

ORIGINAL RESEARCH PAPER

Prediction of Mineral Oil Concentrations Using Fourier Transform Infrared (FTIR) and Modeling Methods

Fateme Baghdadi¹, Rezvan Zendehei¹, Zahra Panjali², Alireza Hajighasemkhan^{1,*}

¹Department of Occupational Health and Safety Engineering, Faculty of Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran

²Department of Occupational Health and Safety Engineering, Faculty of Health, Islamic Azad University of Medical Sciences, Tehran, Iran

Received: 27-8-2024

Accepted: 3-5-2026

ABSTRACT

Introduction: Mineral oil, a key component of metalworking fluids, is a complex mixture that generates aerosols during industrial processes, posing significant respiratory health risks such as laryngeal cancer, asthma, and lung cancer. The NIOSH 5026 method uses Fourier Transform Infrared Spectroscopy (FTIR) to assess exposure to mineral oils. However, the diverse and complex compositions of mineral oils cause significant spectral interferences. Partial Least Squares (PLS) and Artificial Neural Networks (ANN) are advanced modeling methods used to address these interferences without manual intervention. This study aimed to predict mineral oil concentrations in an automotive industry using FTIR and modeling methods.

Material and Methods: FTIR spectral data (1500–4000 cm⁻¹) were recorded across 701 wave numbers and analyzed using PLS and ANN models. Input (matrix X) consisted of FTIR data, while output (matrix Y) represented mineral oil concentrations. Model performance was evaluated using Root Mean Square Error (RMSEp).

Results: The ANN model significantly outperformed the PLS model. The overall RMSEp for ANN was 0.0036, compared to 5.01 for PLS. ANN achieved a regression of 0.997 in the test set, with an average error percentage of 3.01%, while PLS yielded an error of 4.792. ANN modeling used 15% of data for validation and required fewer than 11 hidden layers to achieve optimal performance.

Conclusion: ANN modeling effectively predicted mineral oil concentrations despite spectral interferences, outperforming PLS in accuracy and error reduction. Both methods are viable for evaluating mineral oil exposure, but ANN offers superior predictive capabilities.

Keywords: Mineral oil, NIOSH 5026 method, Artificial neural network, Partial least square regression, Anticipation

HOW TO CITE THIS ARTICLE

Baghdadi F., Zendehei R., Panjali Z., Hajighasemkhan A. Prediction of Mineral Oil Concentrations Using Fourier Transform Infrared (FTIR) and Modeling Methods. *J Health Saf Work*. 2026; 16(1): 99-118.

1. INTRODUCTION

Mineral oil, a critical component of metalworking fluids, plays a significant role in reducing wear and heat during industrial operations. However, these oils, when subjected to high temperatures, aeration, and rapid surface contact, produce aerosols that pose severe health risks, such as laryngeal cancer, asthma, and lung cancer. Fourier Transform Infrared Spectroscopy

(FTIR), as described in the NIOSH 5026 method, has been widely employed for assessing exposure to mineral oils. However, the diverse and complex chemical compositions of mineral oils often result in extensive spectral interferences, complicating accurate analysis.

To address these challenges, advanced multivariate modeling methods, including Partial Least Squares (PLS) regression and Artificial Neural Networks (ANN), have been developed.

* Corresponding Author Email: hajighasemkhan@sbm.ac.ir

This study aimed to evaluate and compare the predictive capabilities of these two modeling methods in determining the concentration of mineral oil aerosols, using FTIR spectra in the automotive industry.

2. MATERIAL AND METHODS

The study utilized FTIR spectral data from the range of 1500–4000 cm^{-1} , recorded across 701 wave numbers. Standard solutions with concentrations of 5, 10, 50, 100, 200, and 250 ppm were prepared using mineral oil dissolved in carbon tetrachloride. Each concentration was analyzed five times, creating a matrix of FTIR spectral data as input (matrix X) and mineral oil concentrations as output (matrix Y). For ANN modeling, data were split into training (70%), validation (15%), and test (15%) sets, and the Levenberg-Marquardt algorithm was used for optimization. The model's architecture included multiple hidden layers, with adjustments to achieve the best performance. For PLS modeling, the same dataset was used, focusing on reducing spectral interferences and enhancing linear relationships between input and output data. Performance metrics, including Root Mean Square Error (RMSE) and relative error percentages, were used to compare the models.

3. RESULTS AND DISCUSSION

The ANN model outperformed the PLS model in all evaluated parameters. The overall RMSE for ANN was 0.0036, significantly lower than the PLS model's RMSE of 5.01. Regression coefficients for the ANN model in training, validation, and test sets were 1, 0.935, and 0.997, respectively. In contrast, the PLS model demonstrated a regression coefficient of 0.9188.

The ANN model achieved a maximum relative error of 0.074%, whereas the PLS model exhibited a maximum relative error of 7.71%. This substantial difference highlights the ANN model's superior ability to address nonlinear relationships and minimize spectral interferences compared to the linear nature of the PLS method.

These findings align with prior studies emphasizing the robustness of ANN in handling complex datasets, such as FTIR spectral data. ANN's ability to self-learn and optimize its structure based on input data contributes to its enhanced predictive accuracy. In practical applications, the ANN model offers a reliable tool for determining occupational exposure levels to mineral oils in industrial environments.

4. CONCLUSIONS

This study demonstrated the efficacy of ANN modeling in accurately predicting the concentrations of mineral oil aerosols despite the challenges posed by spectral interferences. While both ANN and PLS showed acceptable performance, ANN consistently outperformed PLS in terms of precision and error reduction. These findings suggest that ANN can serve as a robust alternative to traditional methods for evaluating occupational exposure to mineral oils. Furthermore, its application can be extended to other industries requiring accurate analysis of complex chemical mixtures.

5. ACKNOWLEDGMENT

The authors express their gratitude to the Research Council of Shahid Beheshti University of Medical Sciences for financial and logistical support, which made this study possible.

بکارگیری روش‌های مدلسازی در پیش بینی نتایج آنالیز روغن‌های معدنی با روش مادون قرمز با تبدیل فوریه (FTIR)

فاطمه بغدادی^۱، رضوان زنده دل^۱، زهرا پنجعلی^۲، علیرضا حاجی قاسمخان^{۱*}

^۱گروه مهندسی بهداشت حرفه‌ای و ایمنی کار، دانشکده بهداشت و ایمنی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

^۲مهندسی بهداشت حرفه‌ای و ایمنی کار، دانشکده بهداشت، دانشگاه علوم پزشکی آزاد اسلامی ایران، تهران، ایران

تاریخ دریافت: ۱۴۰۳/۶/۶، تاریخ پذیرش: ۱۴۰۵/۲/۱۳

چکیده

مقدمه: روغن‌های معدنی، به‌عنوان یکی از اجزای مهم مایعات فلزکاری، در فرآیندهای صنعتی به‌صورت آئروسول در محیط کار منتشر شده و با ورود از طریق استنشاق به بدن، خطرات جدی مانند سرطان حنجره، آسم، و سرطان ریه ایجاد می‌کنند. روش NIOSH ۵۰۲۶ با استفاده از اسپکتروفتومتری FTIR به ارزیابی میزان مواجهه با روغن‌های معدنی می‌پردازد. با این حال، ترکیبات شیمیایی متنوع این روغن‌ها باعث تداخلات گسترده در طیف FTIR می‌شود. در این مطالعه از دو روش مدلسازی حداقل مربعات جزئی (PLS) و شبکه‌های عصبی مصنوعی (ANN) برای پیش‌بینی صحیح غلظت مواجهه‌ی تنفسی با روغن‌های معدنی استفاده شد.

روش کار: داده‌ها از طیف FTIR در بازه $1500-4000 \text{ cm}^{-1}$ استخراج و در 701 عدد موجی ثبت شد. ماتریس طیف (X) به‌عنوان ورودی و ماتریس غلظت روغن معدنی (Y) به‌عنوان خروجی برای مدلسازی PLS و ANN در نظر گرفته شد. عملکرد مدل‌ها براساس خطای جذر میانگین مربعات (RMSEP) مقایسه شد.

یافته‌ها: در مدل PLS، ضریب رگرسیون 0.9188 و میانگین خطا $4/792$ محاسبه شد. در مقابل، مدل ANN با استفاده از کمتر از 11 لایه پنهان و 15% داده‌ها برای اعتبارسنجی، به RMSEP معادل $0/036$ و میانگین خطای $3/01\%$ دست یافت. همچنین، ANN در مقایسه با PLS دقت بالاتری در تخمین غلظت نمونه‌های آزمایشی نشان داد.

نتیجه‌گیری: مدلسازی به روش ANN امکان پیش‌بینی دقیق غلظت روغن‌های معدنی را علی‌رغم تداخلات طیفی فراهم می‌کند. در حالی که هر دو مدل ANN و PLS عملکرد مناسبی داشتند، روش ANN به دلیل دقت بالاتر و خطای کمتر، برتری بیشتری در پیش‌بینی غلظت نشان داد.

کلمات کلیدی: روغن‌های معدنی، روش NIOSH ۵۰۲۶، شبکه عصبی مصنوعی، رگرسیون حداقل مربعات

جزئی

مقدمه

امروزه صنایع مختلف از جمله صنایع فلز کاری به دنبال پیشرفت در همه ابعاد تولید هستند و وضعیت اقتصادی آنها عامل مهمی در این امر محسوب می‌شود. این صنایع از تمام منابع مادی و نیروی انسانی جهت افزایش رشد اقتصادی استفاده می‌کنند (۱، ۲). به منظور ساخت تجهیزات، ابزار، اشیاء، قطعات و سازه‌ها، تجهیزات و فرآیندهای متعددی در فلزکاری استفاده می‌شوند که یکی از مهم‌ترین آنها عملیات برشکاری است. راندمان بالای این عملیات به مقدار زیادی به شرایط برش (سرعت و عمق برش) و نرخ تغذیه وابسته است (۳). برای کاهش آسیب به تجهیزات، کاهش سایش و گرما، شستن تراشه‌ها، ذرات و آلاینده‌ها در نقاط عملیاتی، از مایعات فلزکاری (MWFs) استفاده می‌شود که می‌توان بر اساس ساختار، آنها به سه دسته زیر طبقه بندی کرد:

- ۱) مایعات برش خالص (روغن‌های معدنی، بدون آب)^۲
- ۲) مایعات برش محلول (روغن‌های معدنی امولسیون شده در آب)^۲
- ۳) مایعات برش سنتزی (آب با ترکیبات محلول، بدون روغن‌های معدنی)^۲ (۴، ۵)

یکی از مهم‌ترین اجزای مایعات فلزکاری نوع اول و دوم، روغن‌های معدنی^۲ می‌باشد که مخلوط پیچیده‌ای از مواد شیمیایی مختلف است (۶). در فرآیندهای صنعتی به دلیل هوادهی، تماس با سطوح با حرکت سریع و دمای بالا (۷)، میست‌های روغنی به صورت آئروسول‌هایی با قطرهای ۰/۱ تا ۱ میکرون تولید و در هوای محیط کار آزاد می‌شوند (۸) و به طوری که انستیتوی ملی بهداشت و ایمنی کار آمریکا (NIOSH)^۲ تخمین می‌زند ۱/۲ میلیون نفر در ایالات متحده آمریکا در معرض میست‌های مذکور هستند (۹).

