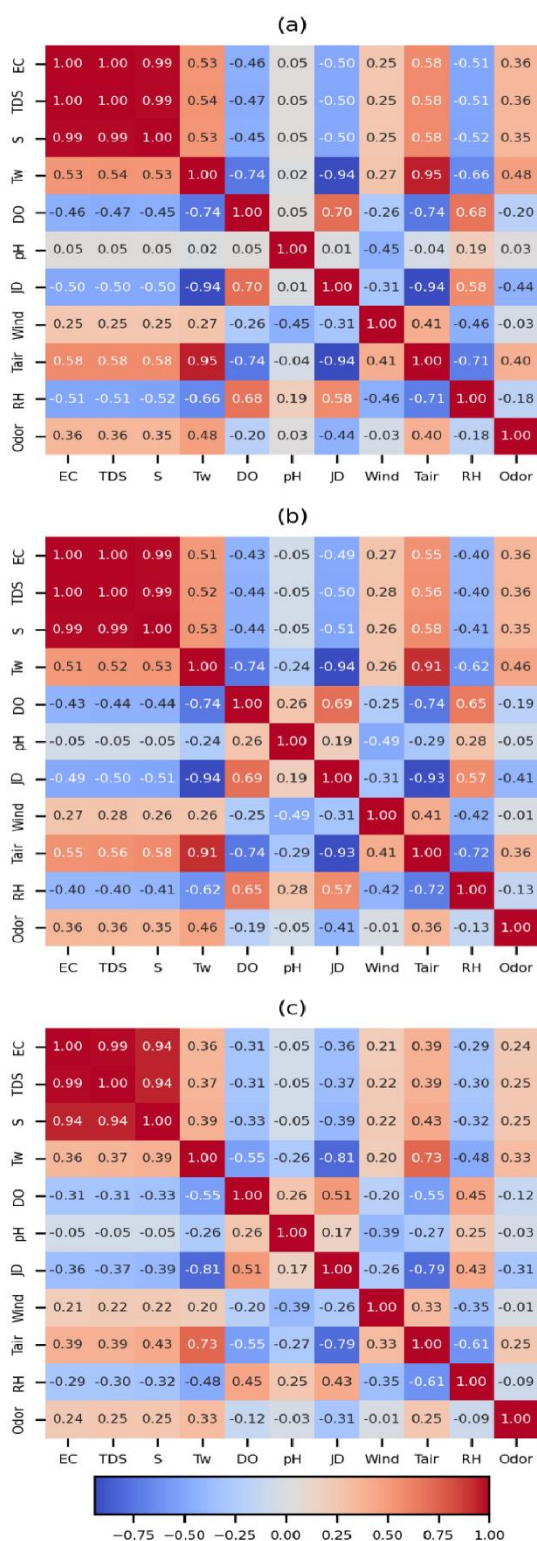
شده‌اند. بنابراین، برای کاهش ابعاد مسئله و افزایش دقت مدل، تنها متغیر هدایت الکتریکی برای مدلسازی انتخاب شده است. شوری و pH در این مدت تغییرات کمی داشته و به ترتیب بین ۰.۳۹ تا ۰.۴۳ PPT و ۷.۶۷ تا ۸.۳۶ ثبت شده است. نتایج همبستگی بین pH و شدت بو نشان می‌دهد همبستگی معنی‌داری وجود ندارد (ضرایب حدود ۰.۰۳-۰.۰۴). در طول داده‌برداری سرعت باد در صبح بین ۰ تا ۴ متر بر ثانیه متغیر بوده، اما در بعد از ظهر نوسانات کمی بین ۱ تا ۲ متر بر ثانیه داشته است. تحلیل ضرایب همبستگی بین سرعت باد و شدت بو نشان می‌دهد سرعت باد تأثیر یا ارتباط قابل توجهی با تغییرات شدت بو ندارد.

جدول ۱: خلاصه متغیرهای کیفی آب و متغیرهای هواشناسی در دوره پایش در کانال غیاثوند

Variables	Unit	Notation	Min	Avg	Max
Electrical Conductivity	$\mu\text{S/cm}$	<i>EC</i>	۷۷۴	۸۵۴	۹۵۰
Total Dissolved Solids	mg/L	<i>TDS</i>	۵۵۰	۶۰۶	۶۷۴
Salinity	PPT	<i>S</i>	۰.۳۹	۰.۴۳	۰.۴۷
Water Temperature	$^{\circ}\text{C}$	<i>T_w</i>	۱۵.۶	۱۹.۱	۲۳.۲
Dissolved Oxygen	mg/L	<i>DO</i>	۱.۵۶	۵.۳۷	۸.۸۰
pH	-	<i>pH</i>	۷.۶۷	۸.۳۶	۸.۸۳
Relative Humidity	%	<i>RH</i>	۹.۹	۲۸.۷۵	۶۲.۰
Air Temperature	$^{\circ}\text{C}$	<i>T_{air}</i>	۱۰.۰	۲۰.۳	۳۴.۰
Wind Velocity	m/s	<i>Wind</i>	۰	۱.۶۱	۴.۰
Odor Intensity	-	<i>Oder</i>	۰	۲.۱	۶.۱

دمای هوا در اولین روز داده‌برداری حدود ۳۴ درجه سانتی‌گراد و به تدریج روندی کاهشی داشته و در آخرین روز داده‌برداری دمای هوا بین ۱۰ تا ۱۲ درجه سانتی‌گراد ثبت شده‌است. نتایج حاصل از آزمون‌های پیرسون، اسپیرمن و کندال نشان می‌دهد که بین دمای هوا و شدت بو یک رابطه همبستگی مثبت متوسط وجود دارد (ضریب پیرسون ۰.۳۹، ضریب اسپیرمن ۰.۳۶ و ضریب کندال ۰.۲۵) و با افزایش دمای هوا معمولاً شدت بو نیز افزایش می‌یابد. روند تغییرات دمای آب نیز مشابه با دمای هوا بوده و از مقدار ۲۳.۲ درجه سانتی‌گراد در اولین روز داده‌برداری به مقداری بین ۱۶ تا ۱۷ درجه سانتی‌گراد در آخرین روز کاهش یافته‌است. نتایج آزمون‌های پیرسون، اسپیرمن و کندال حاکی از آن است که

