MEKATRONIKA
e-ISSN: 2637-0883
VOL. 2, ISSUE 1, 47 – 51
DOI: https://doi.org/10.15282/mekatronika.v2i1.6729

ORIGINAL ARTICLE

Inspection of Crude Oil Condition using Electronic Nose (E-Nose)

Nur Ain Zarani Zailani[1], Muhammad Sharfi Najib[1], Suhaimi Mohd Daud[1], Mujahid Mohamad[1], Saiful Nizam Tajuddin[2] and Che Mohd Aizal Che Mohd[2]

[1]Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, 26600 Pahang, Malaysia.
[2]Pusat Bioaromatik, Universiti Malaysia Pahang, 26300 Pahang, Malaysia.

ABSTRACT – Oil and gas production and distribution processes technologies are highly complex and capital-intensive. Crude oil is a high demand commodity in Malaysia and across the world. Physical and chemical properties are used to classify crude oil in oil and gas industries. The human's nose cannot distinguish the difference of smell among various crude oils grade. Conventional approaches to detect odour are expensive and difficult to operate. Due to declining production and increasing demand, using E-nose technologies to inspect the odour condition of crude oil might be a significant change in the industries. The Case-Based Reasoning (CBR) classification method also is utilised in this project to classify crude oil conditions. As a result, all crude oil samples have their odour profile pattern extracted through the normalisation of data. The performance accuracy of the CBR classifier achieved a high rate, which is 99.31% on average. Hence, the using of E-nose and utilising CBR are excellent methods in investigating odour.

ARTICLE HISTORY
Received: 3rd May 2020
Revised: 1st June 2020
Accepted: 7th June 2020

KEYWORDS
Crude oil
E-nose
CBR
Normalisation
Confusion matrix

INTRODUCTION

Crude oil is a highly sought-after commodity in Malaysia and across the world. As the world's primary fuel sources, oil and natural gas are major industries in the energy sector and significantly impact the global economy. In terms of dollar value, the oil and gas industry is one of the biggest in the world, making these energy sources vital to the global economy, especially for its largest producers [1]. Upstream, midstream, and downstream are the three segments of the oil and gas sector. For oil and gas companies with integrated upstream and downstream business portfolios, the impact of low oil prices can be cushioned [2]. In the past years, downstream players have embarked on various initiatives to maximise value and opportunities in the current low oil price environment by pursuing downstream projects, particularly petrochemical [3].

In the past years, downstream players have embarked on various initiatives to maximise value and opportunities in the current low oil price environment by pursuing downstream projects, particularly petrochemical [4][5]. Unconventional oil and gas resources have significantly impacted global oil and gas supply and are now gradually becoming alternatives to conventional oil and gas resources [6][7]. Oil production in the Association of South-East Asian Nations (ASEAN) is far from insignificant: Indonesia ranks among the top 20 oil-producing countries globally, and Brunei, Malaysia, Thailand and Vietnam also produce significant amounts of oil [8]. ASEAN's regional energy cooperation activities marked by the creation of the ASEAN Economic Community (AEC) in 2015 set energy security and sustainability as major goals for advancing the energy sector and economy [8][9]. In Malaysia, the oil and gas industry is a significant contributor to the economy [5][10].

METHODOLOGY

Data Measurement

Overall set-up is as shown in Figure 1. The measurement of odour data for each sample is done using an E-Nose sensor, and each sample will have five repetitions of the measurement dataset. Between the repetition, the samples odour need to be neutralised using Ethanol to make sure the data reading is stable. Each sample is set up to three different temperatures as an additional variable, which are 40°C, 60°C and 80°C.

Data Pre-Processing

Data pre-processing is a process that enables the extraction of significant features from data. This procedure can be used for and pattern recognition [11]. Data pre-processing is an important step that is required to get high-quality data. It also refers to the steps that are necessary to measure the quality of the data. The quality of the data is a measure based on accuracy, sensitivity and specificity.