بسته به کیفیت فرآیندهای پالایش، روغن‌های معدنی

1. Metal Working Fluids
2. Straight Cutting Fluids
3. Soluble Cutting Fluids
4. Synthetic Cutting Fluids
5. Mineral Oil
6. Mineral Oil Institute for Occupational Safety and Health (NIOSH)

بازیافتی، حاوی مواد شیمیایی سمی مانند ترکیبات N-Nitroso، تری اتانول آمین^۷، دی اتانول آمین^۸، فرمالدئید^۹، هیدروکربن‌های پلی آروماتیک^{۱۰} و مشتقات آلکانول آمین^{۱۱} هستند که می‌توانند برای سلامت افراد در معرض مواجهه، مضر باشد (۱۰). این آئروسول‌ها به راحتی از طریق استنشاق وارد بدن شده (۴) و سبب بروز بیماری‌های تنفسی مختلف شوند (۱۱، ۱۲). مطالعات اپیدمیولوژیک و حیوانی نشان داده‌است که قرار گرفتن در معرض میست‌های روغنی ممکن است منجر به سرطان حنجره، آسم، حساسیت بیش از حد برونش، ذات الریه لیپویدی و سرطان ریه شود (۱۳). بر اساس استانداردهای اعلام شده از سوی انجمن متخصصین بهداشت صنعتی آمریکا (ACGIH)^{۱۲}، اداره ایمنی و بهداشت شغلی آمریکا (OSHA)^{۱۳}، اداره بهداشت و ایمنی بریتانیا (HSE)^{۱۴} و دولت تایوان، مواجهه با غلظت‌های بیش از ۵ میلی گرم در متر مکعب هوا از روغن‌های معدنی پالایش شده در محیط کار، می‌تواند سبب مسمومیت‌ها و بیماری‌های مختلفی گردد (۱۴). انجمن بهداشت حرفه‌ای ژاپن (JSOH)^{۱۵} محدودیت کمتری برای مواجهه اتخاذ کرده است (۳ میلی گرم بر متر مکعب). لازم به ذکر است روغن‌های استفاده شده در فرآیندهای صنعتی دارای ناخالصی بوده و سرطانزا هستند و از این رو می‌بایست تا حد امکان مواجهه با روغن‌های تصفیه نشده کنترل شود (۴).

مهم‌ترین پیش نیاز برای هر طرح کنترلی در مواجهه با روغن‌های معدنی، ارزیابی دقیق مواجهه شغلی با آنها است. در این مطالعه روش NIOSH 5026 مبنای قرار داده شد و روش OSHA-128 مد نظر نمی‌باشد. در روش NIOSH 5026، میزان مواجهه با روغن‌های معدنی بر

7. Triethanol Amine

8. Diethanol Amine

9. Formaldehyde

10. Polyaromatic Hydrocarbons

11. Alkanol Amines

12. The American Conference of Governmental Industrial Hygienists (ACGIH)

13. Occupational Safety and Health Administration (OSHA)

14. Health and Safety Executive (HSE)

15. The Japan Society for Occupational Health (JSOH)

روغن های معدنی می تواند در فرآیند آنالیز این ترکیبات توسط دستگاه FTIR ایجاد تداخل نماید. هر نوع آئروسولی (به عنوان مثال دود سیگار) که در طول موج نزدیک به 2950 cm^{-1} اشعه مادون قرمز را جذب کند، در آنالیز آزمایشگاهی تداخل ایجاد می نماید چراکه بهترین دامنه ی جذبی روغن های معدنی براساس روش NIOSH 5026 در هنگام استفاده از دستگاه طیف سنجی IR، محدوده 2700 cm^{-1} - 1320 cm^{-1} اعلام شده است (۱۹). لذا بررسی مدل ها و روش هایی برای پیش بینی و تفسیر نتایج FTIR، جهت کمک به حذف اثر تداخلات مذکور تا حد امکان، از الزامات بررسی میزان مواجهه با روغن های معدنی است. اخیراً شبکه های عصبی مصنوعی (ANN) به مدلی محبوب و مفید برای طبقه بندی، خوشه بندی، تشخیص الگو و پیش بینی در بسیاری از رشته ها تبدیل شده اند. به گفته دکتر رابرت هخت- نیلسن (مخترع یکی از اولین کامپیوترهای عصبی)، ANN یک سیستم محاسباتی است که از تعدادی عناصر پردازشی ساده و بسیار به هم پیوسته تشکیل شده است و می تواند اطلاعات را با کمک ورودی های خارجی پردازش کند (۲۰).

کاربردهای کامل شبکه های عصبی مصنوعی را می توان با توجه به عوامل تجزیه و تحلیل داده ها مانند دقت، سرعت پردازش، تأخیر، عملکرد، تحمل خطا، حجم، مقیاس پذیری و همگرایی ارزیابی کرد (۹، ۲۱) و برای تشخیص تصویر، پردازش زبان طبیعی آنها را توسعه داده و استفاده کرد (۹). همچنین شبکه های عصبی مصنوعی قادر به رسیدگی به مشکلات در کشاورزی، علوم پزشکی، آموزش، مالی، مدیریت، امنیت، مهندسی، تجارت کالا و هنر هستند. در یک توضیح ساده، شبکه های عصبی مصنوعی داده ها را دریافت می کنند، خود را آموزش می دهند تا الگوی موجود در داده ها را تشخیص دهند و با تولید مدل، مجموعه ی جدیدی از داده های مشابه را پیش بینی می کنند (۲۰).

ساختار شبکه عصبی مصنوعی:

شبکه عصبی از چندین لایه تشکیل می شود که

2. Artificial Neural Network (ANN)

اساس اسپکترومتری مادون قرمز تبدیل فوریه (FTIR) تعیین و همچنین ترکیبات شیمیایی به وجود آمده در روغن های معدنی مصرف شده، توسط طیف مذکور تحلیل می شود. لیکن در روش NIOSH 5026، اشاره ای به محدودیت ها و عدم قطعیت ها وجود ندارد و پیشنهادی جهت کاربری مدل سازی بر ای ارتقای نتایج مشاهده نمی شود.

FTIR به صورت عمده در علوم شیمی و مهندسی شیمی، صنعت تولید دارو و صنایع غذایی مورد استفاده قرار می گیرد. همچنین در تجهیزات آزمایشگاهی برای دستیابی به طیف مادون قرمز جذبی یا نشری یک ماده جامد، مایع یا گاز استفاده می شود. در دستگاه طیف سنج FTIR بجای تاباندن یک پرتو تکفام شده به نمونه، یک پرتو با هزاران فرکانس یا طول موج مختلف به طور همزمان به نمونه تابیده می شود و میزان جذب و چگونگی آن آنالیز می گردد. در مرحله بعد، پرتو دیگری با هزاران طول موج دیگر برای تولید نقطه بعدی به نمونه تابانده می شود و این رویه چندین بار تکرار می شود. در پایان، تمامی این اطلاعات توسط کامپیوتر گردآوری شده و برای تشخیص جذب نمونه در هر نقطه، طول موج جذب شده تحلیل می شود (۱۵).

جهت تجزیه کمی به کمک دستگاه FTIR، سطح زیر پیک در یک عدد موجی خاص بررسی می گردد. عدد موجی در نواحی مختلف FTIR شاخص گروه های عاملی همچون عامل اسیدی، کنتی، الکلی می باشد. از آنجاییکه یک گروه عاملی ممکن است در مولکول های مختلف به شکل مشترک وجود داشته باشد، احتمال تداخل طیفی FTIR در مولکول هایی با ساختار شیمیایی مشابه قابل انتظار است. روغن های معدنی با توجه به فرآیند تولید، تصفیه و بازیافت، حاوی مواد افزودنی متنوعی برای بهبود مشخصات و حفظ عملکرد خود می باشند (۱۶، ۱۷). از افزودنی های روغن های معدنی می توان به ترکیباتی چون نیتروآمین، تری اتانول آمین، دی اتانول آمین و مشتقات آلکانول آمین اشاره کرد (۱۸). ناخالصی های

1. Fourier Transform Infrared Spectrometer

حداقل تعداد این لایه‌ها، ۳ لایه است که عبارت اند از:

۱) لایه ورودی یا Input Layer

۲) لایه میانی - پنهان یا Hidden Layer

۳) لایه خروجی یا Output Layer

هر لایه شبکه عصبی شامل مجموعه‌ای از گره‌ها یا نود (Nodes) است که عملکردی مشابه با نورون‌های مغز انسان دارند. تعداد گره‌های هر لایه با توجه به تعداد ابعاد داده‌های ورودی و مسئله‌ای که بناست مدل را برای آن آموزش داده شود، متفاوت خواهد بود (۲۲، ۲۳). در این مطالعه یک شبکه عصبی مصنوعی پیش‌نگر چند لایه که پرسپترون چندلایه (MLP) نیز نامیده می‌شود، با استفاده از مجموعه داده‌های معین شبکه آموزش داده می‌شود که در اصلاح به آن training می‌گویند. شبکه MLP از حداقل سه لایه شامل یک لایه ورودی، یک لایه پنهان و یک لایه خروجی تشکیل شده است. لازم به ذکر است، شبکه عصبی مصنوعی توسط داده‌های ورودی و خروجی از پیش تعیین شده، آموزش داده می‌شود و هر لایه شامل چندین عصب یا نورون است که هر کدام از عصب‌ها دارای مقدار آستانه و تابع فعال‌سازی می‌باشند و طی آموزش توسط داده‌ها بدست می‌آیند. در مرحله اعتبارسنجی شبکه یا Validation، نتایج بدست آمده از داده‌های تست که به صورت تصادفی وارد مدل شده است، با خروجی متناظر مقایسه می‌شود، که این دو مقدار باید نزدیک به هم باشند. بر همین اساس مدل می‌آموزد که وزن‌ها و مقدار آستانه را طوری تنظیم کند که خروجی درستی را دریافت کند. هرچه تعداد لایه‌ها و عصب‌ها در هر لایه پنهان بیشتر باشد، مدل پیچیده‌تر و یادگیری عمیق‌تر می‌شود. در نهایت با تغییر تعداد لایه‌های پنهان به یک حالت بهینه از یادگیری می‌رسد تا به نتیجه دلخواه دست یابد (۲۴).

