

تشخیص و طبقه‌بندی آلودگی نفتی خاک با استفاده از سامانه بینی الکترونیکی و روش‌های تشخیص الگو

محمد سروش امینی^۱، سید سعید محتسبی^{۲*}، شاهین رفیعی^۳ و احمد علی پوربابائی^۴

چکیده

در چند دهه اخیر آلاینده‌های متنوعی از طریق فعالیت‌های انسانی و به مقدار بسیار زیاد وارد محیط‌زیست شده است. آلودگی نفتی یکی از این آلاینده‌ها است و می‌تواند برای انسان و محیط‌زیست زیان‌آور باشد که تشخیص و پاکسازی آن از اهمیت فراوانی برخوردار است. با توجه به قابلیت روش‌های مبتنی بر بینی الکترونیکی در حوزه‌های غذایی، صنعتی و محیط‌زیست، هدف از این پژوهش امکان‌سنجی به کارگیری فناوری بینی الکترونیکی برای تشخیص آلودگی در یک خاک آلوده به نفت خام است. در این پژوهش از یک سامانه بینی الکترونیکی متشکل از ۸ حسگر نیمه‌هادی اکسید فلزی استفاده شد. خاک مورد استفاده در این پژوهش، خاک یکی از مزارع اطراف کرج بود که در آزمایشگاه با نفت خام در ۴ غلظت ۲۰۰ ppm، ۵۰۰ ppm، ۱۰۰۰ ppm و ۵۰۰۰ ppm آلوده شد و مورد آزمایش قرار گرفت. این سامانه به همراه روش‌های تشخیص الگو هم‌چون تحلیل مؤلفه‌های اصلی (PCA)، تحلیل تفکیک خطی (LDA) و ماشین بردار پشتیبان (SVM) مورد تحلیل و بررسی قرار گرفت. نتایج حاصل از روش‌های تشخیص الگو با دقت ۹۲ درصد برای تحلیل مؤلفه‌های اصلی، ۱۰۰ درصد برای تحلیل تفکیک خطی و ۱۰۰ درصد برای روش ماشین بردار پشتیبان به دست آمد. هم‌چنین با به کارگیری تحلیل آماری داده‌ها با روش آزمایش فاکتوریل در قالب طرح بلوک تصادفی و استفاده از الگوریتم‌های رگرسیون درخت تصمیم، رگرسیون بردار پشتیبان، رگرسیون گرادیان تقویت‌یافته و رگرسیون جنگل تصادفی با نرخ تشخیص سطوح آلودگی با میانگین ۹۵/۵ درصد برای حسگر MQ3 با بهترین عملکرد، تفاوت میان غلظت‌های آلوده شده به خوبی نشان داده شد. درخاتمه، با مقایسه نتایج حاصل از حسگرهای انتخابی بینی الکترونیکی با نتایج به دست آمده توسط سامانه کروماتوگرافی گازی بر روی خاک آلوده به نفت خام، حسگرهای مورد استفاده از حساسیت قابل قبولی برخوردار بوده‌اند.

واژه‌های کلیدی: حسگرهای گازی، نفت خام، بویایی ماشین، خاک آلوده، یادگیری ماشین.

ارجاع: امینی م. س. محتسبی س. س. رفیعی ش. و پوربابائی ا. ع. ۱۴۰۳. تشخیص و طبقه‌بندی آلودگی نفتی خاک با استفاده از سامانه بینی الکترونیکی و روش‌های تشخیص الگو. نشریه پژوهش‌های مکانیک ماشین‌های کشاورزی. ۳۱: ۴۳-۵۴. <https://dx.doi.org/10.22034/JRMAM.2024.14581.680>

۱- کارشناسی ارشد، گروه مهندسی ماشین‌های کشاورزی، دانشکده کشاورزی، دانشگاه تهران، کرج، ایران.

۲- استاد گروه مهندسی ماشین‌های کشاورزی، دانشکده کشاورزی، دانشگاه تهران، کرج، ایران.

۳- استاد گروه مهندسی ماشین‌های کشاورزی، دانشکده کشاورزی، دانشگاه تهران، کرج، ایران.

۴- استاد گروه مهندسی علوم خاک، دانشکده کشاورزی، دانشگاه تهران، کرج، ایران.

* نویسنده مسئول: mohdaseb@ut.ac.ir

تاریخ دریافت: ۱۴۰۲/۰۸/۱۱ تاریخ پذیرش: ۱۴۰۲/۱۰/۱۰

مقدمه

امروزه قسمت مهمی از آلودگی خاک مربوط به آلودگی نفتی است که از راه‌های گوناگونی وارد بافت خاک می‌شود. سالانه مقدار زیادی حادثه نشت نفت در جهان رخ می‌دهد که اکثر آن‌ها در ابعاد کوچک بوده و قابل پاکسازی هستند، اما برخی هم زبان‌آور و غیرقابل مهار هستند (Abbasi Maedeh *et al.*, 2017). اجزای سرطان‌زا مانند هیدروکربن‌های آروماتیک چند حلقه‌ای موجود در نفت، پس از آلوده کردن منابع آب و خاک به طور مستقیم، زندگی و سلامت انسان را به خطر می‌اندازند (Kang *et al.*, 2020). برای سنجش و اندازه‌گیری آلودگی خاک استانداردهای مشخصی تعیین شده است (Li *et al.*, 2020) و نمونه خاک آلوده پس از بررسی، با این استانداردها مقایسه شده و سنجش آلودگی خاک صورت می‌گیرد (Oliaei & Heidarpoor, 2013). فرآیند استفاده مجدد از زمین‌های آلوده نیازمند آزمایش سطوح آلاینده و ارزیابی خاک است (Douglas *et al.*, 2017). در ارزیابی خاک، تنفس خاک یک گام کلیدی در چرخه مداوم مواد کربن آلی و معدنی است و بازخورد آن بر روی زمین و بوم‌سازگان‌ها تأثیرگذار است. در طول دهه‌ها، برای ارزیابی خاک از تله‌های قلیایی برای جذب CO₂ تولید شده توسط نمونه خاک مرطوب در یک محفظه بسته استفاده شده است. اخیراً، یک ابزار دقیق و بسیار حساس برای ارزیابی پویای تنفس خاک توسعه یافته است که بسیار گران بوده و نیازمند نیروی متخصص و مجرب دارد. با توجه به قیمت مناسب و سرعت بالا در رسیدن به جواب مطلوب، توجه به سمت حسگرهای نیمه‌هادی اکسید فلزی جلب شده است (Luan *et al.*, 2016). برای ساختن یک سامانه بینی الکترونیک نیاز به اطلاع دقیق از نحوه کارکرد و جزئیات دقیق دستگاه بویایی انسان است، به شکلی که این سامانه بتواند همانند بینی انسان ارزیابی دقیقی از رایحه‌ها و طبقه‌بندی آن‌ها داشته باشد. حس بویایی یکی از حواس شیمیایی است. در واقع این حس با تشخیص ترکیبات شیمیایی و یا مخلوطی از آن‌ها عملکرد خود را نشان می‌دهد (Ramesh *et al.*, 2019). عملکرد حس بویایی را می‌توان به چهار مرحله اساسی زیر تقسیم نمود: تشخیص یک رایحه، تعیین وجود آن رایحه در محیط، تعیین شدت منبع رایحه و ثبت کردن و به خاطر سپردن رایحه‌های متفاوت