همبستگی بین دمای آب و شدت بو کمی قوی‌تر از دمای هوا با شدت بو است (ضریب پیرسون ۰.۴۷، ضریب اسپیرمن ۰.۴۶ و ضریب کندال ۰.۳۳). به عبارت دیگر، تغییرات دمای آب نسبت به تغییرات دمای هوا تأثیر بیشتری در افزایش شدت بو دارند. این نتایج با تحقیقات قبلی همسو است که نشان می‌دهد افزایش دما می‌تواند انتشار ترکیبات بویایی را از طریق افزایش فعالیت میکروبی و نرخ تبخیر تسهیل کند [۳۹]. رطوبت نسبی نیز در طول داده‌برداری روندی افزایش داشته است. در روز اول داده‌برداری مقدار رطوبت نسبی معادل ۱۰ درصد و در صبح داده‌برداری روز هشتم نیز به بیش از ۶۰ درصد می‌رسد. نتایج به‌دست‌آمده از آزمون‌های پیرسون، اسپیرمن و کندال نشان می‌دهند که بین رطوبت نسبی و شدت بو همبستگی منفی ضعیفی وجود دارد (ضریب پیرسون ۰.۱۸-، ضریب اسپیرمن ۰.۱۳- و ضریب کندال ۰.۰۹-). به عبارت دیگر با افزایش رطوبت نسبی، شدت بو تمایل به کاهش دارد، هرچند ارتباط به صورت ضعیف و نسبتاً نامحسوس است. مقدار اکسیژن محلول در آب در اولین روز داده‌برداری کمترین مقدار خود را داشته و مقداری معادل با ۱.۵۶ mg/L ثبت شده است. به تدریج با کاهش دما و در اثر بارندگی‌ها در بالادست مقدار اکسیژن محلول به طور نسبی افزایش پیدا کرده است. به طوری که بالاترین مقدار آن معادل با ۸.۸ mg/L در انتهای داده‌برداری در ۲۵ آبان تا ۸ آذر به ثبت رسیده است. نتایج آزمون‌های پیرسون، اسپیرمن و کندال نشان می‌دهند که بین اکسیژن محلول و شدت بو یک همبستگی منفی ضعیف وجود دارد (ضریب پیرسون ۰.۲۰-، ضریب اسپیرمن ۰.۱۸- و ضریب کندال ۰.۱۲-). این نتیجه به این معنا است که افزایش اکسیژن محلول ممکن است با کاهش کمی در شدت بو همراه باشد، اما تأثیر آن به صورت نسبی ضعیف است. این همبستگی منفی با یافته‌های مطالعات پیشین سازگار است و نشان می‌دهد سطوح پایین DO با افزایش انتشار H_2S همراه است، که یکی از عوامل اصلی بوی نامطبوع در آب‌های سطحی آلوده محسوب می‌شود [۱۷]. شاخص بو نیز با طیف وسیعی از مقادیر در محدوده ۰ تا ۶/۱ به ثبت رسیده است. میانگین این شاخص ۲.۱ بوده که به صورت بوی نامطبوع متوسط قابل توصیف می‌باشد. نتایج آزمون‌های آماری نشان می‌دهد که بین شماره روز داده‌برداری (JD) و شدت بو، رابطه همبستگی منفی متوسط وجود دارد. به بیان دیگر،



شکل ۴: ماتریس‌های همبستگی پیرسون (a)، اسپیرمن (b) و کندال تاو (c) بین متغیرهای کیفی آب، هواشناسی و شدت بوی نامطبوع در کانال غیاثوند

همان‌طور که انتظار می‌رود با گذر زمان با توجه به کاهش دمای هوا و دمای آب به دلیل تغییر فصل شدت بو کاهش می‌یابد.

در این مطالعه، به منظور ارزیابی تأثیر ترکیب‌های مختلف متغیرهای کیفی و هواشناسی بر دقت مدل‌سازی شدت بو، با استفاده از الگوریتم رگرسیون جنگل تصادفی مدلهایی با ورودی‌های متفاوت از متغیرهای کیفی و هواشناسی ساخته شد تا مشخص شود کدام ترکیب از متغیرها عملکرد بهتری در پیش‌بینی دارد. این رویکرد مشابه با استفاده از شبکه‌های عصبی مصنوعی (ANN) برای پیش‌بینی غلظت بو بر اساس پارامترهای کیفی آب مانند BOD، DO و pH در کارخانه‌های تصفیه فاضلاب است [۲۶]. در جدول (۲) مدلهایی که بر اساس R^2 و RMSE بهترین نتیجه را براساس بهینه‌سازی و ترکیب پارامترها داشته‌اند، ارائه شده‌اند. مقایسه مدل‌های مختلف نشان می‌دهد که ترکیب متغیرهای ورودی تأثیر قابل توجهی بر عملکرد مدل جنگل تصادفی دارد. در برخی از ترکیب‌ها، اضافه کردن و در نظرگیری تعامل متغیرها سبب شکل‌گیری مدل‌های با دقت بالا شده است. برای مثال مدل‌های شماره ۳، ۶ و ۸ متغیرهای تعاملی دارند. در مدل شماره ۶ متغیرهایی از حاصل ضرب متغیر شمارنده روز در متغیرهای رطوبت نسبی، دمای هوا و اکسیژن محلول به وجود آمده است که عملکرد مدل را افزایش داده است. مدل ۸ شامل متغیرهای دمای هوا، شمارنده روز و تعامل آن با متغیرهای اکسیژن محلول، سرعت باد و رطوبت نسبی به بالاترین دقت منجر شده است. مقدار R^2 و RMSE بدست آمده برای داده‌های آموزش ۰.۸۳۵ و ۰.۵۸۹ و برای داده‌های آزمون به ترتیب برابر با ۰.۸۲۴ و ۰.۵۸۱ بوده است. این نتایج نشان می‌دهد که مدل توسعه داده شده نه تنها از توانایی بالایی در تطبیق با داده‌های آموزشی برخوردار بوده، بلکه قابلیت تعمیم مناسبی برای پیش‌بینی در شرایط نادیده نیز دارد. همان‌طور که در جدول (۲) مشخص است در همه مدل‌های منتخب رطوبت نسبی، دمای هوا و اکسیژن محلول