روش حداقل مربعات جزئی (PLS) که گاهی اوقات به آن طرح‌ریزی ساختارهای نهفته نیز گفته می‌شود (۲)، اطلاعات موجود در دو جدول داده‌ای را به هم مرتبط می‌کند (۲). این روش برای اولین بار در اواخر دهه ۱۹۶۰

تا ۱۹۸۰ توسط اقتصاددانی به نام هرمان ولد توسعه یافت. اما حوزه اصلی توسعه اولیه آنها، شیمی سنجی می‌باشد (۲۵-۲۷). رگرسیون PLS (PLSR) یک راه حل موثر در شرایطی است که امکان به دست آوردن یک سیگنال انتخابی از یک طیف چند جزئی پیچیده، وجود ندارد (۲۸).

روش PLS با استخراج متغیرهای پنهان که ترکیب‌های خطی (بهینه) هستند، عمل می‌نماید (۲). این رویکرد معادل یک مسئله همبستگی است که تکنیک همبستگی حداقل مربعات جزئی (PLSC) نامیده می‌شود (۲). حداقل مربعات جزئی (PLS) و شبکه‌های عصبی مصنوعی (ANN) دو روش مدل‌سازی چند متغیره برای غلبه بر مشکلات تداخل طیفی در دستگاه‌هایی مثل FTIR بدون نیاز به مداخله تحلیلیگر بوده (۲۰) و ابزارهای محاسباتی غیرخطی پیچیده‌ای برای مدل‌سازی روابط غیرخطی بسیار پیچیده بین داده‌های ورودی و خروجی هستند (۲۰). مدل PLS با تمرکز بر بهترین دستاورد کوواریانس بین ماتریس X و ماتریس Y امتیازهای متغیر پنهان را محاسبه می‌کند (۲۸). یکی از اصول PLS به حداقل رساندن تأثیر گونه‌های نامربوط همچون تداخل‌ها می‌باشد و در این مطالعه مدل PLS رابطه خطی بین جذب و غلظت آنالیت را به حداکثر می‌رساند. PLS با استفاده از جعبه ابزار ChemoAC در MATLAB انجام شد. سپس از همان مجموعه داده آموزشی مورد استفاده در برازش مدل ANN برای ساخت مدل PLS استفاده شد. داده‌های مشخصات محلول‌ها در مناطق طیف FTIR انتخاب شده، به عنوان X استخراج شد و مقادیر مربوط به غلظت روغن‌های معدنی در محلول‌ها به عنوان پاسخ در یک بردار (y) تنظیم شد. داده‌ها با کدهای تعریف شده برای PLS و ANN در MATLAB مدل‌سازی شدند.

این مطالعه با بکارگیری دو روش PLS و ANN در مدل‌سازی طیف FTIR برای پیش‌بینی غلظت مواجهه‌ی تنفسی روغن‌های معدنی بر اساس روش NIOSH 5026 در یک صنعت خودروسازی طراحی شد. بر اساس روش

1. Multilayer Perceptron

2. Partial Least Squares (PLS)

مورد استفاده در این مطالعه به شرح زیر است؛

الف) تعیین دقت داده ها

انحراف معیار و درصد ضریب تغییرات از جمله شاخص هایی است که با کمک آنها می توان ضریب تغییرات و دقت یک روش را بررسی نمود (فرمول ۱).

$$CV = \frac{S(\sigma)}{\bar{X}} \times 100 \quad \text{فرمول ۱}$$

%	ضریب تغییرات	CV
	انحراف معیار	$S(\sigma)$

\bar{X} میانگین غلظت روغن معدنی ppm

جهت بررسی دقت داده های بدست آمده از آزمایشات از ضریب تغییرات درون-روز و برون-روز استفاده شد. به همین منظور؛ محلول های استاندارد در دامنه غلظتی روش NIOSH 5026 از روغن معدنی تهیه و هر غلظت را ۳ مرتبه در یک روز و یک مرتبه در ۳ روز متوالی، آنالیز کرده و سپس میانگین، انحراف معیار و ضریب تغییرات درون-روز و برون-روز محاسبه شد (۲۹).

ب) شاخص های حد تشخیص (LOD) و حد تعیین کمی (LOQ)

در این مطالعه برای تعیین LOD^۱ و LOQ^۲ (فرمول ۲ و ۳) از رابطه های زیر استفاده گردید (۲۹):

$$LOD \cong X_{bi} + 3 S_{bi} \quad \text{فرمول ۲}$$

ppm	حد تشخیص	LOD
	انحراف معیار نمونه شاهد	S_{bi}
ppm	میانگین غلظت نمونه شاهد	X_{bi}

NIOSH5026 جهت تعیین کمی روغن معدنی، بررسی سطح زیر پیک برای هر غلظت در ناحیه ی عدد موجی $2700-3200 \text{ cm}^{-1}$ توصیه شده است. هدف این مطالعه کاهش تداخلات در ناحیه ی طول موج پیشنهاد شده در روش بین المللی است لذا به جای یک عدد سطح زیر پیک در هر غلظت، میزان پاسخ در طیف FTIR برای بازه طیفی $400-4000 \text{ cm}^{-1}$ برای غلظت ها در نظر گرفته شده است تا با افزایش تعداد عددهای مورد استفاده در هر غلظت از روغن معدنی اثر نویزها یا هم پوشانی پیک ها کاهش یابد. براساس بررسی های به عمل آمده تحقیقی در زمینه تهیه مدل برای نتایج آنالیز روغن های معدنی یافت نگردید لذا در این مطالعه مدل سازی طیف FTIR روغن های معدنی بررسی شد.

روش کار

داده های مورد نیاز جهت مدل سازی براساس خروجی طیف FTIR تهیه گردید. براین اساس جهت مدل سازی از طیف FTIR در ناحیه ی 400 تا 4000 cm^{-1} استفاده شد. از آنجایی که جهت ترسیم منحنی استاندارد ۶ غلظت ارزیابی گردید هر غلظت ۵ تکرار شد و ماتریس داده ای با ابعاد 30×180 برای مدل سازی استفاده شد.

اعتبارسنجی روش تجزیه و تحلیل کمی

بر اساس روش استاندارد NIOSH 5026، اعتبارسنجی آنالیز کمی برای شش غلظت استاندارد ۵، ۱۰، ۵۰، ۱۰۰، ۲۰۰ و ۲۵۰ ppm تهیه شده از روغن معدنی، انجام شد. جهت ارزیابی نمونه ها، غلظت سازی در ۱۰ میلی لیتر تتراکلرید کربن انجام شد. برای تجزیه و تحلیل کمی نمونه ها، از دستگاه FTIR استفاده شد. برای تهیه محلول های استاندارد و رسم منحنی استاندارد در آزمایشگاه، روغن معدنی خالص و استفاده نشده از صنعت مورد نظر دریافت شد. روش اندازه گیری غلظت نمونه های حاوی روغن های معدنی در آزمایشگاه، به کمک پارامترهای سنجش پارامترهای دقت و صحت، اعتبارسنجی شد. پارامترهای

1. Limit Of Detection
2. Limit Of Quantification

$$\text{LOQ} \cong X_{bi} + 10 S_{bi} \quad \text{فرمول ۳}$$

LOQ	حد تعیین کمی	ppm
S_{bi}	انحراف معیار نمونه شاهد	
X_{bi}	میانگین غلظت نمونه شاهد	ppm

ج) تعیین صحت داده‌ها

برای تعیین صحت روش آنالیز نمونه‌ها از شاخص بازیافت نمونه استفاده شد. بدین منظور؛ غلظت‌های استاندارد از روغن معدنی در بازه غلظتی ۵ تا ۲۵۰ ppm تهیه شد و درصد بازیافت با استفاده از رابطه زیر محاسبه گردید (۲۹).

$$\text{Recovery} = C_{actual} / C_{Theoric} \times 100 \quad \text{فرمول ۵}$$

Recovery	درصد بازیافت	%
C_{actual}	غلظت واقعی	ppm
$C_{Theoric}$	غلظت تئوری	ppm

آماده‌سازی نمونه‌های مجهول

به منظور بررسی نمونه مجهول، ۱۰ نمونه هوا از سالن تراشکاری یک صنعت موتورسازی بر اساس روش NIOSH 5026 تهیه شد و با کمک مدل‌های پیشنهاد شده تعیین غلظت گردید. بدین منظور مقدار ۱۰ میلی‌لیتر از محلول کربن تتراکلرید به نمونه‌ها اضافه و نمونه‌ها به مدت ۲۰ دقیقه در دستگاه شیکر قرار داده شدند. تعیین مقدار کمی آلاننده‌ها با دستگاه طیف‌سنج FTIR در ناحیه cm^{-1} ۳۲۰۰-۲۷۰۰ انجام شد و جهت اعتباربخشی مدل، نزدیکی داده‌های حاصل از آنالیز با داده‌های پیش‌بینی شده از روش مدل‌سازی بررسی گردید.

روش اجرای مدل‌سازی

ماتریسی از داده‌های طیف FTIR به عنوان ورودی (ماتریس X) و ماتریسی از غلظت‌های روغن معدنی (ماتریس Y) به عنوان خروجی مدل‌ها، در هر دو روش مدل‌سازی، در نظر گرفته شد. جهت مدل‌سازی شبکه

عصبی مصنوعی، ماتریس داده‌ها به شکل تصادفی بین سه ماتریس validation، train و test تقسیم شد. هفتاد درصد از داده‌ها به طور تصادفی به عنوان مجموعه داده train، حداقل پانزده درصد از داده‌ها به طور تصادفی به عنوان مجموعه داده test و پانزده درصد از داده‌ها به طور تصادفی به عنوان مجموعه داده validation برای ساخت مدل‌هایی جهت پیش‌بینی تفسیر داده‌های FTIR اختصاص داده شد. در نهایت، داده‌ها با روش ANN با الگوریتم (levenberg-marquardt algorithm) و در نرم افزار MATLAB (R2023a) کدگذاری شدند.