با فهم چگونگی شناسایی و تشخیص تفاوت موجود در بین رایحه‌ها (Mascini *et al.*, 2018). اولین وسیله‌ای که برای سنجش ویژگی‌های بویایی و تغییرات شیمیایی مواد مورد استفاده پژوهش‌گران قرار گرفت، بینی انسان بود. اما استفاده از بینی انسان محدودیت‌های متفاوتی را به همراه داشته و یک رویکرد علمی برای غلبه بر آن، استفاده از سامانه بینی الکترونیک است. اصول عملکرد این سامانه براساس سامانه بویایی پستانداران برنامه‌ریزی و ساخته شده است (Tan & Xu, 2020). مجموعه حسگرهای بینی الکترونیک شبیه گیرنده‌های بویایی طبیعی است که مواد فرار به آن متصل می‌شوند و در نتیجه تغییرات فیزیکی و شیمیایی منجر به ایجاد سیگنال می‌گردند. حسگرها، این سیگنال الکتریکی کالیبره شده را به رایانه وارد می‌کنند که در واقع منجر به تشخیص الگو می‌شود (Mortazavi *et al.*, 2023). نرم‌افزار E-nose olfaction مقدار محاسبه شده‌ای به نام شاخص عطر عادی (NAI¹) ارائه می‌دهد. هر چه شدت ترکیبات فرار بیشتر باشد، NAI بالاتر است (Wilson, 2013). با توجه به پژوهش Pineda & Perez، قرائت حسگرها با توجه به غلظت نسبی گاز بر حسب ولتاژ خروجی حسگر که متناسب با غلظت گاز است، انجام می‌شود. مقدار ولتاژ توسط یک میکروکنترلر خوانده می‌شود و در یک مقیاس بدون بعد بین ۰ تا ۱۰۲۳ نشان داده که معمولاً به‌عنوان تبدیل آنالوگ به دیجیتال (ADC) شناخته می‌شود. این محدوده در بین انواع حسگرهای MQ متفاوت است زیرا حسگرها دامنه تشخیصی متفاوتی دارند. قبل از شروع اندازه‌گیری‌ها، حسگرها طبق توصیه سازنده برای ۲۴ ساعت پیش‌گرم می‌شوند. در مطالعه اول از یک سامانه آرایه‌ای سه حسگری برای نمونه‌های خاک کلمبیا و اکوادور استفاده شد، که به‌منظور مطالعه الگوهای انتشار گاز آن‌ها در آزمایش‌های تنفسی القا شده با گلوکز D(+). به دو گروه تقسیم شدند. برای آزمایش دوم با استفاده از یک آرایه حسگر چندگانه در مجموعه‌ای متفاوت از نمونه‌های خاک کلمبیا، الگوهای انتشار گاز از خاک خشک شده با کاهش فعالیت میکروبی به‌دست آمد و نمونه‌ای دیگر با روغن دیزل آلوده شد. نتایج مطالعه دوم، زمانی که فعالیت میکروبی با خشک کردن خاک، کاهش یافت الگوی متفاوتی را از انتشار گاز در نمونه‌های آلوده

1- Norm Aroma Index

شد. انتخاب این عمق خاک به این دلیل بود که نفت بیش‌ترین تأثیر را در این عمق دارد. خاک برداشت شده سپس از الک عبور داده شد و پس از آن خواص فیزیکی و شیمیایی آن مورد بررسی قرار گرفت. خاک برای ارزیابی آلودگی با استفاده از سامانه بویایی ماشین به نسبت ۱/۳ در ظرف مورد آزمایش قرار گرفت (Li et al., 2020). رطوبت قابل قبول مزرعه‌ای (FC)، ۷۰ درصد است و در این تحقیق خاک در ۴ غلظت ۲۰۰ ppm، ۵۰۰ ppm، ۱۰۰۰ ppm و ۵۰۰۰ ppm به نفت خام آلوده شد که غلظت ۱۰۰۰ ppm حد مجاز بوده و ۵۰۰۰ ppm موجب تخریب خاک می‌گردد. به‌طور کلی یک نمونه هم بدون آلودگی به عنوان نمونه شاهد در نظر گرفته شد. همچنین از لحظه آلوده کردن خاک مورد نظر، که زمان صفر در نظر گرفته شد تا ۲۴ ساعت بعد، نمونه آزمایش مهر و موم شده نگهداری گردید تا به تعادل برسد و سپس آزمایش ادامه یافت. تعداد کل نمونه‌ها ۸ عدد بود، که در دمای محیط یعنی ۲۴ تا ۲۷ درجه سلسیوس نگهداری شدند.

نگهداری خاک آلوده

پس از آلوده کردن خاک، نمونه‌ها در یک کیسه‌ی تمیز پلاستیکی داخل یک ظرف درب دار قرار داده شدند. سپس، درب ظرف کاملاً بسته شده و در جای خشک، خنک و به دور از نور قرار گرفت. به منظور اطمینان از جذب کامل آلاینده توسط خاک، آزمایش‌های اولیه، ۲۴ ساعت پس از آلوده سازی انجام گرفت.

طراحی سامانه بین‌الکترونیک

سامانه بین‌الکترونیکی استفاده شده در این پژوهش با هشت حسگر طراحی و ساخته شد و قسمت‌های اصلی این سامانه شامل: سامانه تحصیل داده، محفظه نمونه، حسگرها، محفظه حسگرها که در این دستگاه با محفظه نمونه به صورت یکپارچه ساخته شد، پمپ هوا، منبع تغذیه و اتصالات و لوازم جانبی است (شکل ۱).

حسگرها

حسگرهای به کار رفته در این پژوهش از نوع حسگرهای نیمه‌هادی اکسید فلزی و از گروه حسگرهای الکتروشیمیایی هستند. هر کدام از هشت حسگر استفاده شده (MQ5، MQ136، MQ135، TGS813، TGS2610، MQ135،

به روغن دیزل نشان داد (Pineda & Perez, 2016). در مطالعه‌ای دیگر، خاک‌های آلوده به مواد نفتی از نظر ویژگی‌های فیزیکی و شیمیایی که از پنج پالایشگاه تهران، آبادان، اصفهان، شیراز و تبریز تهیه شده بودند مورد بررسی و مطالعه قرار گرفتند. خاک‌های آلوده شده به‌عنوان نمونه جمع‌آوری شد و سپس غلظت کل هیدروکربن‌های نفتی (TPH)، مقدار دی‌اکسیدکربن و فلزات سنگین شامل کادمیوم، کروم، سرب و نیکل اندازه‌گیری و از نظر آماری مورد بررسی قرار گرفتند. در این تحقیق برای بررسی خواص فیزیکی و شیمیایی که شامل بافت خاک، میزان آهک یا همان کربنات کلسیم، مقدار کربن آلی و میزان فسفر است به ترتیب از روش‌های هیدرومتری، تیتراسیون، والکی‌وبلاک و اولسون استفاده و EC، PH و میزان کل نیتروژن خاک به روش کدال اندازه‌گیری شد. برای اندازه‌گیری غلظت فلزات سنگین از اسید نیتریک و اسید پرکلریک و دستگاه جذب اتمی، و برای استخراج کل هیدروکربن‌های خاک (TPH) از دستگاهی به نام سوکسله و محلول ان‌هگزان، استفاده شد. مقادیر به‌دست‌آمده برای هر کدام از این ویژگی‌های فیزیکی و به‌خصوص شیمیایی خاک، بیش‌تر از مقادیر استاندارد تعریف شده بوده و با بررسی میزان غلظت فلزات مشخص شد که دو پالایشگاه تهران و آبادان بیشترین آلودگی را دارند (Asadabadi et al., 2015). در نهایت با توجه به مطالعات فراوانی که در زمینه تشخیص آلودگی خاک انجام شد، این‌طور نتیجه‌گیری شد که هنوز خلاء یک روش کارآمد، مقرون به صرفه و سریع در این زمینه احساس می‌شود. هدف از انجام این تحقیق، آزمایش نمونه خاک‌های آلوده شده به نفت خام با غلظت‌های مختلف با استفاده از سامانه بین‌الکترونیکی است که موجب تخریب یا عدم تخریب خاک می‌شود. تا نشان‌داده شود که این روش آزمایش، برای تشخیص آلودگی خاک مؤثر است.