به عنوان متغیر ورودی استفاده شده‌اند. این موضوع نشان دهنده اهمیت این سه متغیر در پیش‌بینی شدت بوی نامطبوع می‌باشد. همچنین، مدل‌هایی که شماره روز را در ترکیب نگه داشته‌اند، معمولاً عملکرد بهتری نسبت به مدل‌های فاقد آن داشته‌اند که نشان از نقش مهم روند زمانی در بروز بوی نامطبوع دارد. در مقابل، مدل‌هایی که متغیر هدایت الکتریکی و دمای آب را حذف کرده‌اند دقت کمتری داشته‌اند. به طور کلی، ترکیبی از متغیرهای کیفی کلیدی، متغیرهای هواشناسی و شاخص زمانی، بهترین کارایی را در مدلسازی شدت بو

جدول ۲: بهترین ترکیبات مدلسازی با استفاده از شاخص R^2 و RMSE

Model	Feature Combinations	RMSE Metric		R^2 Metric	
		Test	Train	Test	Train
۱	$EC, T_w, DO, JD, T_{air}, RH$	۰.۸۴۱	۰.۶۵۱	۰.۶۳۱	۰.۷۹۸
۲	$pH, DO, JD, Wind, T_{air}, RH$	۰.۷۴۷	۰.۶۷۷	۰.۷۱۰	۰.۷۸۲
۳	$JD, T_{air}, pH \times JD, pH \times T_{air}, DO \times JD$	۰.۶۷۱	۰.۵۰۵	۰.۷۶۵	۰.۸۷۸
۴	pH, DO, T_{air}, RH	۰.۶۳۵	۰.۴۷۲	۰.۷۹۰	۰.۸۹۴
۵	DO, JD, T_{air}, RH	۰.۷۵۹	۰.۶۰۱	۰.۷۰۰	۰.۸۲۸
۶	$JD, T_{air}, DO \times JD, JD \times T_{air}, JD \times RH$	۰.۶۹۸	۰.۶۱۷	۰.۷۴۶	۰.۸۱۹
۷	$DO, JD, Wind, T_{air}, RH$	۰.۷۰۹	۰.۶۸۵	۰.۷۳۸	۰.۷۷۶
۸	$JD, T_{air}, DO \times JD, JD \times Wind, JD \times RH$	۰.۵۸۱	۰.۵۸۹	۰.۸۲۴	۰.۸۳۵
۹	$T_w, pH, DO, JD, Wind, T_{air}, RH$	۰.۸۴۵	۰.۷۸۵	۰.۶۲۸	۰.۷۰۷
۱۰	pH, DO, JD, T_{air}, RH	۰.۶۹۱	۰.۵۶۸	۰.۷۵۱	۰.۸۴۶

استفاده شد. نتایج بهینه‌سازی پارامترهای جنگل تصادفی در جدول (۳) آورده شده است. برای مثال در بهترین مدل (شماره ۸) نتایج بهینه‌سازی با استفاده از این الگوریتم‌ها، تعداد درخت‌های تصمیم‌گیری برابر با ۲۰ و عمق پیشینه درخت‌ها برابر با ۸ تعیین شد که نشان‌دهنده ساختار نسبتاً فشرده اما مؤثر مدل است. حداقل تعداد نمونه‌های لازم برای تقسیم یک گره نیز برابر با ۲ بدست آمد. همچنین، حداقل تعداد نمونه‌هایی که باید در یک برگ درخت قرار گیرد برابر با ۲ بدست آمد که نشان‌دهنده انعطاف بالای مدل در یادگیری الگوهای موجود در داده است. معیار ارزیابی مناسب برای تقسیم گره‌ها در این مدل خطای مطلق پیش‌بینی‌ها تعیین شد و در هر گره از درخت تصمیم‌گیری، فقط به تعداد ریشه دوم آن از ویژگی‌های تصادفی انتخاب شدند تا داده‌ها بر اساس آن‌ها تقسیم شوند. در نهایت، نمونه‌برداری مجدد (Bootstrap) غیرفعال شد تا برای ساخت هر درخت کل داده‌های آموزش به‌جای نمونه‌برداری با

