

Review

Not peer-reviewed version

Intelligent Petroleum Processing: A Short Review on Applying AI/ML to Petroleum Products Optimization

Abdallah Alzahawy and [Hayder Issa](#) *

Posted Date: 5 April 2024

doi: 10.20944/preprints202404.0442.v1

Keywords: Intelligent petroleum processing; petroleum products optimization; machine learning techniques; gasoline property prediction; process control and operational efficiency



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

Intelligent Petroleum Processing: A Short Review on Applying AI/ML to Petroleum Products Optimization

A. H. Alzahawi ¹ and H. M. Issa ^{2,*}

¹ Kirkuk District Governorate, Kirkuk, 36001, Iraq

² Department of Chemistry, College of Science, University of Garmian, Kalar, Sulaymaniyah Province, 46021, Iraq

* Correspondence: author: hayder.mohammed@garmian.edu.krd

Abstract: Improving the quality of petroleum products and refining processes through the use of artificial intelligence and machine learning techniques is the topic of this article. It shows that expert knowledge and conventional empirical models can't get you where you want to go in a refining process. To effectively capture complex interactions and forecast fuel qualities, machine learning techniques such as principal component analysis (PCA), support vector machines (SVM), artificial neural networks (ANN), and partial least squares (PLS) are suggested. Gasoline and other petroleum products, as well as property prediction, process control, product quality, and operational efficiency in refineries, can all be improved with the help of machine learning applied to spectral or distillation curve data. An exciting new direction in optimizing operations, meeting environmental norms, and precisely estimating gasoline quality is offered by advanced machine learning algorithms.

Keywords: intelligent petroleum processing; petroleum products optimization; machine learning techniques; gasoline property prediction; process control and operational efficiency

1. Introduction

Before looking into the new area of using machine learning to improve the quality of petroleum products, it's important to see how it could help the oil refining business become more efficient, profitable, and environmentally friendly (Sircar et al., 2021). When it comes to refining oil, the biggest task is to get the best product quality, process efficiency, and use of resources. In the past, this optimization process relied a lot on empirical models and expert knowledge (Ma et al., 2024). However, because refining processes are complicated and don't work in a straight line, the results are often less than ideal. Machine learning (ML) techniques, on the other hand, have changed the oil business by giving people powerful ways to get useful information from huge amounts of data (Al Jlibawi et al., 2021).

An application of machine learning to the optimization of petroleum products entails the utilization of algorithms and models for the purpose of analyzing complicated datasets that are obtained from the various phases of the refining process. Through the utilization of methodologies like supervised and unsupervised learning, neural networks, and predictive analytics, refineries have the ability to improve their decision-making processes, optimize product yields, increase quality specifications, and streamline operations (Al-Jamimi et al., 2022).

2. Exploring the Application of Various Machine Learning Methods for Predicting Vapor Pressure in Sour Natural Gas

Predicting vapor pressure correctly is very important in the complicated world of natural gas handling, especially when working with sour gas streams that contain a lot of hydrogen sulfide and carbon dioxide. Usually, old-fashioned computer programs can't fully capture how temperature, pressure, and the make-up of a gas interact with each other, which can lead to waste and safety issues (H. M. Issa, 2015). Different machine learning methods, like neural networks (ANN) and ensemble models, can be used to make accurate predictions about vapor pressure in sour natural gas systems.

These machine learning techniques are able to discover intricate patterns and relationships across large datasets that contain a wide variety of gas compositions and operating conditions (H. M. Issa, 2016). Because of this, they are able to generate more precise estimates of the vapor pressure. The information that was acquired has the potential to assist in the enhancement of the design of processes, the acceleration of operations, and the reduction of risks that are associated with variations in vapor pressure. In the long run, machine learning would make the natural gas processing sector more profitable and sustainable than it is currently experiencing (Nimmanterdwong et al., 2021).