مواد و روش‌ها

تهیه و آماده‌سازی نمونه

نمونه خاک مورد استفاده، خاک مزرعه دانشکدگان کشاورزی و منابع طبیعی دانشگاه تهران و خاک باغچه است که از عمق صفر تا ۲۰ سانتی‌متری نمونه برداری

اطلاعات آن‌ها استخراج شده و در جدول ۱ آورده شده است.



شکل ۱- نمای کلی از محفظه یکپارچه حسگرها و نمونه

MQ3، MQ322 و MQ138) به ترکیبات خاصی از نمونه‌ها واکنش نشان می‌دهند. انتخاب حسگرها بر اساس نتایج به دست آمده از دستگاه کروماتوگرافی گازی که برای استخراج ترکیبات خاک آلوده به نفت خام استفاده می‌شود، است. بر اساس نتایج، ترکیباتی که استخراج شده‌اند شامل زنجیره کربنی C_{۱۰}-C_{۴۵} بوده و موادی مانند بنزین، نفت سفید، گازوئیل، متان و ... را نیز دربردارند (Sharifi Hosseini et Aghamohammadi et al., 2017; al., 2009). لازم به ذکر است از حسگرهایی استفاده شد که حساسیت بالایی را به این مواد نشان دادند تا نتایج مطلوبی حاصل گردد. از مزایای این حسگرها می‌توان به مقرون‌به‌صرفه بودن، حساسیت کم به رطوبت، راه‌اندازی آسان، پایداری شیمیایی و حساسیت زیاد به رایحه‌های تشخیص‌دهنده اشاره کرد. مشخصات این حسگرها از برگه

جدول ۱- حسگرهای استفاده شده

محدوده تشخیص (ppm)	گاز تشخیص‌دهنده	حسگر
۲۰۰ - ۱۰۰۰۰	LPG، گاز طبیعی و گازوئیل	MQ5
۱ - ۲۰۰۰	هیدروژن سولفید، آمونیوم، کربن منواکسید	MQ136
۵۰۰ - ۱۰۰۰۰	متان، پروپان، بوتان	TGS813
۵۰۰ - ۱۰۰۰۰	LPG، بوتان، پروپان	TGS2610
۱۰ - ۱۰۰۰	بنزن، الکل، آمونیاک، دی‌اکسیدکربن	MQ135
۵۰ - ۱۰۰۰	بنزین و اتانول	MQ3
۲۰۰ - ۵۰۰۰	حساسیت بالا به گازهای احتراق مانند کربن منواکسید و گازوئیل	TGS822
۵ - ۵۰۰	تولوئن، استون، الکل و هیدروژن	MQ138

استفاده از روش‌های LDA^۱، PCA^۲ و SVM^۳ و الگوریتم‌های یادگیری ماشین، مورد پردازش قرار گرفتند. سپس میزان دقت روش‌ها مورد بررسی قرار گرفت. در این بخش نتایج روش‌های طبقه‌بندی و تفکیک با یکدیگر مقایسه شدند.

در این پژوهش اجرای روش‌های LDA، PCA و SVM توسط نرم‌افزار 64- (The Unscrambler X 10.4) (bit) بود و الگوریتم‌های یادگیری ماشین^۴ نیز با زبان برنامه‌نویسی پایتون پیاده گردیدند.

سامانه تحویل داده

برای دریافت پاسخ حسگرها، از مدار آردوینو استفاده گردید. این مدار مجموعه نرم‌افزاری - سخت‌افزاری بر پایه میکروکنترلرهای Atmel است. راحتی کار با نرم‌افزار آردوینو و لایه باز بودن بردهای این مجموعه دو ویژگی مهم آردوینو هستند. به علاوه کتابخانه‌های کد نویسی فراوان هماهنگ با قطعات و حسگرهای الکترونیکی، این برد را بسیار پر کاربرد کرده‌است. برد آردوینو مورد استفاده در این سامانه از نوع MEGA R3-2560 است.

ارزیابی سامانه

بعد از انجام آزمایش‌ها و به دست آوردن داده‌های مناسب، نمونه‌ها برای طبقه‌بندی و تفکیک آماده شده و سپس با

1- Linear discriminant analysis
2- Principal component analysis
3- Support vector machine
4- Machine learning algorithms

روش تحلیل مؤلفه‌های اصلی (PCA)

اصلی‌ترین کاربرد تحلیل مؤلفه‌های اصلی در تشخیص الگو، کاهش ویژگی است. تحلیل مؤلفه‌های اصلی همان‌طور که از نامش مشخص است، مؤلفه‌های اصلی را تشخیص می‌دهد و به ما این امکان را می‌دهد تا به جای این‌که تمامی ویژگی‌ها را مورد بررسی قرار دهیم، ویژگی‌هایی مورد بررسی قرار داده می‌شوند که ارزش بیشتری دارند. از دیدی دیگر، PCA یک نگاشت خطی متعامد است که داده را به دستگاه مختصات جدید می‌برد؛ به صورتی که در این نگاشت بزرگ‌ترین واریانس داده‌ها بر روی اولین محور مختصات (PC1)، و دومین واریانس از نظر بزرگی بر روی دومین محور مختصات (PC2) قرار می‌گیرد و به همین ترتیب برای بقیه متغیرها ادامه می‌یابد. (Mortazavi et al., 2020).

روش تحلیل تفکیک خطی (LDA)

روش تحلیل تفکیک خطی، روشی آماری است که از جمله در بازشناخت الگوها و یادگیری ماشین برای یافتن ترکیب خطی خصوصیتی که به بهترین شکل ممکن دو یا چند طبقه از اشیاء را از هم جدا می‌کند، استفاده می‌شود. بررسی تفکیک خطی با تحلیل مؤلفه‌های اصلی شباهت دارد. هر دو روش آماری، برای ترکیب خطی متغیرها به شکلی که داده را به بهترین نحو توضیح دهد، به کار می‌روند. یک کاربرد عمده در هر دو روش، کم کردن تعداد بعدهای داده است. با این حال این روش‌ها تفاوت عمده‌ای با هم دارند: در تحلیل مؤلفه‌های اصلی تفاوت میان طبقه‌ها نادیده گرفته می‌شود، در صورتی که در تحلیل تشخیص خطی، تفاوت میان طبقه‌ها مدل‌سازی می‌شود (Perriere & Thioulouse, 2003).

روش تحلیل ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبان، روش یادگیری با نظارت است که در این الگوریتم دسته‌بندی، دو ابر صفحه در مرز بین دو طبقه داده‌ها قرار داده می‌شود. این دو ابر صفحه تا زمانی که به داده‌ها برخورد کنند، از یکدیگر دور می‌شوند. این عمل جداسازی خطی نام دارد و بردار پشتیبان به نمونه‌هایی گفته می‌شود که بر روی این دو ابر صفحه قرار می‌گیرند، خطای طبقه‌بندی، زمانی کمتر می‌شود که فاصله بین دو ابر صفحه بیشتر شود (Mousavi et al., 2021).