در مدل شماره ۴ با در نظرگیری متغیرهای ساده و کم (اکسیژن محلول، رطوبت نسبی، دمای هوا و pH) دقت شاخص R^2 برای داده‌های آموزش و آزمون به ترتیب ۰.۸۹۴ و ۰.۷۹۰ بدست آمده است. مدل‌های شماره ۳، ۶ و ۸ به ترتیب متشکل و تعامل یافته از مدل‌های ۲، ۵ و ۸ می‌باشد که همانطور که ملاحظه می‌شود با در نظرگیری تعامل متغیر شمارنده روز در متغیرهای کیفی و اقلیمی دقت شاخص‌ها در مجموعه آموزش و آزمون افزایش یافته است. این تعامل به نوعی رفتار فصلی و ماهانه متغیرهای کیفی اقلیمی را نشان می‌دهد که سبب افزایش دقت مدل شده است. نکته قابل توجه در مدل ساخته شده در سناریو شماره ۹ این است که با افزایش و در نظرگیری متغیرهای بیشتر، دقت مدل نسبت به سناریوهای دیگر کاهش یافته است. همانطور که پیش‌تر اشاره شد، به منظور افزایش دقت پیش‌بینی و بهینه‌سازی ساختار مدل جنگل تصادفی، از دو الگوریتم بهینه‌سازی جستجوی تصادفی و جستجوی شبکه‌ای

جایگزینی مورد استفاده قرار گیرند. این تنظیمات به طور کلی نسبت به مدل‌های قبلی منجر به توسعه مدلی بهینه‌تر با دقت بالاتر و تعمیم‌پذیری مناسب‌تر بر روی داده‌های آزمون شد. به طوری که علیرغم دقت بالا، اختلاف دقت در داده‌های آموزش و آزمون حداقل شده است و نشانه‌ای از بیش‌برازش در آن وجود ندارد. شکل (۴)، نتایج پیش‌بینی را برای داده‌های آزمون و مقایسه آن با مقادیر واقعی داده‌ها نمایش می‌دهد. مطابق با این نتایج مدل با شاخص R^2 معادل ۰.۸۳ توانسته است شدت بوی نامطبوع را به صورت کمی پیش‌بینی نماید. شاخص‌های آماری R^2 ، MAE و RMSE برای مدل نهایی برای داده‌های آموزش و آزمون در جدول (۲) آورده شده است. نتایج این جدول نشان می‌دهد که دقت مدل قابل قبول بوده و عملکرد

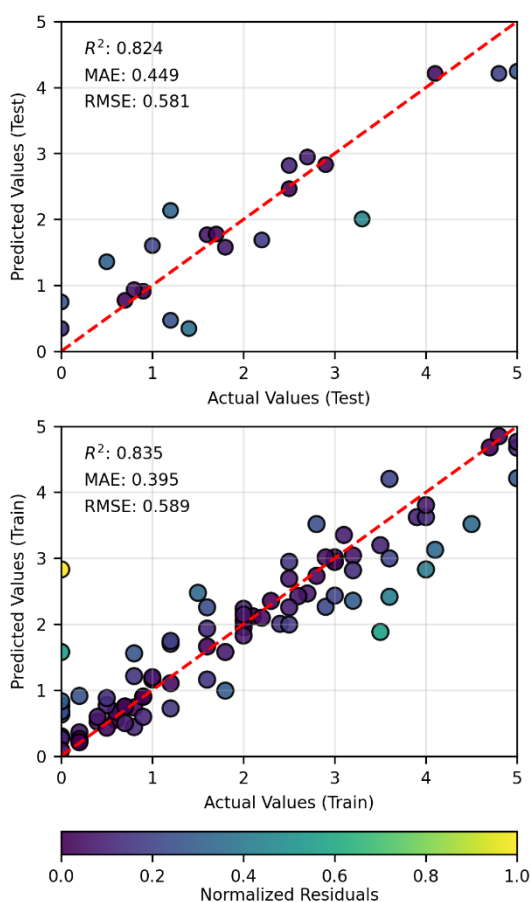
خوبی داشته است. لازم به ذکر است تفاوت بین مقادیر شاخص‌های بدست آمده در داده‌های آموزش و آزمون در برخی از مدل‌ها، به دلیل این است که مجموعه داده‌ها به صورت میدانی در هفته‌های مختلف جمع‌آوری شده‌اند، بنابراین تعداد نمونه‌ها محدود به دوره داده برداری بوده است. بنابراین انتظار می‌رود با پایش بلندمدت و جمع‌آوری مجموعه داده بزرگ (به عنوان مثال با بیش از ۱۰۰۰ داده)، این اختلاف بین این دو مجموعه داده کاهش پیدا کرده و دقت و اعتمادپذیری مدل‌های پیش‌بینی در داده‌های آزمون نیز همچون عملکرد آن در پیش‌بینی داده‌های آموزش به طور قابل توجهی افزایش پیدا کند.

جدول ۳: پارامترهای بهینه الگوریتم جنگل تصادفی در مدل منتخب

Model	Bootstrap	Criterion	Max Features	Min Samples Leaf	Min Samples Split	Max Depth	Number of Trees (n_estimators)
۱	True	Squared error	Sqrt	۱	۴	۱۴	۸
۲	True	Squared error	Sqrt	۱	۲	۱۴	۴
۳	False	Squared error	Sqrt	۲	۲	۱۴	۱۲
۴	False	Absolute error	Sqrt	۱	۴	۸	۸
۵	False	Squared error	Sqrt	۱	۶	۸	۱۶
۶	False	Absolute error	Sqrt	۲	۲	۸	۱۲
۷	False	Squared error	Sqrt	۲	۶	۸	۱۲
۸	False	Absolute error	Sqrt	۲	۲	۸	۲۰
۹	True	Squared error	Sqrt	۲	۶	۱۴	۱۲
۱۰	False	Squared error	Sqrt	۲	۲	۱۴	۴