مدل‌سازی شبکه عصبی مصنوعی یا ANN

شبکه عصبی مصنوعی در واقع دسته‌ای الگوریتم است که برای شناسایی و تشخیص الگوها به کار می‌رود. در این مطالعه یک شبکه عصبی مصنوعی پیش‌نگر چند لایه که پرسپترون چندلایه (MLP) نیز نامیده می‌شود، با استفاده از مجموعه داده‌های معین شبکه آموزش داده شده در اصلاح به آن training می‌گویند. شبکه MLP از حداقل سه لایه شامل یک لایه ورودی، یک لایه پنهان و یک لایه خروجی تشکیل شده است. لازم به ذکر است، شبکه عصبی مصنوعی توسط داده‌های ورودی و خروجی از پیش تعیین شده، آموزش داده می‌شود و هر لایه شامل چندین عصب یا نورون است که هر کدام از عصب‌ها دارای مقدار آستانه و تابع فعال‌سازی می‌باشند و طی آموزش توسط داده‌ها بدست می‌آیند. در مرحله اعتبارسنجی شبکه یا Validation، نتایج بدست آمده از داده‌های تست که به صورت تصادفی وارد مدل شده است، با خروجی متناظر مقایسه می‌شود، که این دو مقدار باید نزدیک به هم باشند. بر همین اساس مدل می‌آموزد که وزن‌ها و مقدار آستانه را طوری تنظیم کند که خروجی درستی را دریافت کند. هرچه تعداد لایه‌ها و عصب‌ها در هر لایه پنهان بیشتر باشد، مدل پیچیده‌تر و یادگیری عمیق‌تر می‌شود. در نهایت با تغییر تعداد لایه‌های پنهان به یک حالت بهینه از یادگیری می‌رسد تا به نتیجه دلخواه دست یابد.

1. Multilayer Perceptron

گذر عقب (به روزرسانی وزن ها) تنظیم شد. افزایش تعداد دوره ها در ابتدا دقت مدل را بهبود بخشید، اما پس از نقطه بهینه، منجر به بیش برآزش و کاهش عملکرد در داده های تست شد که با مکانیزم توقف زود هنگام و تنظیم بیزی از این مسئله جلوگیری گردید.

در فرآیند آزمایش شبکه، هر طیف به نقاط داده تبدیل و به عنوان ورودی به شبکه تغذیه شد. سپس شبکه با توجه به وزن های ایجاد شده، غلظت روغن معدنی را در مخلوط مربوطه تخمین زد.

مدل سازی رگرسیون حداقل مربعات جزئی یا PLS

مدل PLS با تمرکز بر بهترین دستاورد کوواریانس بین ماتریس X و ماتریس Y امتیازهای متغیر پنهان را محاسبه می کند. یکی از اصول PLS به حداقل رساندن تأثیر گونه های نامربوط همچون تداخل ها می باشد و در این مطالعه مدل PLS رابطه خطی بین جذب و غلظت آنالیت را به حداکثر می رساند. PLS با استفاده از جعبه ابزار ChemoAC در MATLAB انجام شد. سپس از همان مجموعه داده آموزشی مورد استفاده در برآزش مدل ANN برای ساخت مدل PLS استفاده شد. داده های مشخصات محلول ها در مناطق طیف FT IR انتخاب شده، به عنوان (X) استخراج شد و مقادیر مربوط به غلظت روغن های معدنی در محلول ها به عنوان پاسخ در یک بردار (y) تنظیم شد. داده ها با کدهای تعریف شده برای PLS در MATLAB مدل سازی شدند.

مرحله ارزیابی قدرت مدل سازی

پس از مدل سازی، عملکرد پیش بینی مدل ANN و PLS بر اساس RMSEP با یکدیگر مقایسه شد. این پارامتر رایج ترین معیار عملکرد هنگام ارزیابی یک مدل پیش بینی کننده است و از مقایسه مقادیر پیش بینی شده با مقادیر واقعی، توسط RMSEP محاسبه می شود.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{act} - X_{pred})^2}{n}} \quad \text{فرمول ۶}$$

1. The Root Mean Squared Error Values of Predictions (RMSEP)

در مطالعه حاضر برای توسعه مدل ANN، هر طیف با فاصله ی نقاط داده ای 5 cm^{-1} (Resolution) با تعداد ۷۰۱ داده و به صورت جداگانه به عنوان داده های ورودی و غلظت روغن معدنی در محلول مربوطه بر حسب ppm به عنوان خروجی در نظر گرفته شد. تصمیم گیری تعداد لایه های پنهان یک عامل بسیار مهم در تعیین ساختار بهینه شبکه عصبی است که مستقیماً بر عملکرد پیش بینی مدل های ANN تأثیر می گذارد. ساختار شبکه عصبی شامل یک لایه پنهان و متشکل از ۱۰ نورون بهینه سازی مدل در نظر گرفته شد و از تابع فعال سیگموئیدی در لایه پنهان استفاده گردید. عامل مهم دیگر تعداد تکرارهای مورد نیاز در طول مرحله آموزش است که در نتایج خروجی مربوط به مدل، نمایش داده می شود و اصطلاحاً با عبارت Epoch نمایش داده می شود. تعداد چرخه های تکرار نشان دهنده تعداد دفعاتی می باشد که پارامترهای الگوریتم در طول فرآیند یادگیری با آن به روز می شوند. در واقع؛ دوره یا Epoch به معنی آموزش شبکه عصبی با مجموعه داده های آموزشی برای یک چرخه است. در یک دوره، از تمام داده ها دقیقاً یک بار استفاده می شود. یک دوره رو به جلو و یک دوره رو به عقب، با هم به عنوان یک چرخه حساب می شوند: یک دوره از یک یا چند دسته تشکیل شده است که در آن از بخشی از مجموعه داده ها برای آموزش شبکه عصبی استفاده می گردد. اساساً، تعداد تکرار بیشتر، قابلیت برآزش بهتر شبکه را ایجاد می کند. با این حال، مشکل این است که تعداد بیش از حد چرخه های تکرار گاهی اوقات عملکرد مدل را مختل می کند و می تواند منجر به خطای "بیش از حد برآزش" شود. در این مطالعه از ۵۰۰۰ دوره برای آموزش استفاده گردید. الگوریتم Bayesian Regularization به طور خودکار، جهت تنظیم آموزش استفاده شد. معیار توقف آموزش داده ها، در درجه اول تکمیل بیشینه دوره های آموزش و پس از آن بهبود خطای اعتبار بخشی ناچیز در شش دوره متوالی (Early Stopping) مد نظر بود. در هر دوره، تمام داده های آموزشی یک بار به شبکه ارائه و وزن ها طی دو مرحله گذر جلو (محاسبه خروجی) و

(ماتریس Y) به عنوان خروجی مدل‌ها، در هر دو روش مدل‌سازی، در نظر گرفته شد. در نهایت پس از تعیین مدل‌های بهینه‌سازی شده، غلظت ۱۰ نمونه آزمایشی تهیه شده از صنعت با استفاده از مدل‌های توسعه‌یافته PLS و ANN برآورد گردید و عملکرد دو مدل مذکور در پیش‌بینی غلظت آنالیت با یکدیگر مقایسه شد.

اعتبارسنجی روش تجزیه و تحلیل کمی

شاخص‌های مربوط به دقت و صحت داده‌ها، کمترین حد تشخیص (LOD) و حد تعیین کمی (LOQ) محاسبه شد. نتایج حاصل از ارزیابی تکرارپذیری داده‌های آنالیز به طور خلاصه در جدول ۱ آمده است. داده‌های ذکر شده در جدول فوق حاکی از آن است که، داده‌های این مطالعه از اعتبار و تکرارپذیری مناسبی برخوردار است.

بررسی‌ها نشان داد؛ مقادیر دو پارامتر حد تشخیص (LOD) و حد تعیین کمی (LOQ) برای روش مذکور به ترتیب ۰/۰۵ و ۰/۱۵ میلی‌گرم به ازای هر نمونه می‌باشد. صحت روش آنالیز نمونه‌ها نیز با محاسبه میانگین میزان درصد بازیافت برای غلظت‌های استاندارد برابر ۹۴/۸۶±۸/۴ تعیین شد.

مدل‌سازی به روش PLS

پس از بکارگیری ماتریس ورودی و خروجی در نرم‌افزار MATLAB از آنجاییکه تعدادی از پاسخ‌های ثبت شده در دستگاه FTIR منفی گزارش شده بود در دو

N	تعداد نمونه‌های آزمایشی	
X_{act}	غلظت واقعی	ppm
X_{pred}	غلظت پیش‌بینی شده توسط شبکه	ppm

علاوه بر این، عملکرد مدل‌ها با محاسبه درصد خطای نسبی در تخمین مقدار روغن معدنی در هر یک از داده‌ها با استفاده از معادله زیر با یکدیگر مقایسه شدند.

$$\text{Relative error (\%)} = \left(\frac{X_{pred} - X_{act}}{X_{act}} \right) \times 100 \quad \text{فرمول ۷}$$

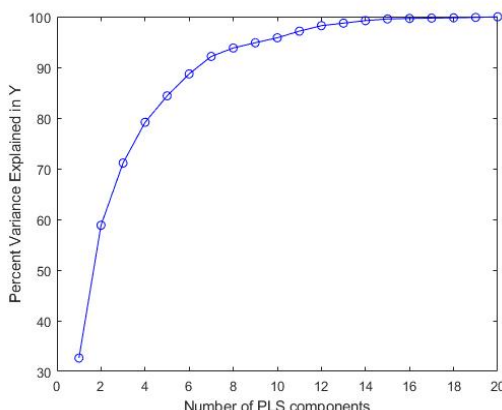
X_{act}	غلظت واقعی	ppm
X_{pred}	غلظت پیش‌بینی شده توسط شبکه (۱۲)	ppm

یافته‌ها

در این مطالعه عملکرد دو مدل PLS و ANN جهت پیش‌بینی غلظت روغن‌های معدنی به روش NIOSH 5026 بررسی و با یکدیگر مقایسه شد. برای این منظور ۶ غلظت استاندارد مطابق روش توصیه شده توسط NIOSH ساخته شد و توسط دستگاه FT-IR آنالیز شد. فرآیند مذکور برای هر یک از غلظت‌های استاندارد، ۵ بار تکرار گردید. برای ارزیابی اعتبارسنجی روش آنالیز، شاخص‌های مربوط به دقت و صحت داده‌ها، کمترین حد تشخیص (LOD) و حد تعیین کمی (LOQ) محاسبه شد. سپس ماتریسی از داده‌های طیف FTIR به عنوان ورودی (ماتریس X) و ماتریسی از غلظت‌های روغن معدنی

جدول ۱: ضریب تغییرات روش تجزیه نمونه روغن معدنی

ضریب تغییرات (CV)		غلظت روغن معدنی (ppm)
برون روز	درون روز	
۲/۴۴۳	۱/۲۹۷	۵
۰/۹۶۲	۲/۰۸۳	۱۰
۶/۰۷۷	۱/۴۷۸	۵۰
۵/۶۵۳	۶/۳۴۹	۱۰۰
۳/۳۱۹	۰/۴۰۱	۲۰۰
۲/۷۸۵	۱/۶۶۵	۲۵۰



شکل ۱: انتخاب تعداد اجزای اصلی در مدل PLS

رگرسیون های بدست آمده از منحنی های پاسخ در مدل PLS و مقدار RMSEp محاسبه شده، مدل با تعداد اجزای اصلی ۱۰ و به کارگیری کل داده های FTIR، به عنوان بهترین مدل جهت پیش بینی غلظت روغن های معدنی انتخاب شد.