روش آزمایش فاکتوریل در قالب طرح بلوک تصادفی با دو فاکتور اصلی

در این پژوهش از روشی استفاده شد که با کمک آن بتوان حسگرها را به‌طور جداگانه مورد بررسی قرار داده و دقت هر کدام را با استفاده از الگوریتم‌های یادگیری ماشین به دست آورد. روش‌های یادگیری ماشین، روش‌هایی با نظارت هستند. به عبارت دیگر، روشی است که خروجی مطلوب داده‌ها موجود بوده و با کمک این خروجی‌ها و ورودی‌های متناظر با آن‌ها، رابطه‌ای را برای مدل‌سازی می‌توان ایجاد کرد و برای مسائل رگرسیون و طبقه‌بندی به کار گرفته می‌شود. آزمایش فاکتوریل با دو فاکتور اصلی روز و زمان انجام گرفت. آزمایش هر نمونه سه مرتبه تکرار شد و برای تحلیل، ۸۰٪ داده‌ها برای آموزش و ۲۰٪ برای آزمون انتخاب شدند. همچنین از رگرسیون درخت تصمیم، رگرسیون بردار پشتیبان، رگرسیون گرادیان تقویت‌یافته و رگرسیون جنگل تصادفی استفاده شد.

نتایج و بحث

نتایج حاصل از بینی الکترونیک و جمع‌آوری داده‌ها

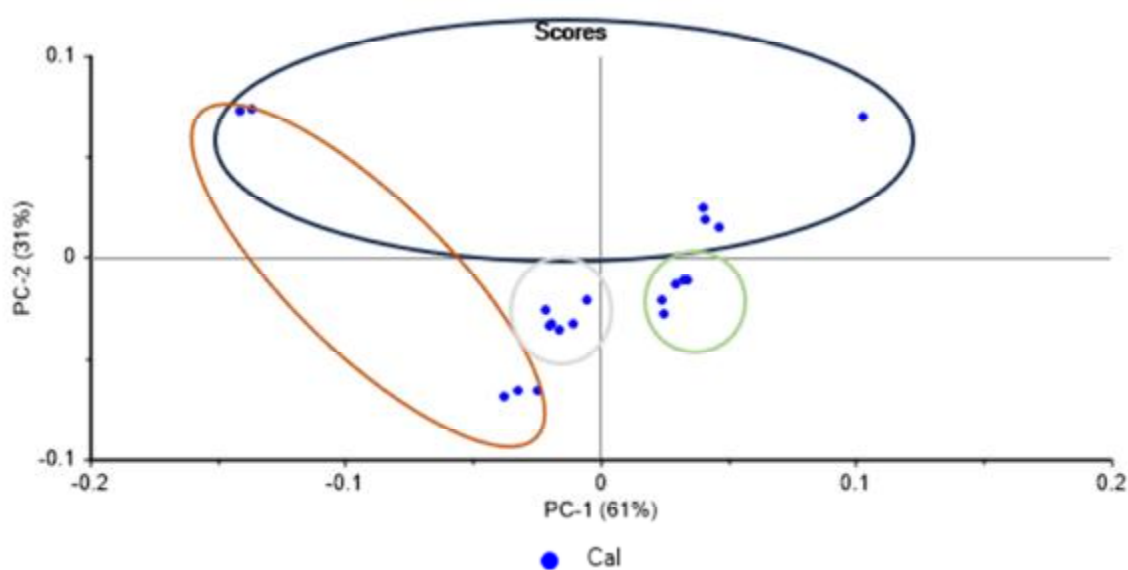
به‌طور فرآیند اندازه‌گیری و جمع‌آوری سیگنال‌ها، در سه مرحله تصحیح خط مبنا، فرایند تزریق گاز فضای بالای نمونه خاک تیمار شده، به محفظه حسگرها و پاک‌سازی هوای محفظه حسگرها زمان‌بندی شد. در نتیجه با توجه به نتایج آزمون و خطا در زمان‌بندی‌های متفاوت، مدت زمان هر کدام از این مراحل به ترتیب، ۱۰۰، ۱۲۰ و ۱۰۰ ثانیه انتخاب گردید.

داده‌های استخراجی که در واقع همان تغییر ولتاژ خروجی حسگرها است با استفاده از روش تفاضلی، نرمال شدند. نتایج به دست آمده از داده‌برداری‌ها با استفاده از سامانه بینی الکترونیک، نشان داد که پاسخ حسگرهای استفاده شده در این تحقیق به خوبی توانسته‌اند تمام نمونه‌ها را تفکیک کنند.

در این تحقیق، آزمایشات با استفاده از یک سامانه بینی الکترونیک بر روی ۴ نمونه خاک با غلظت‌های آلودگی متفاوت در دو نوبت به فاصله ۱۴ روز و یک نمونه شاهد صورت گرفته و طبقه‌بندی شدند.

نتایج روش PCA

از روش تحلیل مؤلفه‌های اصلی برای طبقه‌بندی غلظت‌های آلوده خاک و در نتیجه تشخیص آلودگی خاک استفاده گردید. در شکل ۲، نتایج حاصل از نمودار امتیازی برای طبقه‌بندی کل داده‌ها که شامل ۳ تکرار برای هر نمونه در دو روز اول و چهاردهم و نمونه شاهد در روز اول نشان داده شده است. با توجه به نمودار امتیازی مشخص شد که واریانس استخراجی به وسیله دو مؤلفه PC1 و PC2 که همان دو مؤلفه اول و بالاترین مقدار



شکل ۲- نمودار امتیازی تحلیل مؤلفه‌های اصلی برای کل داده‌ها در ۴ غلظت آلوده

جدول ۲- ماتریس آشفتگی حاصل از PCA

واقعی / پیش‌بینی	ppm۲۰۰	ppm۵۰۰	ppm۱۰۰۰	ppm۵۰۰۰
۲۰۰ppm	۵	۱	۰	۰
۵۰۰ppm	۱	۵	۰	۰
۱۰۰۰ppm	۰	۰	۶	۰
۵۰۰۰ppm	۰	۰	۰	۶