*CORRESPONDING AUTHOR | Muhammad Sharfi Najib | sharfi@ump.edu.my

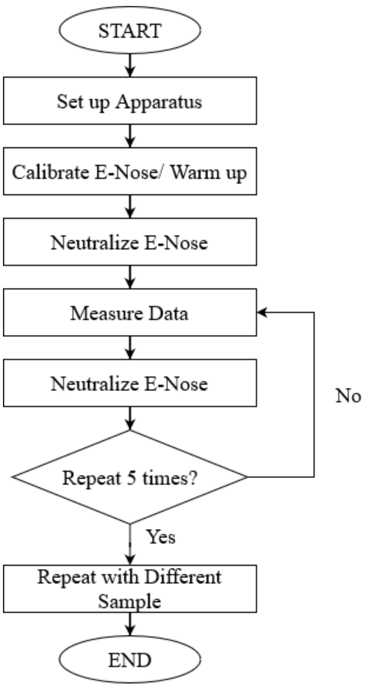


Figure 1. Data measurement flow chart.

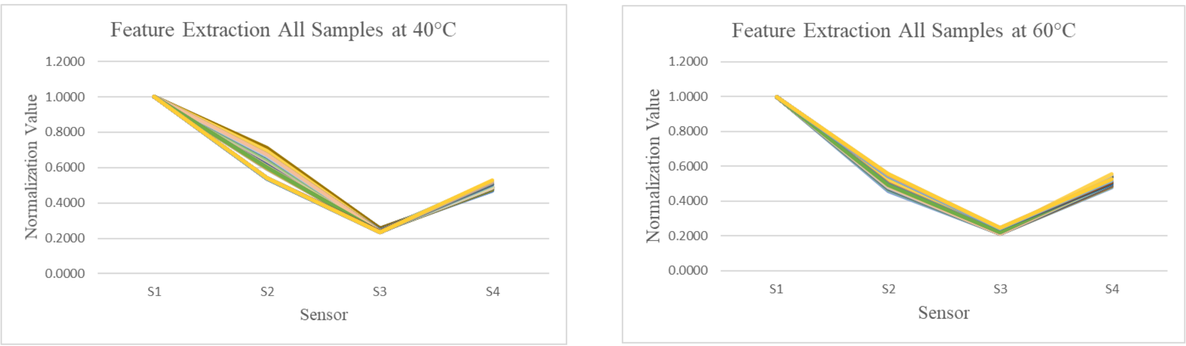
All raw data measurements need to be normalised in the beginning of data pre-processing before proceeding to another step. The mean calculation technique is used to normalise all data values in which R_{max} is the maximum raw data and R is the raw data using the given formula:

$$R' = \frac{R}{R_{max}} \tag{1}$$

The normalised data is recorded accordingly and for each sensor, the normalised data for all five repetitions will be combined to find the average. The values of normalised data are standardised to the 0-1 range. Thus, it leads to extract unique features efficiently.

Feature Extraction

Feature extraction is a simple process that aims to identify the predominant features in a given dataset. It is usually performed by extracting the relationship between the features and a target feature [11]. The normalised data values from previous data pre-processing are plotted in a graph against the number of the sensor to do feature extraction. Based on the graph generated, different graph patterns for every sample shows that every sample has its unique odour profile. Figure 2 a and Figure 2b shows the odour profile pattern graphs for all samples at 40°C and 60°C, whereas Figure 3 illustrates sampe at 80°C.



(a) (b)

Figure 2. Feature extraction odour profile pattern for (a) 40°C sample (b) 60°C sample.

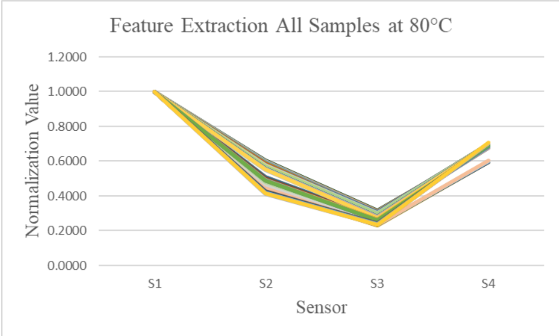


Figure 3. Feature extraction odour profile pattern for 80°C samples.

Case-Based Reasoning (CBR)

Case retrieval, case reuse, case revise, case retain and case base are the five parts of a conventional CBR. As seen in Figure 4 is the basic flow chart of how CBR works.