3. Utilizing Machine Learning for Quantitative Prediction of Reid Vapor Pressure in Crude Oil

Crude oil volatility and vapor-liquid equilibrium depend on Reid vapor pressure (RVP). Safe RVP storage, transportation, and refining require accurate quantitative prediction. It's possible that traditional empirical correlations and equations of state are wrong because they can't show the complicated relationship between RVP and the complex makeup of light crudes (Nascimento et al., 2018). Using machine learning to create reliable RVP forecasting models for light and heavy crude oils. Using a lot of raw data and advanced algorithms like neural networks, random forests, and gradient boosting, it is possible to find complicated patterns and nonlinear connections between RVP and API gravity, molecular weight distribution, and compositional analysis. The more accurate RVP predictions from the ML models allow for more precise process design, the best ways to blend oils, and better management of risks related to vapor pressure in both upstream and downstream light crude oil operations (Hayder M. Issa & Albarzanji, 2020).

The ability to precisely anticipate the RVP in crude oil is of critical importance in the petroleum industry, particularly in the refining process, as it plays a significant role in guaranteeing product quality, safety, and operating efficiency. While traditional methods frequently rely on simplified correlations, it is possible that these correlations do not fully convey the intricate interaction of variables that influence vapor pressure. The incorporation of machine learning techniques, on the other hand, presents a potentially fruitful route for improving the accuracy and dependability of predictions (Kongkiatpaiboon, 2024).

The application of machine learning approaches for the purpose of quantitatively forecasting RVP in light crude oil is being investigated more thoroughly. Exploring a correlation method that is both streamlined and accurate makes use of the capacity of sophisticated algorithms to analyze complex datasets and create correct predictions (Lamberg, 2021). Through the utilization of machine learning, refineries have the potential to enhance their processes, enhance the quality of their products, and streamline their operations, thus contributing to the advancement of effective and sustainable practices within the petroleum industry.

4. Utilizing Machine Learning Techniques for Predicting Gasoline Properties

Predicting accurately the properties of gasoline is one of the most important tasks in optimizing petroleum products. This is necessary to ensure the quality of the products and that they meet strict industry standards. Usually, old ways of doing things depend on regression analysis or interpreting spectral or distillation curve data by hand, which might not give you a lot of information about the complex relationships in the dataset. Adding advanced machine learning methods, on the other hand, looks like a good way to improve the accuracy of predictions and get useful data from complicated data structures (Correa Gonzalez et al., 2021).

It is crucial to estimate gasoline qualities accurately to optimize operations, maintain product quality, and comply with strict environmental standards. Conventional laboratory testing methods, although dependable, can be slow and take a lot of resources, which can impede the flexibility needed in today's ever changing refining industry. Many works have examined the use of sophisticated machine learning methods to quantitatively forecast important gasoline characteristics by utilizing easily accessible data sources, including spectral analysis and distillation curves (Hayder Mohammed Issa, 2024).

Spectral or distillation curve data is combined with regression analysis, PCA, and PLS to make quantitative predictions about the properties of gasoline. Refineries may be able to better understand

gasoline composition, maximize production, and meet strict quality standards by using these advanced methods. This exploration shows how machine learning may revolutionize the petroleum sector's efficiency and creativity (Hayder M. Issa, 2024b). Partial least squares (PLS), regression analysis, and principal component analysis (PCA) form a robust ensemble in the suggested methodology (Li & Qin, 2021; Wang et al., 2022). They quantitatively established correlations between input factors and target gasoline attributes by capturing the underlying latent variables that govern the complex interactions between input data and desired outputs. PLS further enhances the predictive power, while PCA enables dimensionality reduction and identification of the most salient features within the spectral or distillation curve data (Cai et al., 2021). Usually, studies in this trend have the overarching goal of creating computationally efficient and very accurate predictive models for characteristics like octane rating, vapor pressure, and compositional analysis by integrating various methods (Baird & Oja, 2016).

