VOL. 11, NO. 10, MAY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

CLASSIFICATION OF HONEY ODOR-PROFILE USING CASE-BASED REASONING TECHNIQUE (CBR)

Nurdiyana Zahed, Muhammad Sharfi Najib, Nik Fatin Nadiah Nik Mohd Azhani
Fakulti Kejuruteraan Elektrik and Elektronik, Universiti Malaysia Pahang, Pekan, Malaysia
E-Mail: nurdiyanazahed92@gmail.com

ABSTRACT

Honey is one a well-known healthy natural products. By image visualization, honeys look quite similar between each other. Hence an electronic device is one of significant instrument that can be employed for detection namely an E-Nose. E-Nose comprises of an array of electronic chemical sensors that capable of recognizing odor profiles. This device is integrated by microcontroller, software and hardware design. The datasets that had been collected in this work were normalized and analyzed using boxplot for feature extraction. The unique features found from the feature extraction were used in Case-Based Reasoning (CBR) as attributes. The result shows that the classification rate of CBR is 100% in accuracy, specificity and sensitivity. These results conclude that different types of honey were able to be classified based on odor profile employed with CBR.

Keywords: honey, odor-profile classification, E-nose, case-based reasoning.

INTRODUCTION

Honey is the natural sweet substance the have been used as the food, medicine and also as the food preservative which can differentiate it type from its botanical origin [1]. It had been produce by honey bees contain execrates from excretion produce of plant which are nectar or blossom honeys or produce from insect called Hemiptera that result to the honeydew honeys [2]. Honey bees that produce honey are social insects works in a whole colony and divide their work according to their age, in protect their honeycomb, while others done their activity to search food resources, and they need critical information to done their work mainly about super organism [3]. In an investigation involve thirteen different samples of honey, the result achieve is forty-six odor of compounds is detected from the honey. The aroma profile for most honey is show quite similar and most common is aroma of flora, honey-like and green [4]. It is hard to classify the honey without used any instrument to calibrate it to show it types.

Electronic nose is provided as an instrument be use to check the data of classification of the odor of the honey. It is an instrument to comprises an array of electronic chemical sensors with specifically and correct pattern-recognition system [5][6]. The data from Electronic Nose than is analysed by using Case-Based Reasoning method.

Case-Based reasoning is the aim of method to complete the task of the classification of odor of two chosen sample of honey. Intelligent Classification using Case Based Reasoning method is introduced by Aamodt and Plaza [7]. CBR is an approach to solve problem by use the cases and experiences from previous investigation which look quite similar to the current case. In CBR system, the data and knowledge from previous case are use as the stored case [8]. The stage of elaborate, retrieve, reuse and retain from the data that achieve is apply to classify the odor of the things in CBR system [5][9][10]. It is applied when the data is recorded from Electronic Nose software. The data achieve for five times observation for each sample is exceed to 4000 data. Because of the number of data is too large, all the data is normalized before find its similarity, accuracy and sensitivity compare from result of both sample of honey. The average of data is calculated. By applying a few formulas, the data is inserting in a template of CBR to test it similarity. The expected result from this similarity test is, four other data from same sample of current case is show highest similarity compare to 5 five data observation from another sample. The calculation than proceed to find accuracy and specificity. Good accuracy is when the value is show 1 (all true as expected).

ODOR PROFILE CLASSIFICATION OF HONEY

E-Nose

For the classification of honey, method based on odor profile using E-Nose has been applied. E-Nose system consists of hardware structures that include chamber, printed circuit board (PCB) and Arduino circuit. The completed E-Nose chamber unit has been employed with Arduino software meant for data acquisition and data logging. Figure below show the free body diagram of E-Nose:

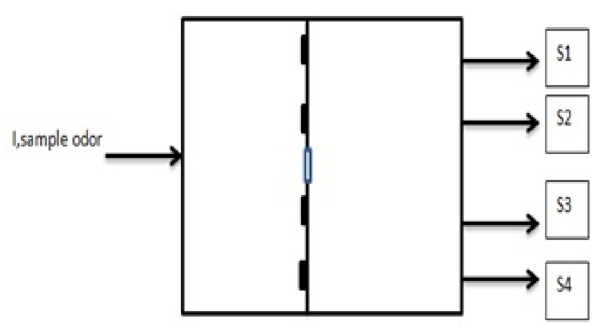

Figure-1. Block diagram of e-nose system.

VOL. 11, NO. 10, MAY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

The input of the system that represented as I is the odor-profile of the tested sample of honey. The output is the reading for four sensors, S1, S2, S3 and S4. The array of sensors in Electronic Nose is to obtain the olfactory fingerprints of the compound test to the electronic device system. The sensors (S1, S2, S3 and S4) were used for detection having different sensitivity for different volatile compound. Two types of honey (pure honey and commercial honey) tested samples were measured using E-Nose. Each sample was measured in a test vial in amount of 5 ml. After E-Nose is properly set up in well-controlled condition, Arduino software was deployed and the reading output from Arduino was displayed in a set of sensors array. There were 800 data (200 column X 4 row) collected within 2 minutes and 5 minutes interval measurement without samples. The row represents the set of number data collected and column of table represent the reading for each sensors that detect the odor profile measurement. The data was collected 5 times for each sample and the total data result achieved altogether with 8000 data.