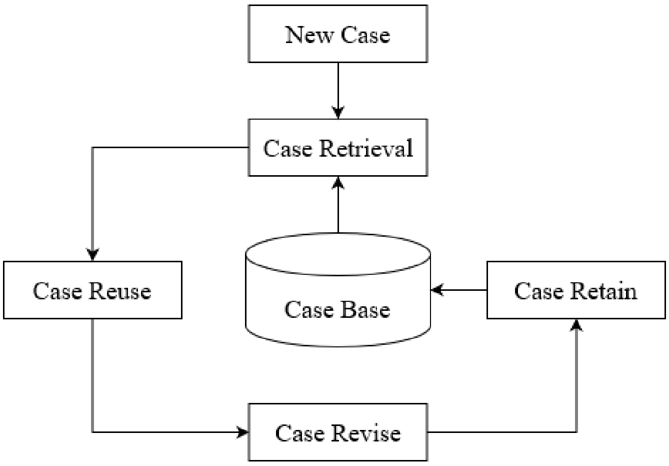


Figure 4. Basic flow chart of CBR.

The CBR retrieval process employs similarity measures. The purpose of similarity is to choose cases that can be easily adapted to the current issue and choose cases with (almost) the same solution as the current issue. The basic assumption for similarity measure is similar problems have similar solutions. The formula to calculate similarity is given as follows:

$$sim(A,B) = \frac{1}{\sum w_i} \cdot \sum_{i=1}^{p} w_i \cdot sim_i(a,b) \qquad (2)$$

A is the new case, *B* is the previous cases, *a* and *b are* denoted as new and previous features, respectively. The number of attributes is *p*, *i* is the iteration, w_i is the weight of attributes, and sim_i is the local similarity. Similarity measure is used in solving and reasoning a new crude oil case to match previous cases in the CBR case library. The available cases might fit efficiently or almost identical to the new current case. The similarity results in Figure 5 indicate the similarity between cases in the library and new cases. The first, second, and third similarity rankings for each case are tabulated in case-based reasoning voting table to measure classification performance. Accuracy performance is acquired with the help of similarity measures.

Confusion Matrix

The confusion matrix method is used to measure the performance of classification. Table 1 depicts the confusion matrix in a two-class classification task [12]. The chart shows that four different outcome forecasts can be made. False positive and false negative results are two forms of errors, while true positive and true negative results are valid classifications. False positive examples are negative examples of classes that have been incorrectly classed as positive. In contrast, false negative examples are positive examples of classes that have been incorrectly classed as negative.

				K1				K2			
ATTRIBUTION	LOCAL WEIGHT	NORMALIZED WEIGHT	CURRENT CASE	STORED CASE	SIMILARITY FUNCTION	NORMALIZED SIMILRAITY FUNCTION	WEIGHTED SIMILARITY	STORED CASE	SIMILARITY FUNCTION	NORMALIZED SIMILRAITY FUNCTION	WEIGHTED SIMILARITY
S1	1	0.25	1.0000	1.0000	1	1	0.25	1.0000	1	1	0.25
S2	1	0.25	0.4136	0.6556	0.758	0.758	0.1895	0.6411	0.7725	0.7725	0.193125
S3	1	0.25	0.2337	0.2467	0.987	0.987	0.24675	0.2467	0.987	0.987	0.24675
S4	1	0.25	0.7054	0.5096	0.8042	0.8042	0.20105	0.5046	0.7992	0.7992	0.1998
	4				MAX		SUM		MAX		SUM
					1		0.8873		1		0.889675
					SIMILARITY		88.73		SIMILARITY		88.9675

Figure 5. CBR similarity table.

Table 1. Confusion matrix table concept.

		Predicted Class	
		Negatives	Positives
Actual Class	Negatives	TN	FP
	Positives	FN	TP

Based on the confusion matrix, accuracy, sensitivity and specificity performance can be calculated for each samples. The formula to calculate accuracy, sensitivity, and specificity is given in the following equations:

$$Accuracy = \frac{TP + TN}{TN + FP + FN + TP} \tag{3}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{4}$$

$$Specificity = \frac{TN}{TN + FP} \tag{5}$$

EXPERIMENTAL RESULTS

This experiment investigates the performance of a crude oil condition classifier utilising Case-Based Reasoning. The experimental parameters used in this experiment are seven different samples of crude oil at three temperature variants. Table 2 demonstrates that 7 out of 12 chosen samples that have been analysed show 100% accuray performance while the other five samples show a rate higher than 96%. Furthermore, the overall performance result is great as the average accuracy is 99.31%. The odour of crude oil can be classified with a high rate of sensitivity and specificity.

Table 2. Performance results calculated.