ضعف عملکرد مدل PLS نسبت به ANN در این مطالعه عمدتاً به ماهیت خطی روش PLS در برابر پیچیدگی و غیرخطی بودن داده های طیف سنجی FTIR روغن های معدنی بازمی گردد. تنوع ترکیبات شیمیایی و تداخلات گسترده طیفی موجب می شود رابطه بین جذب و غلظت به طور کامل خطی نباشد، در حالی که PLS تنها قادر به استخراج روابط خطی بر اساس متغیرهای پنهان است. در مقابل، مدل ANN با بهره گیری از معماری چندلایه و توابع فعال سازی غیرخطی، توانایی یادگیری الگوهای پیچیده و روابط غیرخطی بین داده های FTIR و غلظت روغن معدنی را داشته و با استفاده از مکانیزم های تنظیم سازی و توقف زود هنگام، خطای پیش بینی به طور چشمگیری کاهش یافته است؛ به گونه ای که مقادیر RMSEp و خطای نسبی در ANN به مراتب کمتر از PLS گزارش شد.

مدل سازی به روش ANN

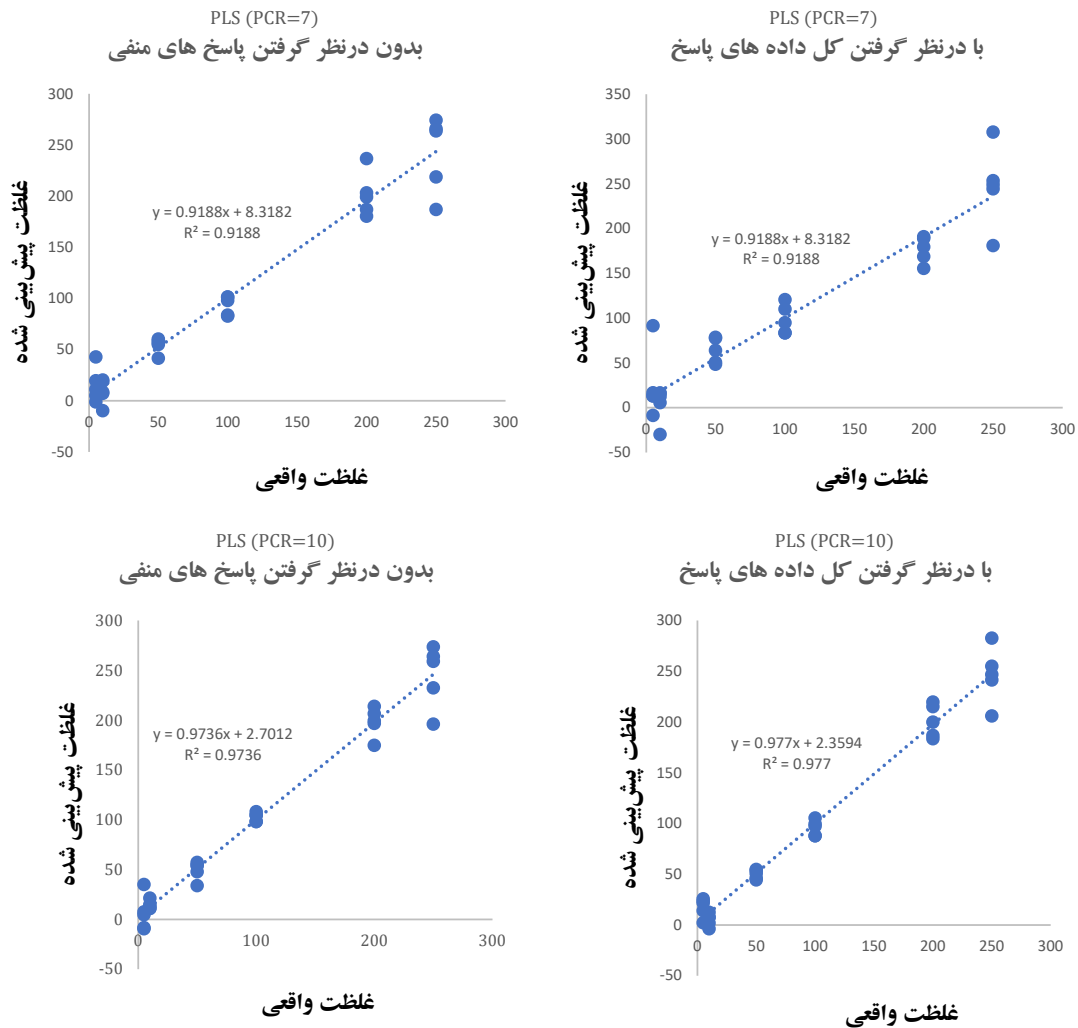
در این مطالعه از ۳۰ نمونه آزمایشگاهی ساخته شده با غلظت های استاندارد و توصیه شده توسط روش

حالت با در نظر گرفتن کل داده ها و حذف داده های منفی مدل ارزیابی شد. شکل ۱ پارامتر percent variance را به ازای مقادیر اجزای اصلی مختلف در روش PLS نشان می دهد. همانطوری که در منحنی نشان داده شده است در جز اصلی ۷ شکستی در منحنی نشان داده شده است که تا جز اصلی دهم ادامه دارد. لذا مدل سازی در تعداد اجزای اصلی ۷ و ۱۰ بررسی گردید.

• رگرسیون های بدست آمده از منحنی های پاسخ مدل PLS به عنوان غلظت پیش بینی شده در مقابل غلظت واقعی، در محدوده ۰/۹۱۸۸ الی ۰/۹۷۷ تعیین شد (شکل ۲).

در مدل PLS زمانی که اجزای اصلی تعداد ۷ استفاده شد؛ مقدار RMSEp در شرایطی که کل داده های FTIR در نظر گرفته شده بود مساوی با ۲۸/۳۱۳ و در شرایط بدون در نظر گرفتن علامت منفی داده های FTIR، مقدار RMSEp به ۱۹/۵۰۸۵ کاهش یافت. از طرف دیگر زمانی که اجزای اصلی مقدار ۱۰ اعمال شد، در شرایط بدون در نظر گرفتن داده های منفی در مقایسه با کل داده ها، مقدار RMSEp به ترتیب ۱۵/۲۱۲ و ۱۴/۱۴۵ محاسبه شد. نتایج فوق نشان داد که انتخاب تعداد اجزای اصلی مناسب در شرایطی که کل داده های FTIR در نظر گرفته شود، عملکرد PLS را در آنالیز کمی روغن معدنی به طور رضایت بخشی بهبود می بخشد. در نهایت با توجه به

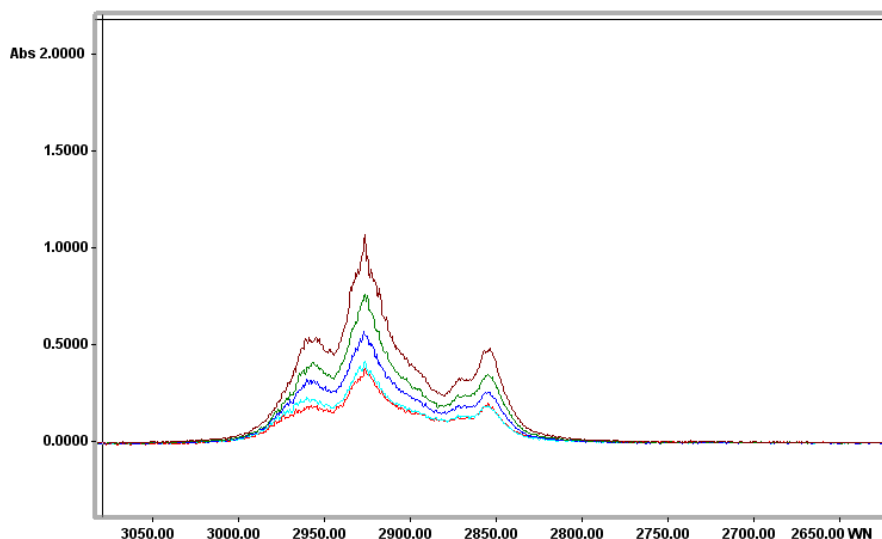
1. Components



شکل ۲: منحنی های پاسخ مدل PLS به عنوان غلظت پیش بینی شده در مقابل غلظت واقعی

کنترلی در این مطالعه به طور همزمان به کار گرفته شد. ابتدا داده ها به صورت تصادفی به مجموعه های آموزش، اعتبارسنجی و آزمون تقسیم شدند (به طوری که ۱۵٪ داده ها به اعتبارسنجی اختصاص یافت) تا تعمیم پذیری مدل در داده های مستقل ارزیابی شود. فرآیند آموزش با استفاده از مکانیزم توقف زود هنگام (early stopping) انجام شد، به گونه ای که در صورت عدم بهبود خطای مجموعه اعتبارسنجی، آموزش متوقف می گردید و از یادگیری نوین جلوگیری می شد. علاوه بر این، محدود کردن تعداد epochها (۸ دوره آموزش) و بهینه سازی

NIOSH 5026، برای ساخت مدل های کالیبراسیون چند متغیره استفاده شد. شکل ۳ طیف FT-IR در ناحیه ی توصیه شده توسط NIOSH را نشان می دهد. مجموعه داده های آموزشی ذکر شده برای برازش مدل ANN جهت پیش بینی غلظت روغن معدنی توسط الگوریتم Levenberg-Marquardt اجرا شد. غلظت روغن معدنی مربوط به هر نمونه استاندارد نیز به عنوان خروجی شبکه در نظر گرفته شد. به منظور جلوگیری از بیش برازش (overfitting) در مدل شبکه عصبی مصنوعی (ANN)، چندین راهکار



شکل ۳: طیف FT-IR مجموعه داده های آموزشی

زمان انجام فرآیند (یا Elapsed Time) به میزان ۲۰ ثانیه تعیین شده است. رگرسیون بدست آمده از رسم منحنی غلظت های نمونه Test و غلظت های پیش بینی شده با مدل ANN مقدار ۰/۹۹۷ می باشد (شکل ۴). جدول ۳ خصوصیات مدل بهینه شده ی ANN را نشان داده است.