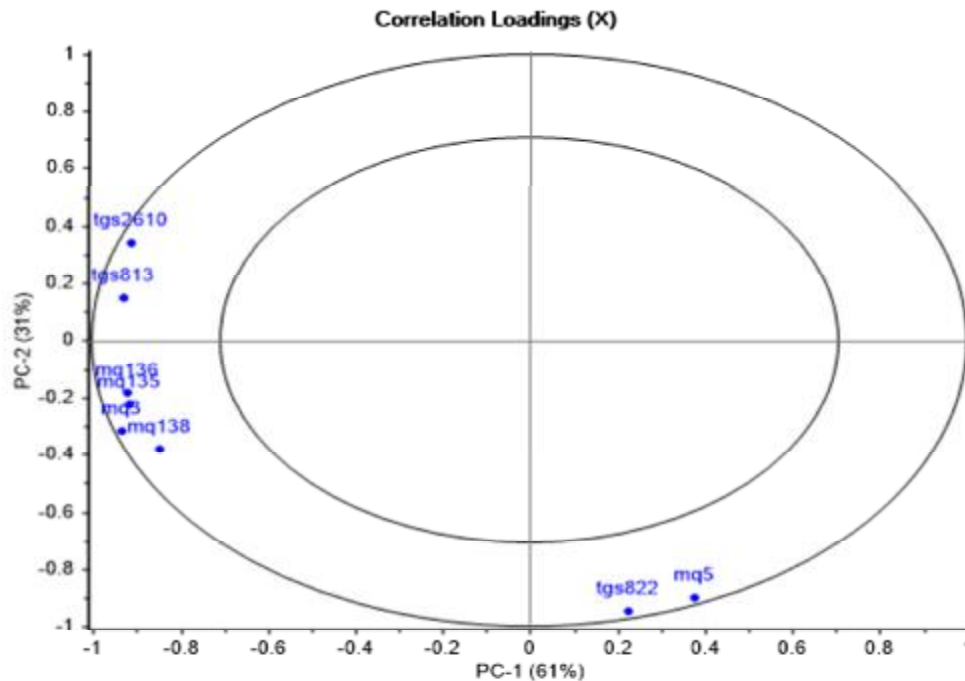
دایره بیرونی و مقدار ۱ و ۱- محورهای نمودار داشته باشد، نشان دهنده این است که آن حسگر نسبت به سایر حسگرها تأثیر بیش‌تری در تفکیک طبقه‌ها از یکدیگر با توجه به آن مؤلفه اصلی دارد.

نمودار بارگذاری در شکل ۳ به وضوح نشان می‌دهد که برای تفکیک غلظت‌های متفاوت نفت در خاک، حسگرهای MQ5، TGS822 و MQ3 بیش‌ترین ضریب

یکی دیگر از نمودارهایی که در تحلیل PCA می‌توان استخراج کرد و به کار گرفت نمودار بارگذاری است. این نمودار با استفاده از ضریب وزنی متغیرها، توزیع آن‌ها را ترسیم می‌کند و بیان‌کننده تأثیر هر حسگر بر روی تفکیک طبقه‌های موجود در راستای مؤلفه اصلی است. هرچه قدر مقدار بارگذاری حسگری روی یک مؤلفه اصلی بیش‌تر باشد و به عبارت دیگر، نزدیکی بیش‌تری به

هم‌چنین حسگر MQ138 ضریب بارگذاری کمتری نسبت به حسگرهای دیگر دارد، اما همچنان هم‌بستگی منفی بالایی با محور اول دارد.

بارگذاری و در نتیجه بیش‌ترین تأثیر را داشته‌اند. حسگرهای بعدی که به فاصله خیلی کمی قرار گرفته‌اند عبارتند از: MQ135، MQ136، TGS2610 و TGS813.



شکل ۳- نمودار بارگذاری برای ویژگی‌های استخراج شده

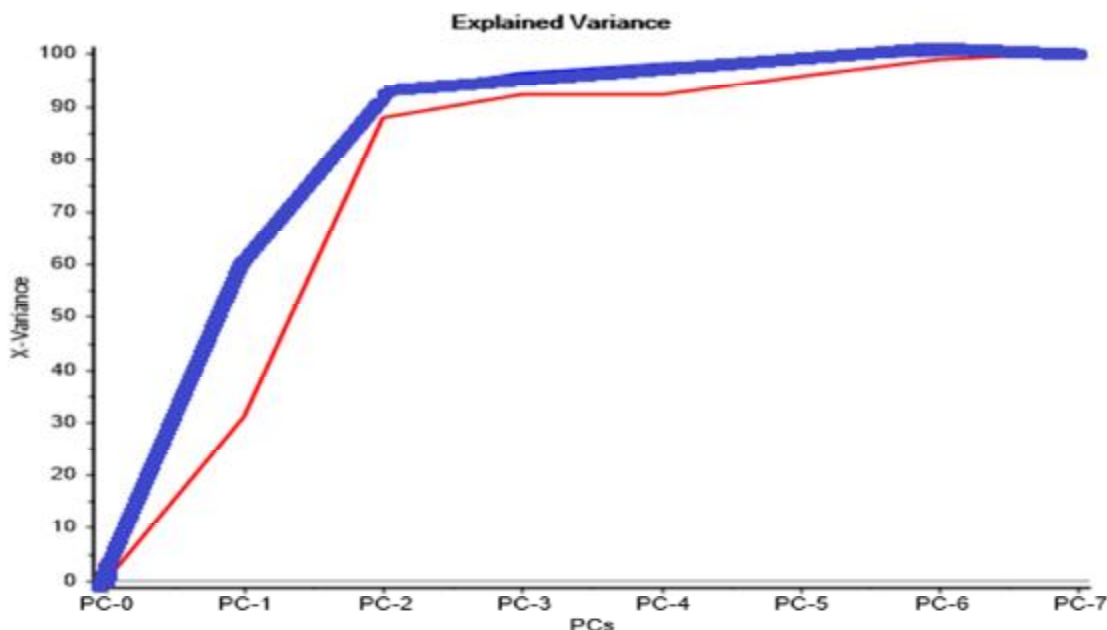
شده) نشان می‌دهد که دو مؤلفه اول، ۹۲٪ واریانس و پنج مؤلفه بعدی حدود ۸٪ واریانس را تشکیل می‌دهند. به همین علت است که دو مؤلفه اول سهم بیشتری در توضیح واریانس دارند و انتخاب می‌شوند و مرحله مدل‌سازی آغاز می‌شود.

نتایج روش LDA

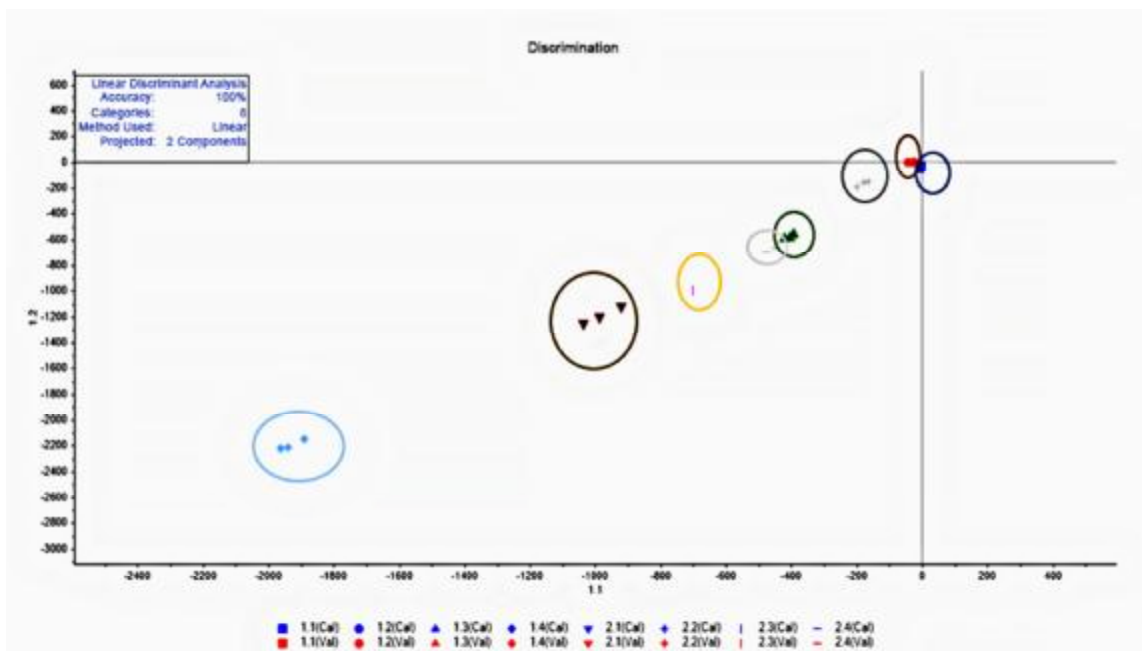
نتایج به‌دست آمده از روش تحلیل تفکیک خطی با درصد تفکیک ۱۰۰ درصد و نمودار LDA در شکل ۵ نشان داده شده است. همان‌طور که مشخص است، هدف تشخیص غلظت‌های متفاوت است که در نمودار مذکور غلظت‌های ۲۰۰ ppm، ۵۰۰ ppm، ۱۰۰۰ ppm و ۵۰۰۰ ppm در روز اول به ترتیب با ۱/۱، ۱/۲، ۱/۳ و ۱/۴ نمایش داده شده و برای روز چهاردهم همین غلظت‌ها به ترتیب با ۲/۱، ۲/۲، ۲/۳ و ۲/۴ نمایش داده شده است.