Combining machine learning with spectral or distillation curve data could improve gasoline property prediction. Spectral analysis, which captures gasoline samples' unique molecular fingerprint, can be used to feed prediction models quickly and cheaply. Distillation curve data, obtainable from refinery operations, also reveals gasoline stream boiling range dispersion and composition. These data sources and the synergistic combination of regression, PCA, and PLS methods can help refineries monitor and optimize gasoline properties in real time, improving process control, product quality, and operational efficiency (Hayder M. Issa, 2024a).

Artificial neural networks (ANN) and support vector machines (SVM) provide further methods for predictive modeling and analysis in the field of petroleum product optimization. ANNs are adept at capturing intricate connections in intricate datasets, resembling the human brain's structure. This makes them ideal for forecasting gasoline properties. Its capacity to adapt and learn from data allows it to reveal complex patterns and relationships that could be missed by conventional methods (Tipler et al., 2022). SVM is a robust supervised learning technique that excels in managing high-dimensional data and is especially successful when dealing with non-linearly separable data. By integrating ANN and SVM techniques into the predictive modeling framework, refineries can improve the precision and reliability of gasoline property forecasts (Leal et al., 2020). Advanced machine learning algorithms enhance regression analysis, PCA, and PLS techniques to create a comprehensive toolkit for optimizing gasoline manufacturing operations and maintaining product quality.

Conclusions

The study delves into the topic of how to optimize the quality of petroleum products and refining operations using artificial intelligence and machine learning approaches. It shows how conventional empirical models and specialized expertise can't guarantee perfection in the refining process. We suggest using machine learning approaches including support vector machines (SVM), artificial neural networks (ANNs), principal component analysis (PCA), and partial least squares (PLS) to capture complicated interactions and forecast fuel qualities. Refineries can improve their process control, product quality, and operational efficiency by integrating machine learning with spectral or distillation curve data. This can improve petroleum products like gasoline, as well as property prediction. To optimize operations, predict gasoline quality accurately, and comply with environmental norms, advanced machine learning technologies offer a viable option.

The study's findings demonstrate that machine learning methods are useful for improving the quality of petroleum products, refining operations, and predicting gasoline properties. Figure 1 shows how these methods enhance operational efficiency, process control, and decision-making in refineries.

Research can move further by investigating and creating more sophisticated machine learning algorithms that are optimized for petroleum processing, taking into account the particular features and intricacies of the sector. Data availability, model interpretability, and scalability are just a few of the constraints and obstacles that need to be better examined and discussed in order to utilize AI/ML approaches in the petroleum business.