- (a) Gradient، گرادیان مربع تابع خطا است.
 - (b) Mu پارامتر کنترلی تأثیرگذار بر همگرایی خطا در الگوریتم مورد استفاده جهت آموزش شبکه عصبی است که در مورد الگوریتم LMS به حداکثر مقدار ویژه ماتریسمبستگی ورودی وابسته است.
 - (c) Validation Check به تعداد چرخه های یادگیری پایان یافته شبکه عصبی مصنوعی اشاره دارد که به تعداد چرخه های تکرار بستگی دارد.
- پس از تعیین مدل های بهینه سازی شده برای پیش بینی غلظت روغن های معدنی، داده های طیف های FT-IR در ۱۵ نمونه ی Test که در فرآیند برازش مدل ها مورد استفاده قرار نگرفتند، بررسی شدند. خطای نسبی مدل ها در تخمین غلظت نمونه های Test تعیین گردید (جدول ۴).

مشاهده می شود که روش سنتی PLS غلظت روغن

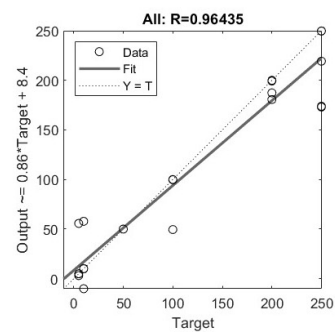
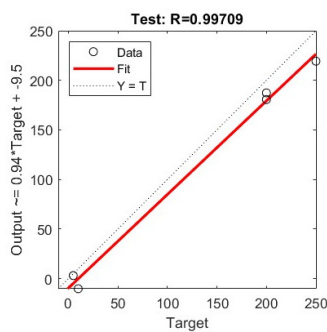
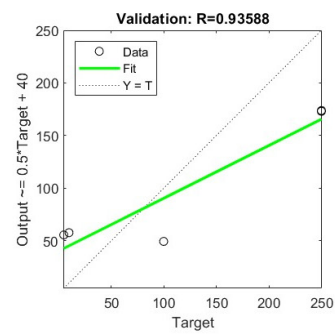
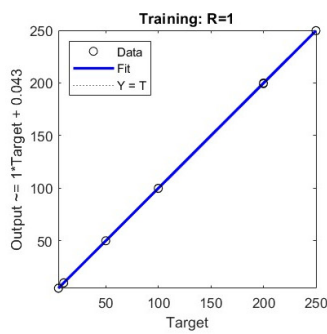
معماری شبکه (تعداد لایه های پنهان کمتر از ۱۱) به عنوان نوعی regularization ساختاری عمل کرده و از افزایش غیرضروری پیچیدگی مدل جلوگیری نمود. عملکرد پایدار مدل در داده های آزمون و نزدیکی معیارهای خطا در مجموعه های آموزش و آزمون نشان می دهد که مدل ANN بدون بیش برآزش، توانسته است الگوی واقعی رابطه بین طیف FTIR و غلظت روغن معدنی را به طور مؤثر یاد بگیرد.

نتایج تجربی به دست آمده در تغییر فاکتورهای مؤثر بر عملکرد ANN جهت رسیدن به بهترین پیش بینی برای تعیین غلظت آنالیت در جدول ۲ نشان داده شده است. بطوریکه تغییرات مربوط به سه رزولوشن ۴، ۵ و 6 cm^{-1} با تعداد لایه های پنهان ۲، ۵، ۷، ۹، ۱۰ و ۱۱ برای داده های Test و Validation با تعداد اجراهای ثابت، نشان داده شده است. نتایج رگرسیون در رسم منحنی داده های واقعی Test برحسب داده های پیش بینی شده از ۰/۸۷۴ تا ۰/۹۹۷ متفاوت بوده است.

نتایج نشان می دهد بهترین شرایط در مدل سازیدر روش ANN مقدار رزولوشن 5 cm^{-1} ، ۵، ۱۱ لایه پنهان می باشد که رگرسیون ۱، ۰/۹۳۵ و ۰/۹۹۷ به ترتیب برای داده های Validation، Training و Test، با مدت

جدول ۲: خلاصه نتایج تغییرات اعمال شده بر فاکتورهای مؤثر بر عملکرد ANN

ردیف	تعداد لایه‌های پنهان	درصد داده‌های Validation (%)	درصد داده‌های Test (%)	مدت زمان انجام فرآیند Test (s)	رگرسیون Training: (R)	رگرسیون Validation: (R)	رگرسیون Test: (R)
۱	۲	۱۵	۱۵	۴	۰/۹۹۹	۰/۹۹۰	۰/۹۶۵
۲	۵	۱۵	۱۵	۱۰	۰/۹۹۷	۰/۹۸۵	۰/۹۷۲
۳	۷	۱۵	۱۵	۴۳	۰/۹۹۹	۰/۴۶۰	۰/۸۷۴
۴	۹	۱۵	۱۵	۵۰	۰/۹۹۶	۰/۹۳۳	۰/۹۰۰
۵	۱۰	۱۵	۱۵	۵۳	۰/۹۹۵	۰/۹۸۹	۰/۸۹۰
۶	۱۱	۱۵	۱۵	۶۹	۱	۰/۹۸۴	۰/۹۵۶
۷	۲	۱۵	۱۵	۱	۰/۹۶۹	۰/۹۷۹	۰/۹۷۲
۸	۵	۱۵	۱۵	۲۲	۰/۹۹۶	۰/۹۵۶	۰/۹۶۰
۹	۷	۱۵	۱۵	۱۵	۰/۹۹۷	۰/۹۹۳	۰/۹۶۷
۱۰	۹	۱۵	۱۵	۲۳	۱	۰/۸۶۰	۰/۸۷۸
۱۱	۱۰	۱۵	۱۵	۳۰	۰/۹۹۸	۰/۸۹۶	۰/۹۹۲
۱۲	۱۱	۱۵	۱۵	۲۰	۱	۰/۹۳۵	۰/۹۹۷
۱۳	۲	۱۵	۱۵	۳	۰/۶۲۲	۰/۹۴۳	۰/۹۰۸
۱۴	۵	۱۵	۱۵	۸	۰/۹۸۸	۰/۹۷۵	۰/۹۸۷
۱۵	۷	۱۵	۱۵	۸	۰/۹۷۹	۰/۹۴۹	۰/۹۷۷
۱۶	۹	۱۵	۱۵	۱۲	۰/۹۶۳	۰/۹۲۳	۰/۹۸۱
۱۷	۱۰	۱۵	۱۵	۱۹	۱	۰/۹۲۱	۰/۹۶۵
۱۸	۱۱	۱۵	۱۵	۱۷	۰/۹۷۹	۰/۷۸۱	۰/۹۱۵



شکل ۴: رگرسیون منحنی‌های Training، Validation و Test مدل ANN

جدول ۳: مشخصات مدل ANN انتخابی بر اساس بهترین رگرسیون

Unit	Initial value	Stopped value	Target value
تعداد چرخه های تکرار Epoch	۰	۸	۱۰۰۰
مدت زمان انجام فرآیند تست Elapsed Time	-	۰۰:۰۰:۲۰	-
عملکرد Performance	۱.۹۲e+۰۴	۵.۱۹e-۲۷	۰
پیاده سازی مشتق دوم Gradient	۱.۶۳e+۰۷	۴.۶۷e-۱۱	۱e-۰۷
Mu	۰.۰۰۱	۱e-۰۸	۱e-۱۰
Validation Check	۰	۳	۶

جدول ۴: غلظت واقعی و پیش بینی شده روغن معدنی توسط مدل های ANN و PLS در مجموعه داده های Test

ردیف	غلظت واقعی (ppm)	مدل ANN		مدل PLS	
		غلظت پیش بینی شده	خطای نسبی (%)	غلظت پیش بینی شده	خطای نسبی (%)
۱	۵	۴/۹۹۶۳	-۰/۰۷۴	۴/۷۲۶	-۵/۴۸
۲	۱۰	۱۰/۰۰۱۳	۰/۰۱۳	۱۰/۵۶۱	۵/۶۱
۳	۵۰	۵۰/۰۰۰۹	۰/۰۰۱۸	۴۶/۸۳۱	-۶/۳۴
۴	۱۰۰	۱۰۰/۰۰۲۲	-۰/۰۰۲۲	۹۲/۲۹۳	-۷/۷۱
۵	۲۰۰	۱۹۹/۹۹۳۵	-۰/۰۰۳۳	۲۰۰/۰۳۴	۰/۰۲
۶	۲۵۰	۲۴۹/۹۹۶	-۰/۰۰۱۶۸	۲۵۸/۹۸۷	۳/۵۹
RMSEP _{Total}		۰/۰۰۳۶		۵/۰۱	