در پایان، نتایج این مطالعه نشان داد که استفاده از داده‌های میدانی در کنار مدل‌سازی داده‌محور می‌تواند رویکرد مؤثری برای پیش‌بینی شدت بوی نامطبوع در کانال‌های آب سطحی شهری باشد. مدل نهایی توسعه‌یافته با بهره‌گیری از الگوریتم جنگل تصادفی و بهینه‌سازی دقیق پارامترها، توانست با دقتی قابل قبول ($R^2 \approx 0.83$) شدت بو را پیش‌بینی کند. این سطح از دقت، به‌ویژه با توجه به محدودیت تعداد نمونه‌های میدانی، بسیار امیدوارکننده است و نشان می‌دهد که مدل ارائه‌شده ظرفیت بالایی برای تعمیم‌پذیری دارد. با

توسعه شبکه‌های پایش و افزایش حجم داده‌ها در آینده، می‌توان دقت مدل را بیش از پیش ارتقا داد. یافته‌های این بخش پایه‌ای علمی برای طراحی سامانه‌های پایش مستمر بو در شبکه آب‌های سطحی فراهم کرده و می‌تواند به عنوان ابزاری تصمیم‌یار برای مدیریت شهری و کنترل آلودگی‌های بویایی مورد استفاده قرار گیرد. در تحقیقات آتی، مطالعاتی با پایش‌های بلندمدت و حجم داده بیشتر می‌توانند بستر مناسبی برای به‌کارگیری مدل‌های ترکیبی پیشرفته و یادگیری



شکل ۵: نتایج واقعی و پیش‌بینی شاخص بو

۴ - نتیجه‌گیری

بوی نامطبوع منتشرشده از کانال‌های آب سطحی شهری یکی از معضلات روبه‌رشد در کلان‌شهرهایی همچون تهران است که نه تنها بر کیفیت محیط زیست و سلامت عمومی تأثیرگذار است، بلکه می‌تواند موجب نارضایتی گسترده شهروندان شود. در این مطالعه، با تمرکز بر کانال غیاثوند به‌عنوان مطالعه موردی، برای نخستین بار تلاش شد چارچوبی عملیاتی برای کمی‌سازی و پیش‌بینی شدت بوی نامطبوع بر اساس پارامترهای قابل اندازه‌گیری کیفیت آب و شرایط هواشناسی ارائه شود. تحلیل همبستگی نشان داد که متغیرهایی مانند دمای آب، دمای هوا، هدایت الکتریکی و اکسیژن محلول نقش مؤثری در شدت بوی نامطبوع دارند. به‌ویژه، دمای آب بیشترین همبستگی مثبت و اکسیژن محلول بیشترین همبستگی منفی را با

عمیق فراهم کنند و به ارتقاء بیشتر دقت و قابلیت تعمیم چارچوب پیشنهادی منجر شوند.

این مطالعه با ارائه چارچوبی ترکیبی مبتنی بر داده‌های میدانی، شاخص‌های کیفی آب، متغیرهای هواشناسی و مدل یادگیری ماشین جنگل تصادفی، گامی نوآورانه در جهت پیش‌بینی و مدیریت بوی نامطبوع در کانال‌های آب سطحی شهری برداشته است. نوآوری اصلی این پژوهش در تلفیق داده‌برداری میدانی با مدلسازی هوشمند و دستیابی به دقت بالا در پیش‌بینی شدت بو، آن هم در شرایط محدودیت داده، نهفته است. مدل نهایی با R^2 برابر با ۰.۸۲۴ برای داده‌های آزمون، قابلیت تعمیم‌پذیری مناسبی را نشان داده و می‌تواند مبنای تصمیم‌سازی برای اقدامات مدیریتی هدفمند در شهرهای بزرگ قرار گیرد. این چارچوب نه تنها در سطح منطقه‌ای، بلکه در مقیاس وسیع‌تر نیز قابل توسعه است و می‌تواند در سایر رودخانه‌ها، کانال‌های شهری و حتی در محیط‌های صنعتی برای پایش سریع و کم‌هزینه بوی نامطبوع توسعه یافته و مورد استفاده قرار گیرد. پیشنهاد می‌شود در مطالعات آتی با افزایش حجم داده‌ها از طریق پایش‌های بلندمدت، استفاده از حسگرهای آنلاین و ترکیب روش‌های پیشرفته یادگیری عمیق، دقت و قابلیت اطمینان این چارچوب پیش از پیش ارتقاء یابد و به‌عنوان بخشی از سیستم‌های هوشمند مدیریت کیفیت محیط‌زیست شهری به کار گرفته شود. همچنین در تحقیقات آینده، توسعه یک مدل مفهومی و یکپارچه که علاوه بر پیش‌بینی، قابلیت تصمیم‌سازی و کنترل را نیز پوشش دهد، با بهره‌گیری از شبکه‌های پایش گسترده‌تر و داده‌های بلندمدت پیشنهاد می‌شود.

اگرچه حجم داده‌های این مطالعه به یک دوره ۱۰ هفته‌ای محدود است و پیش‌بینی‌های بلندمدت و تعمیم مدل به کانال‌های دیگر ممکن است با عدم قطعیت‌هایی همراه باشد، نتایج نشان‌دهنده عملکرد قابل اعتماد مدل در شناسایی روندهای کوتاه‌مدت و ویژگی‌های کیفی آب در محدوده مطالعه است. این امر اهمیت مطالعه را به عنوان پایه‌ای محکم برای توسعه مدل‌های جامع‌تر و اعمال در کانال‌های دیگر افزایش می‌دهد. برای تعمیم و تقویت پیش‌بینی‌های بلندمدت، جمع‌آوری داده‌های طولانی‌تر و متنوع‌تر، به‌کارگیری تکنیک‌های تقویت داده یا انتقال یادگیری و اعتبارسنجی مدل با داده‌های مستقل از زمان یا مکان دیگر، مسیرهای پژوهشی آینده و گام‌های منطقی برای بهبود و توسعه این مدل محسوب می‌شوند.