همچنین موضوع مهمی که در رابطه با نمودار بارگذاری لازم به ذکر است، متغیرهایی است که روی یک نمودار در جهت هم و به فاصله نزدیک از یکدیگر قرار دارند. این متغیرها دارای همبستگی مثبت هستند و می‌توان گفت تقریباً اثر مشابهی در تفکیک طبقه‌ها دارند. در نتیجه حسگرهایی که دارای همبستگی مثبتی هستند می‌توان برای کاهش ابعاد داده‌ها و کمتر کردن تعداد حسگرها، حسگر با مقدار حساسیت کمتر را حذف کرد تا در کاهش هزینه‌ها و کاهش زمان تحلیل و محاسبات هم مؤثر واقع گردد. با این تفاسیر و با توجه به نمودار بارگذاری می‌توان گفت که دو حسگر MQ135 و MQ136 دارای همبستگی مثبت هستند و برای سادگی کار می‌توان یکی از آن دو را حذف نمود.

در تحلیل PCA علاوه بر نمودار بارگذاری و امتیازی، نمودار دیگری هم می‌توان استخراج کرد که در شکل ۴ آمده است. این نمودار (واریانس توضیح داده



شکل ۴- نمودار واریانس توضیح داده شده



شکل ۵- نمودار تحلیل تفکیک خطی برای داده‌ها در ۴ غلظت

نتایج روش SVM

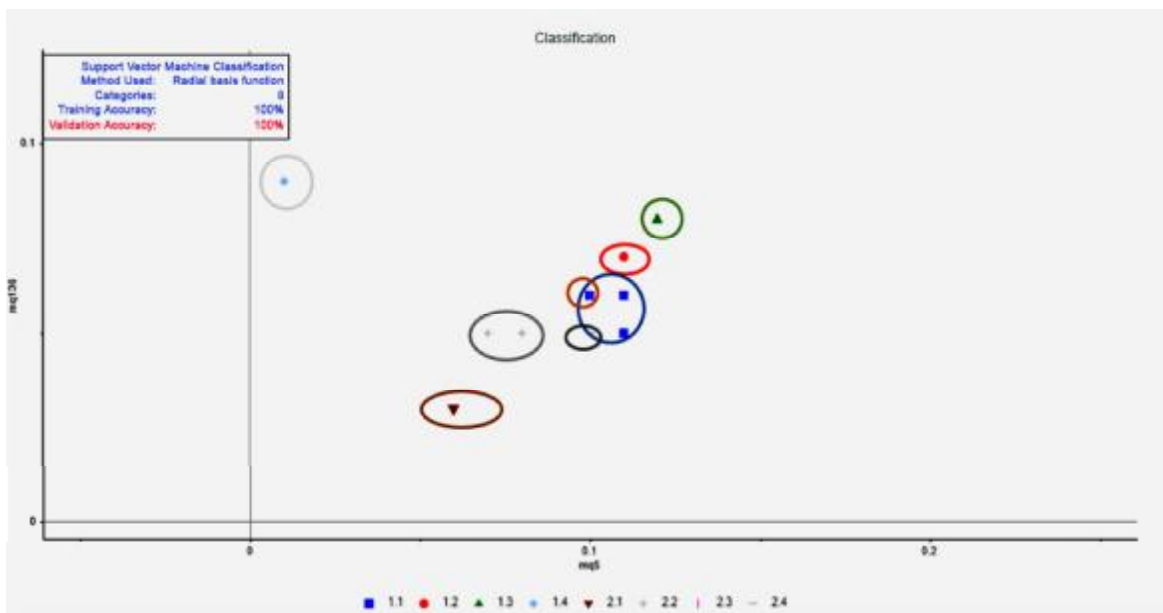
گروه‌های نمونه، همانند روش LDA، به ۸ گروه تقسیم شدند مطابق روش ذکر شده در هر مجموعه، ۸۵ درصد داده‌ها برای آموزش و ۱۵ درصد باقی مانده برای ارزیابی مدل استفاده گردید. مطابق شکل ۶ عملکرد این روش به خوبی بررسی شده و دقت طبقه‌بندی برای آموزش و

ارزیابی ۱۰۰٪ به دست آمد.

در نهایت بالاترین دقت و بهترین تفکیک با استفاده از روش‌های SVM و LDA انجام شدند. باتوجه به شکل ۵ مشاهده می‌شود که در روش LDA دو غلظت ۱/۱ و ۱/۲ که مربوط به گروه‌های آلودگی خاک با غلظت‌های ۲۰۰ و ۵۰۰ ppm هستند، نزدیک‌ترین فاصله را به هم

این الگوها به وضوح رفتار ثابت کل سامانه را نشان می‌دهند (Pineda & Perez, 2016). اما در پژوهش انجام شده تنها وجود آلودگی نفتی حائز اهمیت نیست و تشخیص تفاوت میان غلظت‌ها هدف اصلی است. به همین علت استفاده از روش‌های تجزیه و تحلیل آماری به علت مشابه بودن نمودارهای خروجی امری ضروری بود و در این پژوهش با استفاده از این روش‌ها نتایج قابل قبولی حاصل شد.

دارند که به علت نزدیک بودن مقدار آلودگی نفتی این دو گروه نسبت به یکدیگر است. بدیهی است که هر چه غلظت‌ها کمتر و به یکدیگر نزدیک‌تر باشند، تشخیص آن‌ها سخت‌تر و همپوشانی آن‌ها در نمودار بیشتر می‌شود. همچنین برای تفاوت میان خاک آلوده شده به مواد نفتی و خاک بدون تیمار یا تیمارهای متفاوت تنها استفاده از نمودارهای خروجی حاصل از سامانه بین‌الکترونیک کافی است و نیازی به تجزیه و تحلیل‌های آماری نیست، زیرا



شکل ۶- نمودار ماشین بردار پشتیبان برای داده‌ها در ۴ غلظت

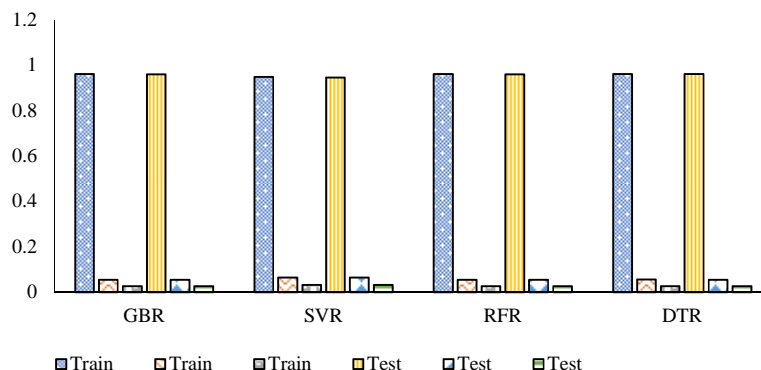
می‌دهد و صحیح بودن این مقدار بیانگر قابل تشخیص و تفکیک‌پذیر بودن هر طبقه از هم توسط سامانه بین‌الکترونیک است.