	ACCURACY	SENSITIVITY	SPECIFICITY
2 (40)	99.17	100.00	99.10
2 (60)	100.00	100.00	100.00
2 (80)	100.00	100.00	100.00
4 (40)	96.94	78.79	98.78
4 (60)	100.00	100.00	100.00
4 (80)	100.00	100.00	100.00
5 (40)	97.78	86.67	98.79
5 (60)	100.00	100.00	100.00
5 (80)	100.00	100.00	100.00
7 (40)	98.89	93.33	99.39
7 (60)	98.89	93.33	99.39
7 (60)	100.00	100.00	100.00
AVERAGE	**99.31**	**96.01**	**99.62**

CONCLUSION

To conclude, all crude oil samples in three different temperatures have their unique odour profile pattern. Thus, it is proven that crude oil has different properties that can be sensed by using E-nose. CBR approach shows a positive accuracy result of 99.31% on average, nearly 100% for classification performance rate.

Other than that, the more samples used and analysed are better as the CBR case library can contain a much higher number of cases than any new case. In this project, there are seven samples in total. However, only four out of seven samples were thoroughly analysed. It cannot validate if the results of four analysed samples can represent the complete seven samples if they are to be analysed.

ACKNOWLEDGEMENT

This project research is supported by Bio-Aromatic Centre of Excellent, Universiti Malaysia Pahang (UMP) Pekan.

REFERENCES

[1] I. Statistics, "Technically Recoverable Shale Oil and Shale Gas Resources: Other Western Europe," no. September, 2015.

[2] K. A. Rahim and A. Liwan, "Oil and gas trends and implications in Malaysia," *Energy Policy*, vol. 50, pp. 262–271, 2012, doi: 10.1016/j.enpol.2012.07.013.

[3] Petroliam Nasional Berhad (PETRONAS) (20076-K), “PETRONAS 2018-2020 Activity Outlook,” vol. 1, 2020.

[4] H. WANG *et al.*, "Assessment of global unconventional oil and gas resources," *Pet. Explor. Dev.*, vol. 43, no. 6, pp. 925–940, 2016, doi: 10.1016/S1876-3804(16)30111-2.

[5] P. Yatim, M. N. Mamat, S. H. Mohamad-Zailani, and S. Ramlee, "Energy policy shifts towards sustainable energy future for Malaysia," *Clean Technol. Environ. Policy*, vol. 18, no. 6, pp. 1685–1695, 2016, doi: 10.1007/s10098-016-1151-x.

[6] C. Zou *et al.*, "Formation, distribution, potential and prediction of global conventional and unconventional hydrocarbon resources," *Pet. Explor. Dev.*, vol. 42, no. 1, pp. 14–28, 2015, doi: 10.1016/S1876-3804(15)60002-7.

[7] BP, "BP Energy Outlook Brasil." 2017, [Online]. Available: http://www.bp.com/content/dam/bp-country/pt_br/PDFs/bp-energy-outlook-2017-country-insight-brazil_portugues.pdf.

[8] S. Tongsopit, N. Kittner, Y. Chang, A. Aksornkij, and W. Wangjiraniran, "Energy security in ASEAN: A quantitative approach for sustainable energy policy," *Energy Policy*, vol. 90, pp. 60–72, 2016, doi: 10.1016/j.enpol.2015.11.019.

[9] X. Shi, H. M. P. Variam, and Y. Shen, "Trans-ASEAN gas pipeline and ASEAN gas market integration: Insights from a scenario analysis," *Energy Policy*, vol. 132, no. April, pp. 83–95, 2019, doi: 10.1016/j.enpol.2019.05.025.

[10] D. K. H. Tang, F. Leiliabadi, E. U. Olugu, and S. Z. binti Md Dawal, "Factors affecting safety of processes in the Malaysian oil and gas industry," *Saf. Sci.*, vol. 92, pp. 44–52, 2017, doi: 10.1016/j.ssci.2016.09.017.

[11] V. C. Pezoulas, T. P. Exarchos, and D. I. Fotiadis, *Machine learning and data analytics*. 2020.

[12] J. Novakovic, A. Veljovi, S. Iiic, Z. Papic, and M. Tomovic, "Evaluation of Classification Models in Machine Learning," *Theory Appl. Math. Comput. Sci.*, vol. 7, no. 1, pp. 39–46, 2017, [Online]. Available: https://uav.ro/applications/se/journal/index.php/TAMCS/article/view/158.