محیط زیست مرکز مطالعات و برنامه ریزی شهر تهران که مسئولیت راهبری و هدایت پروژه را داشتند و همچنین از ناظر علمی پروژه جناب آقای دکتر مجید حسین زاده نیز تشکر و قدردانی می شود. دستگاه سنجش بو که در این پروژه پژوهشی استفاده شده است با استفاده از اعتبار اعطایی جایزه دکتر کاظمی آشتیانی از سوی بنیاد ملی نخبگان تهیه شده است و لازم است از این بنیاد گرانقدر به دلیل حمایت های ارزشمندشان صمیمانه قدردانی شود. در انتها از حمایت های معاونت پژوهشی دانشگاه صنعتی شریف نیز تشکر و قدردانی می نمایم.

منابع

[۱] United Nations, The sustainable development goals report ۲۰۱۸, ۲۰۱۸. [Online]. Available: <https://unstats.un.org/sdgs/report/۲۰۱۸/goal-۰۶/>.

[۲] Yue Sun, Zihan Xu, Nan Zhang, Dongdong Hu, Yao Liang, Chang Liu, Yinxi Zhu, Xiaomeng Chen, Unveiling the intrinsic relationship between ammonia and hydrogen sulfide generation during composting, *Environmental Technology & Innovation*, Volume 39, ۲۰۲۵, ۱۰۴۲۹۸, ISSN ۲۳۵۲-۱۸۶۴, <https://doi.org/10.1016/j.eti.2025.104298>.

[۳] Lu, M. Lamichhane, P. Liang, F., Identification of odor causing compounds in a commercial dairy farm. *Water Air Soil Pollut: Focus* 8, ۳۵۹-۳۶۷ ۲۰۰۸. <https://doi.org/10.1007/s11267-007-9150-x>.

[۴] Han, B. Liu, Yt. Wu, Jh., Characterization of industrial odor sources in Binhai New Area of Tianjin, China. *Environ Sci Pollut Res* 25, ۱۴۰۰۶-۱۴۰۱۷ ۲۰۱۸. <https://doi.org/10.1007/s11356-018-1596-z>.

[۵] Riaz, Sobia & Kareem, Hafiz Abdul & Mehmood, Rizwan & Sadia, Haleema. ۲۰۲۱. Livestock-waste, types, sources, pollution potential, *Country wise comparisons*. <https://doi.org/10.1201/9781003283621-9>.

[۶] Valipoori Goodarzi, H. Fadaei Tehrani, M.R. Mohammad Zadeh, S., Investigation regarding assessment of potentially toxic elements (PTEs) contamination risk in Rasht city with emphasis on the two rivers, Goharroud and Zarjoub, Gilan province, Iran. *Sustain. Water Resour. Manag.* ۱۰, ۵۰ ۲۰۲۴. <https://doi.org/10.1007/s40899-023-01025-8>.

[۷] Khalili, R. Parvinnia, M. and Motaghi, H., ۲۰۲۰. Evaluation of Bashar River water quality using CCME water quality index. *Journal of Environmental Science Studies*, ۵(۳), pp. ۲۸۰۷-۲۸۱۴.

شاخص بو نشان دادند. همچنین مدلی بر پایه الگوریتم یادگیری ماشین جنگل تصادفی توسعه یافت که با استفاده از ترکیب بهینه ای از متغیرهای کیفی و هواشناسی، توانست شدت بو را با دقت بالا پیش بینی کند ($R^2 = 0.824$ برای داده های آزمون). این مدل پس از بهینه سازی با استفاده از الگوریتم های جستجوی شبکه ای و تصادفی، به ساختاری فشرده و کارآمد با قابلیت تعمیم پذیری مطلوب دست یافت. همچنین، نتایج نشان داد که متغیرهایی مانند دمای هوا، رطوبت نسبی، اکسیژن محلول و شماره روز داده برداری به عنوان ورودی در تمام مدل های برتر حضور داشته و نقش کلیدی در افزایش دقت مدل ایفا کرده اند.

یافته های این پژوهش نشان می دهد که بوی نامطبوع آب سطحی نه تنها تابع کیفیت لحظه ای آب است، بلکه تحت تأثیر روندهای زمانی، اقلیمی و تغییرات فصلی نیز قرار دارد. چارچوب پیشنهادی این مطالعه می تواند به عنوان مدلی کاربردی برای شناسایی، پهنه بندی و پیش بینی بوی نامطبوع در دیگر کانال ها و رودخانه های شهری توسعه یابد و در آینده به عنوان پایه ای برای طراحی سامانه های پایش هوشمند و تصمیم یار در مدیریت شهری مورد استفاده قرار گیرد. این چارچوب نه تنها در سطح منطقه ای، بلکه در مقیاس وسیع تر نیز قابلیت پیاده سازی دارد و می تواند در سایر رودخانه ها، کانال های شهری و حتی محیط های صنعتی برای پایش سریع، مقرون به صرفه و هدفمند بوی نامطبوع استفاده شود. از آنجا که این پژوهش به صورت پایلوت و با تعداد محدودی نمونه انجام شده است، انتظار می رود که با پایش بلندمدت و افزایش حجم داده ها، دقت مدل و اعتبار نتایج به طور چشمگیری ارتقاء یابد. در نهایت، این پژوهش گامی عملیاتی در جهت رفع یک نیاز واقعی در حوزه مدیریت شهری است و پایش بوی نامطبوع می تواند مسیر طراحی سیاست ها و اقدامات پیشگیرانه در برابر این معضل را هموارتر سازد.