بر این اساس، حسگرهای انتخابی ناشی از نتایج به دست آمده از سامانه کروماتوگرافی گازی بر روی خاک آلوده به نفت خام (Aghamohammadi *et al.*, 2017; Sharifi Hosseini *et al.*, 2009)، از حساسیت قابل قبولی برخوردار بودند و در مجموع سامانه بین‌الکترونیک را می‌توان به عنوان روشی کم هزینه‌تر و آسان‌تر در این موضوع، جایگزین روش‌های دیگر نمود. همچنین در پژوهش (Luan *et al.*, 2023) نتایج سامانه بین‌الکترونیک برای تشخیص هیدروکربن‌های خاک تنها با استفاده از دو الگوریتم ماشین بردار پشتیبان و جنگل تصادفی مورد بررسی قرار گرفت، ولی در پژوهش انجام

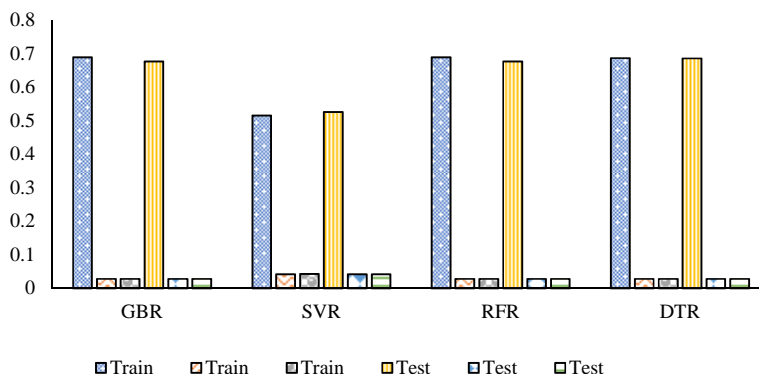
نتایج تحلیل آماری داده‌ها با روش آزمایش فاکتوریل

بعد از به دست آمدن نتایج، مشخص شد که حسگر MQ3 مجدداً بالاترین دقت و کم‌ترین ضریب خطا را داشته است شکل ۷ و چهار الگوریتم رگرسیون درخت تصمیم (DTR)، رگرسیون بردار پشتیبان (SVR)، رگرسیون گرادینت تقویت‌یافته (GBR) و رگرسیون جنگل تصادفی (RFR) با بهترین عملکرد با میانگین دقت ۹۵/۵ درصد نرخ تشخیص سطوح آلودگی را نشان می‌دهند. حسگر MQ136 کمترین دقت و حساسیت را در بین حسگرها داشته است شکل ۸. همچنین جدول ۳ معنی‌دار بودن (مطابق ANOVA^۱) اختلاف بین پاسخ‌های هر حسگر به نمونه‌های متفاوت در روز اول و چهاردهم آزمایش را نشان

شده علاوه بر مقایسه نتایج به دست آمده از سامانه الگوریتم‌های یادگیری ماشین بیشتری استفاده شد تا بین‌الکترونیک و سامانه کروماتوگرافی گازی از اطمینان کامل از صحت نتایج به دست آمده حاصل شود.



شکل ۷- نتایج الگوریتم‌های یادگیری ماشین برای حسگر MQ3



شکل ۸- نتایج الگوریتم‌های یادگیری ماشین برای حسگر MQ136

جدول ۳- پاسخ‌های حسگر MQ3 به نمونه‌های مختلف

۶	۵	۴	۳	۲	۱	۰	MQ3
reject	upper	lower	p-adj	meandiff	group۲	group۱	۰
True	-/۰.۵۷۹	-/۰.۱۱۹۳	۰/۰	-/۰.۸۸۶	Day_1 & ppm_200	Day_1 & ppm_1000	۱
True	-/۰.۶۷۵	-/۰.۱۲۸۶	۰/۰	۰/۰.۹۸	Day_1 & ppm_500	Day_1 & ppm_1000	۲
True	۰/۱.۶۱۳	۰/۰.۹۹۸	۰/۰	۰/۱.۳۵	Day_1 & ppm_5000	Day_1 & ppm_1000	۳
False	۰/۰.۱۶۷	-/۰.۴۹۹	۰/۰.۸۶۳۲	-/۰.۱۴۵	Day_14 & ppm_1000	Day_1 & ppm_1000	۴
True	-/۰.۲۴۸	۰/۰.۸۴۹	۰/۰	۰/۰.۱۵۴	Day_14 & ppm_200	Day_1 & ppm_1000	۵
True	-/۰.۵۲۵	-/۰.۱۱۶۱	۰/۰	-/۰.۰۸۳	Day_14 & ppm_500	Day_1 & ppm_1000	۶
True	-/۰.۴۶۴۵	-/۰.۵۲۷۴	۰/۰	-/۰.۴۹۵	Day_14 & ppm_5000	Day_1 & ppm_1000	۷
False	۰/۰.۲۰۷	-/۰.۳۹۴	۰/۰.۹۸۱۲	-/۰.۰۹۴	Day_1 & ppm_500	Day_1 & ppm_200	۸
True	۰/۰.۲۴۹۴	۰/۱.۸۸۹	۰/۰	۰/۰.۲۱۹۱	Day_1 & ppm_5000	Day_1 & ppm_200	۹
True	۰/۰.۱۴۹	۰/۰.۴۴۱	۰/۰	۰/۰.۷۴۵	Day_14 & ppm_1000	Day_1 & ppm_200	۱۰
True	۰/۰.۶۳۳	۰/۰.۰۱۶	۰/۰.۳۰۹	۰/۰.۳۲۵	Day_14 & ppm_200	Day_1 & ppm_200	۱۱

(معنی‌دار بودن اختلاف آن‌ها به صورت True و معنی‌دار نبودن اختلاف به صورت False بیان شده است.)