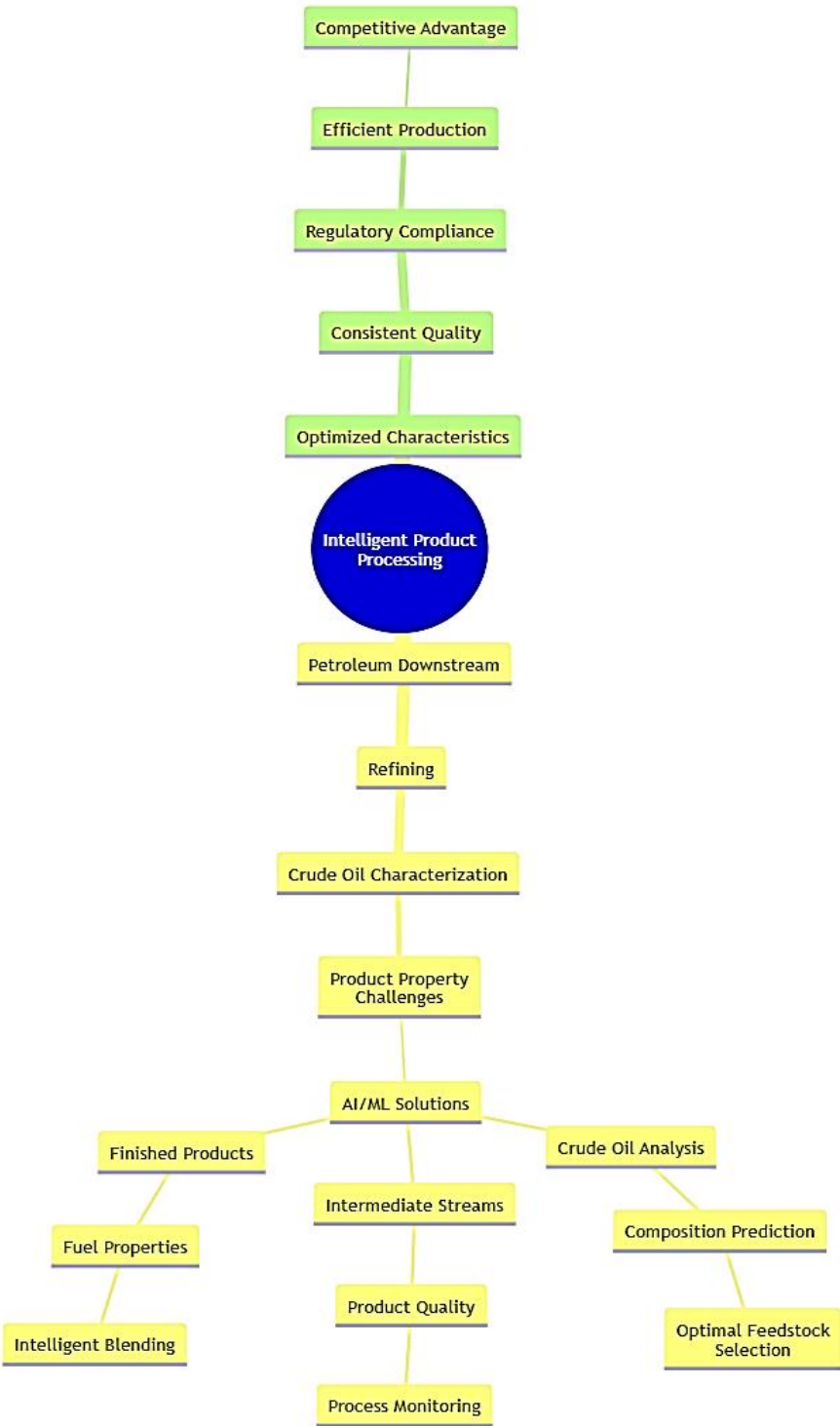


Figure 1. A mind map illustrates the petroleum processing concepts of an AI-driven refinery.

Taking into account aspects like implementation costs, maintenance expenses, and potential return on investment, future studies might also investigate the economic viability and cost-effectiveness of using AI/ML systems in petroleum processing optimization.

References

1. Al-Jamimi, H. A., BinMakhashen, G. M., & Saleh, T. A. (2022). Multiobjectives optimization in petroleum refinery catalytic desulfurization using Machine learning approach. *Fuel*, 322, 124088.
2. Al Jlibawi, A. H. H., Othman, M. L., Ishak, A., Noor, B. S. M., & Al Huseiny, M. (2021). Optimization of distribution control system in oil refinery by applying hybrid machine learning techniques. *IEEE Access*, 10, 3890-3903.