۵- تشکر و قدردانی

این مقاله برگرفته از نتایج پروژه پژوهشی انجام شده به شماره قرارداد ۱۳۷/۳۰۵۰۲۶ در مرکز مطالعات و برنامه ریزی شهر تهران است و نویسندگان این مقاله از حمایت های این مرکز کمال تشکر را دارند. از سرکار خانم دکتر زهرابرزگر مدیر گروه محترم مطالعات

- [۱۸] Hawko, C. Verrielle, M. Hucher, N. Crunaire, S. Leger, C. Locoge, N. and Savary, G., ۲۰۲۱. A review of environmental odor quantification and qualification methods: The question of objectivity in sensory analysis. *Science of The Total Environment*, ۷۹۵, p.۱۴۸۸۶۲. <https://doi.org/10.1016/J.SCITOTENV.2021.148862>
- [۱۹] Bax, C. Sironi, S. and Capelli, L., ۲۰۲۰. How can odors be measured? An overview of methods and their applications. *Atmosphere*, ۱۱(۱), p.۹۲. <https://doi.org/10.3390/atmos11010092>
- [۲۰] Sado-Inamura, Y. and Fukushi, K., ۲۰۱۸. Considering water quality of urban rivers from the perspectives of unpleasant odor. *Sustainability* 10 (3), 650 [online]. <https://doi.org/10.3390/su10030650>
- [۲۱] Wang, Y. Shao, L. Kang, X. Zhang, H. Lü, F. and He, P., ۲۰۲۳. A critical review on odor measurement and prediction. *Journal of Environmental Management*, ۳۳۶, p.۱۱۷۶۵۱. <https://doi.org/10.1016/j.jenvman.2023.117651>
- [۲۲] Lebrero, R. Bouchy, L. Stuetz, R. and Muñoz, R., ۲۰۱۱. Odor assessment and management in wastewater treatment plants: a review. *Critical Reviews in Environmental Science and Technology*, ۴۱(۱۰), pp.۹۱۵-۹۵۰. <https://doi.org/10.1080/10643389.2011.600000>
- [۲۳] Wei, L. Huang, C. Wang, Z. Wang, Z. Zhou, X. and Cao, L., ۲۰۱۹. Monitoring of urban black-odor water based on Nemerow index and gradient boosting decision tree regression using UAV-borne hyperspectral imagery. *Remote Sensing*, ۱۱(۲۰), p.۲۴۰۲. <https://doi.org/10.3390/rs11202402>
- [۲۴] Yang, F. Pluth, T.B. Fang, X. Francq, K.B. Jurjovec, M. and Tang, Y., ۲۰۲۱. Advanced machine learning application for odor and corrosion control at a water resource recovery facility. *Water Environment Research*, ۹۳(۱۱), pp.۲۳۴۶-۲۳۵۹. <https://doi.org/10.1002/wer.1618>
- [۲۵] Mulrow, J. Kshetry, N. Brose, D.A. Kumar, K. Jain, D. Shah, M. Kunetz, T.E. and Varshney, L.R., ۲۰۲۰. Prediction of odor complaints at a large composite reservoir in a highly urbanized area: A machine learning approach. *Water Environment Research*, ۹۲(۳), pp. ۴۱۸-۴۲۹. <https://doi.org/10.1002/wer.1191>
- [۲۶] Kang, J.H. Song, J. Yoo, S.S. Lee, B.J. and Ji, H.W., ۲۰۲۰. Prediction of odor concentration emitted from wastewater treatment plant using an artificial neural network (ANN). *Atmosphere*, ۱۱(۸), p. ۷۸۴. <https://doi.org/10.3390/atmos11080784>
- [۲۷] Qiu, S. Hou, P. Huang, J. Han, W. and Kang, Z., ۲۰۲۱. The monitoring of Black-Odor River by electronic nose with chemometrics for pH, COD, TN, and TP. *Chemosensors*, ۹(۷), p.۱۶۸. <https://doi.org/10.3390/chemosensors9070168>
- [۲۸] Iran. Tehran Municipality. ۲۰۱۵. Operational studies for implementation of the comprehensive surface water management plan and preparation of improvement
- [۲۹] Shakeri, A. Moore, F., The impact of an industrial complex on freshly deposited sediments, Chener Rahdar river case study, Shiraz, Iran. *Environ Monit Assess* 169, ۳۲۱-۳۳۴ ۲۰۱۰. <https://doi.org/10.1007/s10661-009-1173-5>
- [۳۰] seyednezhad, N. Salimi Gol Sheikhi, M. Barati, R., ۲۰۲۲. 'Analytical note: Discussion to determine the network of problems of Kashaf Rood river', *Journal of Water and Sustainable Development*, ۹(۲), pp. ۱۳۵-۱۳۷.
- [۳۱] Cao, J. Sun, Q. Zhao, D. Xu, M. Shen, Q. Wang, D. Wang, Y. and Ding, S., ۲۰۲۰. A critical review of the appearance of black-odorous waterbodies in China and treatment methods. *Journal of hazardous materials*, ۳۸۵, p.۱۲۱۵۱۱. <https://doi.org/10.1016/j.jhazmat.2019.121511>
- [۳۲] Choe, L.J. Jung, S.W. Kim, D.G. Baek, M.J. Kang, H.J. Lee, C.Y. and Bae, Y.J., ۲۰۱۴. Temporal changes in benthic macroinvertebrates and their interactions with fish predators after restoration in the C heonggyecheon, a downtown stream in S eoul, K orea. *Entomological Research*, ۴۴(۶), pp.۳۳۸-۳۴۸. <https://doi.org/10.1111/1374-5967.12080>
- [۳۳] Gerner, N.V. Nafu, I. Winking, C. Wencki, K. Strehl, C. Wortberg, T. Niemann, A. Anzaldúa, G. Lago, M. and Birk, S., ۲۰۱۸. Large-scale river restoration pays off: A case study of ecosystem service valuation for the Emscher restoration generation project, *Ecosyst. Serv.*, ۳۰, ۳۲۷-۳۳۸, <https://doi.org/10.1016/j.ecoser.2018.03.002>
- [۳۴] Romero, E. Le Gendre, R. Garnier, J. Billen, G. Fisson, C. Silvestre, M. and Riou, P., ۲۰۱۶. Long-term water quality in the lower Seine: Lessons learned over ۴ decades of monitoring. *Environmental Science & Policy*, ۵۸, pp.۱۴۱-۱۵۴. <https://doi.org/10.1016/j.envsci.2016.01.016>
- [۳۵] Chițescu, C.L. Ene, A. Geana, E.I. Vasile, A.M. and Ciucure, C.T., ۲۰۲۱. Emerging and persistent pollutants in the aquatic ecosystems of the Lower Danube Basin and North West Black Sea Region—A review. *Applied sciences*, ۱۱(۲۰), p.۹۷۲۱. <https://doi.org/10.3390/app11209721>
- [۳۶] Krishna, C.R. Klein, C.R. Jones, K.W. Clesceri, N.L. and Stern, E.A., ۱۹۹۵. Human exposure to toxic materials. *The Mt Sinai J of Med*, ۶۲, p.۵.
- [۳۷] Liang, K. Feng, X. Yang, B. Dou, Z. and Chen, W., ۲۰۲۱, February. Research on treatment measures and treatment effects of black and smelly water body in shenyang area. In IOP Conference Series: Earth and Environmental Science (Vol. 651, No. ۴, p. ۰۴۲۰۴۳). IOP Publishing. <https://doi.org/10.1088/1755-1315/651/4/042043>
- [۳۸] Cruz, L.P. Alves, R.S. da Rocha, F.O. Moreira, M.S. and dos Santos Júnior, A., ۲۰۲۲. Atmospheric levels, multivariate statistical study, and health risk assessment of odorous compounds (H₂S and NH₃) in areas near polluted urban rivers in the city of Salvador, in Northeastern Brazil. *Air Quality, Atmosphere & Health*, ۱۵(۱), pp.۱۵۹-۱۷۶. <https://doi.org/10.1007/s11869-021-1095-۷>