جدول ۵: ارزیابی شاخص های ارزیابی مدل های ANN و PLS

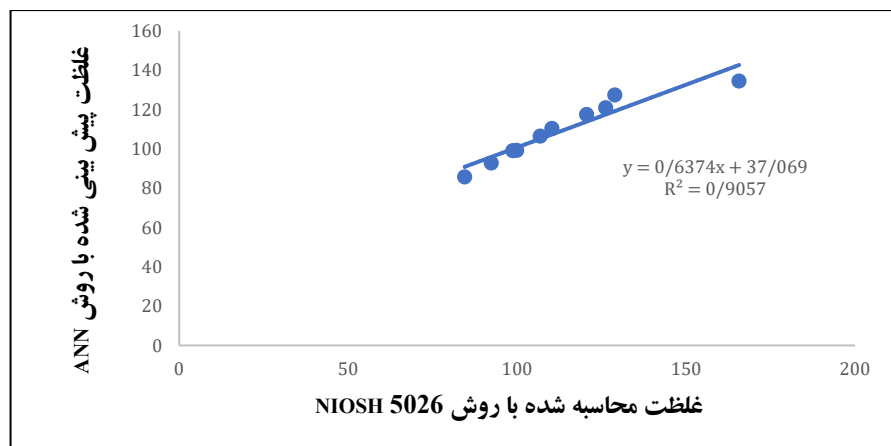
شاخص ارزیابی	مدل ANN	مدل PLS
RMSEP	۰/۰۰۳۶	۵/۰۱
MAE	۰/۰۸۶	۴/۲۳
MAPE (%)	۰/۱۲۳	۵/۸۷
R ² (آموزش)	۱/۰۰۰	۰/۹۷۷
R ² (آزمون)	۰/۹۹۷	۰/۹۱۹
ضریب همبستگی	۰/۹۹۹	۰/۹۵۸

در مدل ANN حداکثر خطای نسبی ۰/۰۷۴ گزارش شد. در جدول ۵ شاخص های مختلف جهت مقایسه عملکرد مدل های ANN و PLS ارایه شده است. باقیمانده های دو مدل ANN و PLS توزیع نرمال داشتند. (p-value > 0.05) توزیع باقیمانده ها در ANN به طور تصادفی بدون الگوی سیستماتیک حول صفر توزیع شده اند در حالیکه در PLS الگوی خفیف سیستماتیک

معدنی را در مجموعه داده های Test با میانگین خطای ۴/۷۹۲ تخمین زده است، در حالی که این مقدار برای مدل ANN ۰/۰۹۶ بوده است. علاوه بر این، RMSEP کلی مدل PLS در تخمین غلظت آنالیت نمونه های Test، ۵/۰۱ بود که در مدل ANN به ۰/۰۰۳۶ ارتقا یافت. مشاهده می شود که در مدل PLS حداکثر خطای نسبی در پیش بینی غلظت ۷/۷۱ درصد بوده است در حالی که

جدول ۶: ارزیابی قدرت پیش‌بینی مدل ANN در تعیین غلظت نمونه‌های مجهول

ردیف	غلظت نمونه مجهول	غلظت پیش‌بینی شده توسط مدل ANN	درصد خطا
۱	۱۰۶/۴۳	۱۰۶/۸۶	۰/۴۰۲۴
۲	۹۹/۰۰۳	۹۸/۷۷	۰/۲۳۵۹
۳	۹۲/۷۸	۹۲/۳۸	۰/۴۴۲
۴	۱۱۰/۳۵	۱۱۰/۳۱	۰/۰۳۶۳
۵	۱۲۷/۴	۱۲۸/۹۵	۱/۲۰۲
۶	۱۲۰/۸۴	۱۲۶/۲۴	۴/۲۷۷۵
۷	۹۹/۲۳	۹۹/۸۸	۰/۶۵۰۸
۸	۱۳۴/۴۷	۱۶۵/۷	۱۸/۸۴۷۳
۹	۱۱۷/۴۶	۱۲۰/۵۹	۲/۵۹۵۶
۱۰	۸۵/۷	۸۴/۴۹	۱/۴۳۲۱



شکل ۵: منحنی رگرسیون حاصل از ترسیم نمونه‌های مجهول تعیین شده با روش NIOSH 5026 و غلظت‌های پیش‌بینی شده با روش ANN

تعیین شده توسط مدل به عنوان داده خروجی، درصد خطای مدل محاسبه گردید (جدول ۶). نتایج خلاصه شده در جدول فوق نشان می‌دهد که میانگین درصد خطای محاسبه شده براساس غلظت نمونه‌های مجهول پیش‌بینی شده توسط ANN، $5/71 \pm 3/01$ می‌باشد. شکل ۵ منحنی رگرسیون حاصل از ترسیم نمونه‌های مجهول تعیین شده با روش NIOSH 5026 و غلظت‌های پیش‌بینی شده با روش ANN را مشخص ساخته است.

و در محدوده‌های غلظتی بالاتر مشاهده شد. با نظر به نتایج فوق می‌توان نتیجه گرفت که مدل توسعه یافته ANN در تخمین غلظت روغن‌های معدنی نمونه Test عملکرد بهتری نسبت به مدل PLS داشته است، در نتیجه می‌توان در تعیین غلظت نمونه‌های مجهول صنعت از آن بهره گرفت. بر همین اساس جهت ارزیابی قدرت پیش‌بینی مدل ANN در تعیین غلظت نمونه‌های واقعی، طیف FTIR نمونه‌های مجهول به عنوان داده‌های ورودی به مدل انتخابی تغذیه شد و در نهایت بر مبنای غلظت

هدف از این مطالعه پاسخ به این پرسش است که آیا روش های مدل سازی می توانند راه حل جایگزین و کم خطاتری نسبت به روش های سنتی (روش های رایج منحنی کالیبراسیون تک متغیره) برای تعیین غلظت روغن های معدنی جهت ارزیابی کیفیت محیط کار باشند. مواجهه با ذرات ۱ میکرونی میست روغن های معدنی معلق در هوا ممکن است یک عامل خطر بالقوه برای بیماری های مزمن انسدادی باشند. مواجهه با OMPM می تواند فیبروز ریوی را با هدف قرار دادن اپیتلیوم ریه و فیبروبلاست ها و فعال کردن مسیرهای سیگنال دهی TGF- β 1/Smad3 و TGF- β 1/MAPK p38 ایجاد کند (۴، ۳۰). با توجه به اثرات نام برده شده در مطالعات گذشته، می بایست جهت حفظ سلامت کارکنان شاغل در اینگونه صنایع، ارزیابی دقیقی از مواجهه شغلی با میست روغن های معدنی انجام شود.

در این مطالعه، به منظور تعیین غلظت نمونه های حاوی میست روغن معدنی با دستگاه طیف سنج FTIR در ناحیه $2700 - 3200 \text{ cm}^{-1}$ مطابق با روش NIOSH 5026، از محلول های استاندارد روغن معدنی تهیه شده از یک صنعت موتورسازی استفاده شد. میانگین ضریب تغییرات درون-روز و برون-روز به ترتیب ۲/۲۱ و ۳/۵۴ محاسبه شد، لذا داده های این مطالعه از اعتبار مناسبی برخوردار است. مقادیر دو پارامتر حد تشخیص (LOD) و حد تعیین کمی (LOQ) برای روش مذکور، به ترتیب ۰/۰۵ و ۰/۱۵ میلی گرم به ازای هر نمونه مشخص گردید. صحت روش آنالیز نمونه ها برای غلظت های استاندارد برابر ۹۴/۸۶ تعیین شد.

روش حداقل مربعات جزئی (PLS) یا طرح ریزی ساختارهای نهفته (۲)، اطلاعات موجود در دو دسته داده را به هم مرتبط می کند (۲). PLS یک راه حل موثر در شرایطی است که امکان به دست آوردن یک سیگنال انتخابی از یک طیف چند جزئی پیچیده، وجود ندارد (۲۶). مطالعات بسیاری در این حوزه به بررسی

1. Oil-mist Particulate Matter

روش مدل سازی PLS جهت پیش بینی غلظت ترکیبات و اعتبارسنجی روش پرداخته اند. مطالعه ای جهت آنالیز کمی مخلوط روغن های معدنی به روش PLS با استفاده از داده های طیف سنجی فرابنفش (UV) و مادون قرمز میانی (MIR) توسط زایاکسو ژانگ و همکاران در سال ۲۰۲۱ انجام شد که نتایج حاصل از آن نشان داد مدل MIR-PLSR از دقت بالایی در پیش بینی کمی ترکیبات هدف برخوردار است.

رگرسیون های بدست آمده از منحنی های پاسخ مدل PLS به عنوان غلظت پیش بینی شده در مقابل غلظت واقعی، در محدوده ۰/۹۱۸۸ الی ۰/۹۷۳۶ محاسبه شد. نتایج مطالعه نشان داد تعداد اجزای اصلی مناسب، عملکرد PLS را در آنالیز کمی روغن معدنی به طور رضایت بخشی بهبود می بخشد. از طرفی نتایج برخی مطالعات حاکی از عملکرد ضعیف مدل خطی PLS در قیاس با مدل های غیرخطی است، از جمله این مطالعات می توان به مطالعه انجام شده توسط لی و همکاران اشاره نمود. این مطالعه که با هدف توسعه یک مدل رگرسیون آنالیز جهت تعیین کمی مقادیر آلبومین گلیکوزیله در سرم انسانی پرداخته است. در این مطالعه بر اساس داده های طیف ATR-FTIR و با ترکیب دو روش؛ SiPLS^۲ به عنوان مدل خطی و SVM^۳ به عنوان مدل غیرخطی مطالعه شده است، نتایج نشان داده است که عملکرد مدل سازی SVM به طور قابل توجهی بهتر از SiPLS خطی بوده است.

در راستای بکارگیری رویکرد شبکه عصبی مصنوعی به منظور تحلیل داده ها بر پایه طیف FTIR در سال های اخیر مطالعاتی انجام شده است که از آن می توان به مطالعه ای توسط پوپسک و همکاران اشاره کرد. هدف از این مطالعه، توسعه مدل ANN مبتنی بر داده های طیف سنجی FTIR، ارتقای آنالیز کمی ترکیبات سنگ های کلیه بود که اختلاف بین مقادیر محاسبه شده و پیش بینی شده ترکیبات ارزیابی شد. همچنین مطالعه ای

2. Synergy Interval Partial Least Square
3. support Vector Machine

مدل‌ها فراهم شود. استفاده از چند نمونه واقعی محیطی نیز به منظور بررسی امکان‌پذیری کاربرد مدل در شرایط عملی انجام شد و نتایج نشان داد که مدل ANN توانایی تعمیم قابل قبولی به داده‌های خارج از مجموعه آموزشی دارد. افزایش تعداد نمونه‌های صنعتی در مطالعات آینده می‌تواند به بهبود پایداری آماری مدل، کاهش عدم قطعیت و افزایش تعمیم‌پذیری نتایج به شرایط واقعی محیط کار منجر شود.