3. Baird, Z. S., & Oja, V. (2016). Predicting fuel properties using chemometrics: a review and an extension to temperature dependent physical properties by using infrared spectroscopy to predict density. *Chemometrics and Intelligent Laboratory Systems*, 158, 41-47.
4. Cai, G., Liu, Z., Zhang, L., Shi, Q., Zhao, S., & Xu, C. (2021). Systematic performance evaluation of gasoline molecules based on quantitative structure-property relationship models. *Chemical Engineering Science*, 229, 116077.
5. Correa Gonzalez, S., Kroyan, Y., Sarjovaara, T., Kiiski, U., Karvo, A., Toldy, A. I., . . . Santasalo-Aarnio, A. (2021). Prediction of gasoline blend ignition characteristics using machine learning models. *Energy & Fuels*, 35(11), 9332-9340.
6. Issa, H. M. (2015). An Investigation of the Comparative Performance of Diverse Vapor Pressure Prediction Techniques of Sour Natural Gas. *Petroleum Science and Technology*, 33(13-14), 1443-1448.
7. Issa, H. M. (2016). A new correlation for vapor pressure prediction of natural gas mixture. *Petroleum Science and Technology*, 34(23), 1913-1919.
8. Issa, H. M. (2024a). Parametric and Nonparametric Approaches of Reid Vapor Pressure Prediction for Gasoline Containing Oxygenates: A Comparative Analysis Using Partial Least Squares, Nonlinear, and LOWESS Regression Modelling Strategies with Physical Properties. *Modelling and Simulation in Engineering*, 2024, 8442457.
9. Issa, H. M. (2024). Prediction of octane numbers for commercial gasoline using distillation curves: a comparative regression analysis between principal component and partial least squares methods. *Petroleum Science and Technology*, 42(10), 1233-1249.
10. Issa, H. M. (2024b). Streamlining aromatic content detection in automotive gasoline for environmental protection: Utilizing a rapid and simplified prediction model based on some physical characteristics and regression analysis. *Results in Engineering*, 21, 101771.
11. Issa, H. M., & Albarzanji, A. A. (2020). Quantitative prediction of Reid vapor pressure for a light crude oil using a simplified and proper correlation. *Petroleum Science and Technology*, 1-9.
12. Kongkiatpaiboon, S. (2024). Improving Offshore Condensate Production Optimization Processes with Prediction of Reid Vapor Pressure Using Machine Learning. Paper presented at the Offshore Technology Conference Asia.
13. Lamberg, A. (2021). Predicting distillation properties of fuel blends using Machine Learning.
14. Leal, A. L., Silva, A. M., Ribeiro, J. C., & Martins, F. G. (2020). Using spectroscopy and support vector regression to predict gasoline characteristics: a comparison of ¹H nmr and nir. *Energy & Fuels*, 34(10), 12173-12181.
15. Li, B., & Qin, C. (2021). Predictive analytics for octane number: a novel hybrid approach of KPCA and GS-PSO-SVR model. *IEEE Access*, 9, 66531-66541.
16. Ma, G., Shi, J., Xiong, H., Xiong, C., Zhao, R., & Zhang, X. (2024). Machine learning assisted molecular modeling from biochemistry to petroleum engineering: A review. *Geoenergy Science and Engineering*, 212770.
17. Nascimento, M. H., Oliveira, B. P., Rainha, K. P., Castro, E. V., Silva, S. R., & Filgueiras, P. R. (2018). Determination of flash point and Reid vapor pressure in petroleum from HTGC and DHA associated with chemometrics. *Fuel*, 234, 643-649.
18. Nimmanterdwong, P., Changpun, R., Janthboon, P., Nakrak, S., Gao, H., Liang, Z., . . . Sema, T. (2021). Applied artificial neural network for hydrogen sulfide solubility in natural gas purification. *ACS omega*, 6(46), 31321-31329.
19. Sircar, A., Yadav, K., Rayavarapu, K., Bist, N., & Oza, H. (2021). Application of machine learning and artificial intelligence in oil and gas industry. *Petroleum Research*, 6(4), 379-391.
20. Tipler, S., D'Alessio, G., Van Haute, Q., Parente, A., Contino, F., & Coussement, A. (2022). Predicting octane numbers relying on principal component analysis and artificial neural network. *Computers & Chemical Engineering*, 161, 107784.
21. Wang, H., Chu, X., Chen, P., Li, J., Liu, D., & Xu, Y. (2022). Partial least squares regression residual extreme learning machine (PLSRR-ELM) calibration algorithm applied in fast determination of gasoline octane number with near-infrared spectroscopy. *Fuel*, 309, 122224.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.