Water Environment Research, 92(3), pp.418-429.
<https://doi.org/10.1002/wer.1191>

[38] Pedregosa, F. Varoquaux, G. Gramfort, A. Michel, V. Thirion, B. Grisel, O. Blondel, M. Prettenhofer, P. Weiss, R. Dubourg, V. and Vanderplas, J., 2011. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, pp.2825-2830.

[39] Piccardo, M.T. Geretto, M. Pulliero, A. and Izzotti, A., 2022. Odor emissions: A public health concern for health risk perception. *Environmental Research*, 204, p.112121. <https://doi.org/10.1016/j.envres.2021.112121>

plans for streams and canals in the 22 districts of Tehran City – Conceptual design studies of District 3 – Second Edition, Volume 1. *Tehran: Yekom Consulting Engineers* [In Persian].

[29] Iran. Tehran Municipality. 2015. Operational studies for implementation of the comprehensive surface water management plan and preparation of improvement plans for streams and canals in the 22 districts of Tehran City – Conceptual design studies of District 3 – Second Edition, Volume 2. *Tehran: Yekom Consulting Engineers* [In Persian].

[40] Breiman, L., 2001. Random forests. *Machine learning*, 45, pp.5-32. <https://doi.org/10.1023/A:10109334.4324>

[31] Wang, F. Wang, Y. Zhang, K. Hu, M. Weng, Q. and Zhang, H., 2021. Spatial heterogeneity modeling of water quality based on random forest regression and model interpretation. *Environmental Research*, 202, p.111660. <https://doi.org/10.1016/j.envres.2021.111660>

[32] Alnahit, A.O. Mishra, A.K. and Khan, A.A., 2022. Stream water quality prediction using boosted regression tree and random forest models. *Stochastic Environmental Research and Risk Assessment*, 36(9), pp.2661-2680. <https://doi.org/10.1007/s00477-021-02152-4>

[33] Sakaa, B. Elbeltagi, A. Boudibi, S. Chaffai, H. Islam, A.R.M.T. Kulimushi, L.C. Choudhari, P. Hani, A. Brouziyne, Y. and Wong, Y.J., 2022. Water quality index modeling using random forest and improved SMO algorithm for support vector machine in Saf-Saf river basin. *Environmental Science and Pollution Research*, 29(32), pp.48491-48508. <https://doi.org/10.1007/s11356-022-18644-x>

[34] Fang, X. Li, X. Zhang, Y. Zhao, Y. Qian, J. Hao, C. Zhou, J. and Wu, Y., 2021. Random forest-based understanding and predicting of the impacts of anthropogenic nutrient inputs on the water quality of a tropical lagoon. *Environmental Research Letters*, 16(5), p.055003. 10.1088/1748-9326/abf39d

[35] Xu, J. Xu, Z. Kuang, J. Lin, C. Xiao, L. Huang, X. and Zhang, Y., 2021. An alternative to laboratory testing: random forest-based water quality prediction framework for inland and nearshore water bodies. *Water*, 13(22), p.3262. <https://doi.org/10.3390/w13223262>

[36] Yang, F. Pluth, T.B. Fang, X. Francq, K.B. Jurjovec, M. and Tang, Y., 2021. Advanced machine learning application for odor and corrosion control at a water resource recovery facility. *Water Environment Research*, 93(1), pp.2346-2359. <https://doi.org/10.1002/wer.1618>

[37] Mulrow, J. Kshetry, N. Brose, D.A. Kumar, K. Jain, D. Shah, M. Kunetz, T.E. and Varshney, L.R., 2020. Prediction of odor complaints at a large composite reservoir in a highly urbanized area: A machine learning approach.