ادامه جدول ۳ -

۶	۵	۴	۳	۲	۱	۰	MQ3
False	./۰۳۵۷	./۰۲۷	./۹۹۹۹	./۰۰۴۳	Day_14 & ppm_500	Day_1 & ppm_200	۱۲
True	./۰۳۷۶۳	./۰۴۲۸	./۰	./۰۰۴۸	Day_14 & ppm_5000	Day_1 & ppm_200	۱۳
True	./۰۲۵۸۸	./۰۹۸۴	./۰	./۰۲۲۸۶	Day_1 & ppm_5000	Day_1 & ppm_500	۱۴
True	./۰۱۱۴۲	./۰۵۳۷	./۰	./۰۸۳۶	Day_14 & ppm_1000	Day_1 & ppm_500	۱۵
True	./۰۷۲۷	./۰۱۱۲	./۰۰۱	./۰۴۱۹	Day_14 & ppm_200	Day_1 & ppm_500	۱۶
False	./۰۴۵	./۰۱۷۵	./۰۸۸۵۳	./۰۱۳۸	Day_14 & ppm_500	Day_1 & ppm_500	۱۷
True	./۰۳۶۷	./۰۴۲۸۹	./۰	./۰۳۷۰۱	Day_14 & ppm_5000	Day_1 & ppm_500	۱۸
True	./۰۱۱۴۲	./۰۱۷۵۱	./۰	./۰۱۴۴۶	Day_14 & ppm_1000	Day_1 & ppm_5000	۱۹
True	./۰۱۵۵۸	./۰۲۱۷۶	./۰	./۰۱۸۶۷	Day_14 & ppm_200	Day_1 & ppm_5000	۲۰
True	./۰۱۸۳۴	./۰۲۴۶۲	./۰	./۰۲۱۴۸	Day_14 & ppm_500	Day_1 & ppm_5000	۲۱
True	./۰۵۹۵۴	./۰۶۵۷۶	./۰	./۰۶۲۶۵	Day_14 & ppm_5000	Day_1 & ppm_5000	۲۲
True	./۰۰۳۱۱	./۰۰۷۳	./۰۰۱	./۰۰۴۲	Day_14 & ppm_200	Day_14 & ppm_1000	۲۳
True	./۰۰۳۸۷	./۰۱۰۱۷	./۰	./۰۰۷۰۲	Day_14 & ppm_500	Day_14 & ppm_1000	۲۴
True	./۰۴۵۰۷	./۰۵۱۳	./۰	./۰۴۸۱۹	Day_14 & ppm_5000	Day_14 & ppm_1000	۲۵
False	./۰۰۳۲	./۰۰۶۰۱	./۰۱۳۲۴	./۰۰۲۸۱	Day_14 & ppm_500	Day_14 & ppm_200	۲۶
True	./۰۴۰۸۲	./۰۴۷۱۵	./۰	./۰۴۳۹۸	Day_14 & ppm_5000	Day_14 & ppm_200	۲۷
True	./۰۳۷۹۶	./۰۴۴۳۸	./۰	./۰۴۱۱۷	Day_14 & ppm_5000	Day_14 & ppm_500	۲۸

and technology,[online], 22(2): 93. (In Persian)

Alipour, A. Z., Malekian, M., & Soleimani, M. (2016). Soil pollution of petroleum hydrocarbons and heavy metals in five Iranian oil refineries.

Douglas, R. K., Nawar, S., Alamar, M. C., Coulon, F., & Mouazen, A. M. (2017). Almost 25 years of chromatographic and spectroscopic analytical method development for petroleum hydrocarbons analysis in soil and sediment: state-of-the-art, progress and trends. *Critical Reviews in Environmental Science and Technology*, 47(16): 1497-1527.

Heidarpoor, M., & Oliaei, M. (2013). Oil contamination propagation patterns in soils. *Modares Civil Engineering journal*, 13(2): 39-51. (In Persian).

Kang, C. U., Kim, D. H., Khan, M. A., Kumar, R., Ji, S. E., Choi, K. W., & Jeon, B. H. (2020). Pyrolytic remediation of crude oil-contaminated soil. *Science of the Total Environment*, 713: 136498.

Li, F., Zhang, Y., Wang, S., Li, G., Yue, X., Zhong, D., & Shen, K. (2020). Insight into ex-situ thermal desorption of soils contaminated with petroleum via carbon number-based fraction approach. *Chemical Engineering Journal*, 385: 123946.

Luan, X., Kong, C., Yao, Z., Sun, Y., & Chang, Z. (2023). Hierarchical electronic nose detection and assessment technology for the

نتیجه‌گیری

در این پژوهش از یک سامانه بین‌الکترونیک بر پایه ۸ حسگر نیمه‌هادی اکسید فلزی با الگوریتم‌های تجزیه و تحلیل داده PCA، LDA، SVM و الگوریتم‌های یادگیری ماشین جهت ارزیابی آلودگی نفتی خاک استفاده گردید. همه روش‌ها توانستند با دقت بالایی ارزیابی آلودگی نفت خاک را انجام دهند و طبق این مطالعه می‌توان بیان نمود بین‌الکترونیک ابزاری مناسب برای تشخیص و تعیین غلظت‌های متفاوت آلودگی نفتی در خاک است. همچنین در این سامانه حسگرهای MQ5، MQ3 و TGS822 و بهترین عملکرد را داشته و حسگرهای MQ136 و MQ138 کمترین حساسیت را به ترکیبات نفتی و ضریب بارگذاری پایین‌تری را داشته‌اند.

منابع

- Abbasi Maedeh, P., Nasrabadi, T., Wu, W., & Al Dianty, M. (2017). Evaluation of oil pollution dispersion in an unsaturated sandy soil environment. *Pollution*, 3(4): 701-711.
- Aghamohammadi, R., Siahcheshm, K., Zarrini, G., & Kadkhodaie, A. (2020). Bacterial Bioremediation of Surficial Oil-Contaminated Soil in Tabriz Refinery Campus. *Journal of Environmental Science*

- reusage of land contaminated with petroleum hydrocarbons. *Sensors and Actuators B: Chemical*, 390: 133940.
- Mascini, M., Ram, M. K., Bhethanabotla, V. R., White, B. J., & Bogomolova, A. (2018). *Sensors for chemical and biological applications*. CRC press.
- Mortazavi, N., Mohtasebi, S., Soltani, M., Dehghanbanadaky, M. 2020. Estrus Detection of Cows from its milk using an Electronic Nose System. M.Sc. thesis, Department of Agricultural Machinery, University of Tehran. (In Persian).
- Mortazavi, S. N., Mohtasebi, S., Soltani, M., Deghan, M., & Mohtasebi, S. M. (2023). Feasibility Study of Dairy Cow Estrus detection based on milk using electronic nose system. *Journal of Researches in Mechanics of Agricultural Machinery*, 12(1): 83-94. (In Persian).
- Mousavi, S. M., Khoshkam, M., & Feizi, J. (2021). Comparison of metabolic profile in different saffron samples based on their geographical origin using gas chromatography-mass spectroscopy techniques (GC-MS). (In Persian).
- Perriere, G., & Thioulouse, J. (2003). Use of correspondence discriminant analysis to predict the subcellular location of bacterial proteins. *Computer Methods and Programs in Biomedicine*, 70(2): 99-105.
- Pineda, D. M., & Pérez, J. C. (2017). SENose: An under U \$50 electronic nose for the monitoring of soil gas emissions. *Computers and Electronics in Agriculture*, 133: 15-21.
- Ramesh, B., Mohtasebi, S. S., & Rafiee, S. (2019). Classification of different Iranian rice varieties and frauded rice based on volatile compounds detected by electronic nose method. *Iranian Journal of Biosystems Engineering*, 50(3): 595-606. (In Persian).
- Sharifi Hosseini, S., Shahbazi, A., Yazdipour, A., & Kamranfar, I. (2009). The effect of Agricultural Fertilizers on Bioremediation of a Crude-Oil Polluted. *Water and Soil*, 23(3). (In Persian).
- Tan, J., & Xu, J. (2020). Applications of electronic nose (e-nose) and electronic tongue (e-tongue) in food quality-related properties determination: A review. *Artificial Intelligence in Agriculture*, 4: 104-115.
- Wilson, A. D. (2013). Diverse applications of electronic-nose technologies in agriculture and forestry. *Sensors*, 13(2): 2295-2348.