نتیجه گیری

نتایج این مطالعه حاکی از آن است؛ تعیین غلظت نمونه‌های حاوی روغن‌های معدنی با دقت بالا، از طریق مدل‌سازی به روش ANN امکان پذیر است. از آنجایی که پیچیدگی ترکیبات شیمیایی در ترکیبات روغن معدنی بسیار بالاست لذا جهت افزایش دقت اندازه‌گیری بکارگیری مدلسازی در اندازه‌گیری ترکیبات روغن معدنی در محیط‌های کاری مختلف توصیه می‌شود.

نتایج این مطالعه همچنین نشان می‌دهد که هر دو مدل چند متغیره ANN و PLS در ارزیابی کمی نمونه‌های روغن‌های معدنی عملکرد نسبتاً خوبی داشتند، اما روش ANN عملکرد بهتری از نظر نزدیکی غلظت نمونه‌های پیش‌بینی شده نسبت به غلظت نمونه‌های واقعی نشان داد. علاوه بر این، مدل را می‌توان با طیف گسترده‌ای از داده‌های آموزشی بهینه کرد و به عنوان یک روش مناسب جهت تعیین غلظت نمونه‌های حاوی آلاینده‌های محیط کار مورد استفاده قرار داد. از آنجایی که ابزارهای ساده‌ای همچون nntools برای اجرای مدل شبکه‌ی عصبی در نرم‌افزار MATLAB توسعه یافته است لذا چنین رویکردی را به سادگی در سنجش مواجهات شغلی می‌توان پیاده‌سازی کرد.

دیگر در تایید بکارگیری ANN در غلظت‌سنجی پتاسیم فنوکسی متیل پنی سیلین، توسط بین وانگ و همکاران بر اساس داده‌های طیف‌سنجی FNIRS با استفاده از روش پیش‌پردازش O-PLS همراه با روش غیرخطی ANN انجام شد. مطالعه‌ای توسط صالحی و همکاران در سال ۲۰۲۱ به ارزیابی و مقایسه عملکرد دو روش کالیبراسیون چند متغیره، ANN و PLS، در آنالیز کمی کوارتز پرداخته است. نتایج حاصل از این مطالعه حاکی از آن بود که هر دو مدل در تجزیه و تحلیل کمی کوارتز، عملکرد بسیار خوبی داشتند اما مدل ANN که با توانایی پیش‌بینی غیرخطی ذاتی خود، عملکرد بهتری از خود نشان داده است. در این مطالعه تغییرات مربوط داده‌های Test از ۰/۸۷۴ به ۰/۹۹۷ متفاوت بود. نتایج حاکی از آن بود که انتخاب عوامل مناسب در طول مرحله یادگیری به طور قابل توجهی توانایی پیش‌بینی شبکه را در آنالیز کمی روغن‌های معدنی بهبود می‌بخشد. نتایج حاصل از ارزیابی و مقایسه عملکرد پیش‌بینی مدل‌ها نشان داد مدل ANN در تخمین غلظت روغن‌های معدنی عملکرد بهتری نسبت به مدل PLS داشته است، به همین دلیل برای تعیین غلظت نمونه‌های مجهول از آن بهره گرفته شد. نتایج این مطالعه با یافته‌های قبلی در مورد توانایی منحصر به فرد روش‌های ANN در یافتن روابط بین داده‌ها و حل مسائل پیچیده در آنالیزهای شیمیایی، مطابقت دارد (۲۰).

هدف این پژوهش در گام نخست، ارزیابی قابلیت روش FTIR در ترکیب با مدل‌های داده‌محور برای پیش‌بینی غلظت روغن‌های معدنی و مقایسه عملکرد روش‌های مدل‌سازی بوده است. داده‌های آزمایشگاهی با دامنه وسیعی از غلظت‌ها طراحی شدند تا بیشترین تنوع طیفی و شرایط مواجهه را پوشش دهند و بدین ترتیب، مبنای مناسبی برای آموزش و ارزیابی اولیه

REFERENCES

- Moradpour Z, Ramezanifar S, Asgari Gandomani E, Zendehtdel R. Semi-quantitative respiratory health risk assessment of exposure to metalworking fluids (oil mists) in an automotive industry. *J Chem Health Risks*. 2022;12(4):777-84.
- Abdi H, Williams LJ. Partial least squares methods: partial least squares correlation and partial least square regression. In: Reisfeld B, Mayeno AN, editors. *Computational toxicology: volume II*. New York: Humana Press; 2013. p. 549-79.
- Ozimina D, Kowalczyk J, Madej M, Nowakowski Ł, Kulczycki A. The impact of the type of cutting fluid on the turning process. *Tribologia*. 2017;275(5):79-88.
- Moradpour Z, Ramezanifar S, Asgari Gandomani E, Zendehtdel R. Semi-quantitative respiratory health risk assessment of exposure to metalworking fluids (oil mists) in an automotive industry. *J Chem Health Risks*. 2022;12(4):777-84.
- Verma DK, Shaw DS, Shaw ML, Julian JA, McCollin SA, Tombe Kd. An evaluation of analytical methods, air sampling techniques, and airborne occupational exposure of metalworking fluids. *J Occup Environ Hyg*. 2006;3(2):53-66.
- Asgari M, Azari M, Zendehtdel R, Khodakarim S, Rafiepour A, Tavakol E, et al. Development of a new method for analysis of oil mists. *Health Scope*. 2017;6(3):e15114.
- Asgari M, Azari M, Zendehtdel R, Khodakarim S, Rafiepour A, Tavakol E, et al. Development of a new method for analysis of oil mists. *Health Scope*. 2017;6(3):e15114.
- Park SS, Kang MS, Hwang J. Oil mist collection and oil mist-to-gas conversion via dielectric barrier discharge at atmospheric pressure. *Sep Purif Technol*. 2015;151:324-31.
- Abiodun OI, Jantan A, Omolara AE, Dada KV, Mohamed NA, Arshad H. State-of-the-art in artificial neural network applications: a survey. *Heliyon*. 2018;4(11):e00938.
- Azari M, Tavakol E, Yazdankhah Z, Dadashpour A. Developing and validating a new method in analysis of metal working fluids. *Pajoohande*. 2009;13(6):495-500.
- Huynh CK, Herrera H, Parrat J, Wolf R, Perret V. Occupational exposure to mineral oil metalworking fluid (MWFs) mist: development of new methodologies for mist sampling and analysis. Results from an inter-laboratory comparison. *J Phys Conf Ser*. 2009;151:012062.
- Sauvain JJ, Suarez G, Hopf NB, Batsungnoen K, Charriere N, Andre F, et al. Oxidative potential of aerosolized metalworking fluids in occupational settings. *Int J Hyg Environ Health*. 2021;235:113775.
- Wang YF, Tsai PJ, Chen CW, Chen DR, Dai YT. Size distributions and exposure concentrations of nanoparticles associated with the emissions of oil mists from fastener manufacturing processes. *J Hazard Mater*. 2011;198:182-7.
- Hsu HI, Chen MR, Wang SM, Chen WY, Wang YF, Young LH, et al. Assessing long-term oil mist exposures for workers in a fastener manufacturing industry using the Bayesian decision analysis technique. *Aerosol Air Qual Res*. 2012;12(5):834-42.
- Bahrami A. *Methods of sampling and analysis of pollutants in air*. Hamadan: Hamadan University of Medical Sciences; 2015. 352 p.
- Simpson A, Groves J, Unwin J, Piney M. Mineral oil metal working fluids (MWFs)—development of practical criteria for mist sampling. *Ann Occup Hyg*. 2000;44(3):165-72.
- Malloy EJ, Miller KL, Eisen EA. Rectal cancer and exposure to metalworking fluids in the automobile manufacturing industry. *Occup Environ Med*. 2007;64(4):244-9.
- Azari M, Tavakol E, Yazdankhah Z, Dadashpour A. Developing and validating a new method in analysis of metal working fluids. *Pajoohande*. 2009;13(6):495-500.
- Rafiepour A. Comparison of two methods of Fourier transform infrared (IR-FT) and ultraviolet (UV) spectroscopy in the measurement of mineral oils. *Iran Occup Health*. 2017;14(2):70-9.
- Salehi M, Zare A, Taheri A. Artificial neural networks (ANNs) and partial least squares (PLS) regression in the quantitative analysis of respirable crystalline silica by Fourier-transform infrared spectroscopy (FTIR). *Ann Work Expo Health*. 2021;65(3):346-57.
- Mozaffari A, Emami M, Fathi A. A comprehensive investigation into the performance, robustness, scalability and convergence of chaos-enhanced evolutionary algorithms with boundary constraints. *Artif Intell Rev*. 2019;52:2319-80.
- Eluyode O, Akomolafe DT. Comparative study of

- biological and artificial neural networks. *Eur J Appl Eng Sci Res.* 2013;2(1):36-46.
23. ul Islam B, Baharudin Z, Raza MQ, Nallagownden P. Optimization of neural network architecture using genetic algorithm for load forecasting. In: 2014 5th International Conference on Intelligent and Advanced Systems (ICIAS); 2014 Jun 3-5; Kuala Lumpur, Malaysia. IEEE; 2014.
24. Weakley AT, Miller AL, Griffiths PR, Bayman SJ. Quantifying silica in filter-deposited mine dusts using infrared spectra and partial least squares regression. *Anal Bioanal Chem.* 2014;406:4715-24.
25. Wold H. Estimation of principal components and related models by iterative least squares. In: Krishnaiah PR, editor. *Multivariate analysis.* New York: Academic Press; 1966. p. 391-420.
26. Wold H. Nonlinear iterative partial least squares (NIPALS) modelling: some current developments. In: Krishnaiah PR, editor. *Multivariate analysis-III.* New York: Academic Press; 1973. p. 383-407.
27. Wold H. Soft modelling: the basic design and some extensions. In: Jöreskog KG, Wold H, editors. *Systems under indirect observation, part II.* Amsterdam: North-Holland; 1982. p. 1-54.
28. Gezahegn TL. Fourier transform infrared spectroscopy combined with partial least squared regression and standard addition-net analyte signal method for quantification of chitin in insect-based fish feeds [master's thesis]. Bergen: The University of Bergen; 2018.
29. Shokoohi Y, Kakooei H. Standard methods of measurement and evaluation of hazardous agents in the workplace. Tehran: Tehran University of Medical Sciences Research Deputy; 2011.
30. Nie H, Liu H, Shi Y, Lai W, Liu X, Xi Z, et al. Effects of different concentrations of oil mist particulate matter on pulmonary fibrosis in vivo and in vitro. *Toxics.* 2022;10(11):647.