CASE BASED REASONING FOR SOLVING ODOR PROFILE OF DIFFERENT HONEY

a) CBR computation

Case-Based Reasoning computation is a technique use by comparing the current case with stored case. One of the store case is selected to be the current case and first of the stored case is the same value as the current case and similarity should be 100% shown at CBR computation template. The data case get from the normalize value for all the five observation for each samples. The step is repeated by changing the current case from data of different sample. The expected result is when sample from Pure Honey is selected to be current case, the similarity of all store case from this sample time is higher than another sample from commercial honey type and vice versa.

b) CBR performance

Case-Based Reasoning performance is the technique to measure CBR classifier. It is use to find the classification of accuracy, sensibility, sensitivity and performance of the data result. Data result of similarity from CBR computation after each store case is selected to be current case is copy to CBR performance template. A voting technique using three highest similarity value from each case is selected and assigned as highest vote (k=1), medium vote (k=2) and lowest vote (k=3). After the voting is selected and applied for all the cases in the CBR case library, it is observed weather the value selected is from the same sample type with the case. Then, find accuracy, sensitivity and performance from data result from CBR performance template by applying formula below:

Accuracy = (true case/sum of case) x100%
Sensitivity = (true sample 1 case / (true sample 1 case + false sample 2 case)) x100%

Specificity = (true sample 2 case / (false sample 1 case + true sample 2 case)) x 100%.

RESULT & DISCUSSION

Table-1 and Table-2 shows the raw data collection from Arduino software once the sample of honey is tested in E-Nose system.

Table-1. Raw data of pure honey.

No. of measurement	Sensor Reading(Ω)			
	1	2	3	4
DM_1	795	531	410	931
DM_2	796	533	411	927
DM_3	797	534	411	932
DM_4	797	535	411	932
DM_5	797	537	411	931
DM_6	797	537	412	931
DM_7	796	538	412	932
DM_8	798	538	411	931
DM_9	797	539	413	932
.	.	.	.	.
.	.	.	.	.
DM_{2000}	809	559	412	932

Table-2. Raw data of commercial honey.

No. of measurement	Sensor Reading (Ω)			
	1	2	3	4
DM_1	712	470	421	931
DM_2	711	470	422	931
DM_3	709	470	421	932
DM_4	709	470	421	931
DM_5	711	471	421	932
DM_6	712	470	421	931
DM_7	715	471	421	927
DM_8	718	471	422	932
DM_9	717	471	421	931
.	.	.	.	.
.	.	.	.	.
DM_{2000}	737	482	421	932

Based on graph in Figure-2 until Figure-6, it shows the comparison data for pure and commercial honey. The different in all the graphs is at data in sensor 1 and sensor 2. The changes in data measurement show the ability of both sensors to detect the odor of the sample. Value of variable resistor that present from sensor 3 and sensor 4 is unchanged before and after sample is test in E-Nose system. Even though sensor 3 and sensor 4 is not effective to detect the odor of honey, it is useful to those sensors to make the E-Nose system multi-purpose to detect the odor of other type of sample.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

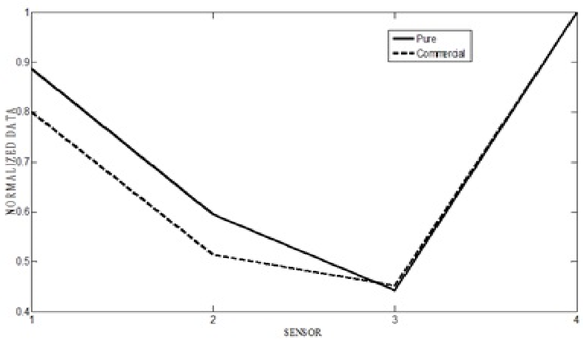

Figure-2. Graph normalized data vs sensor for 1st time.

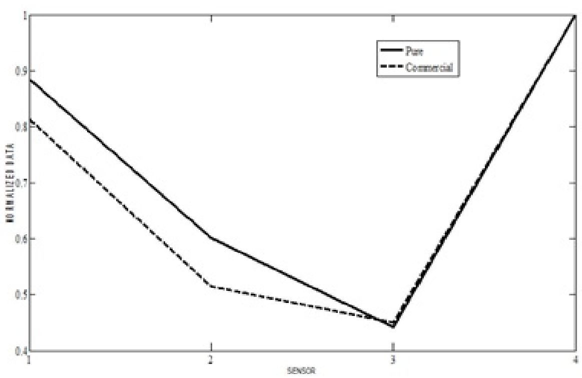

Figure-3. Graph normalized data vs sensor for 2nd time.

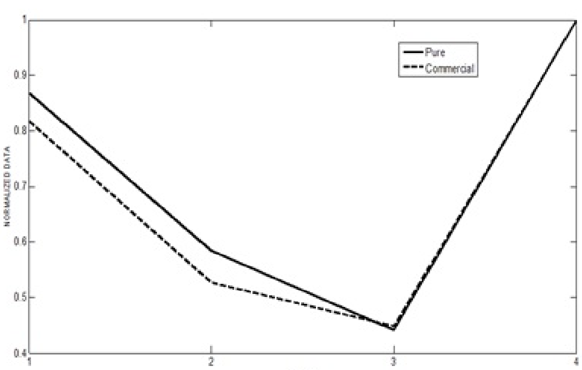

Figure-4. Graph normalized data vs sensor for 3rd time.

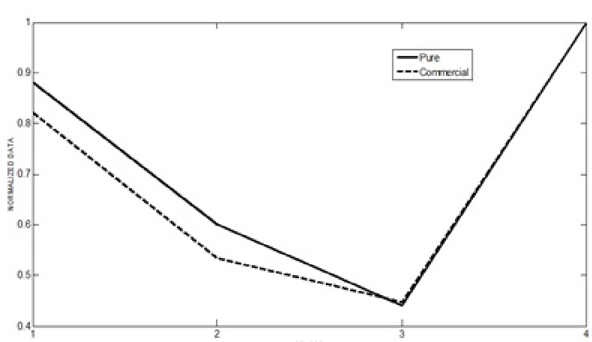

Figure-5. Graph normalized data vs sensor for 4th time.

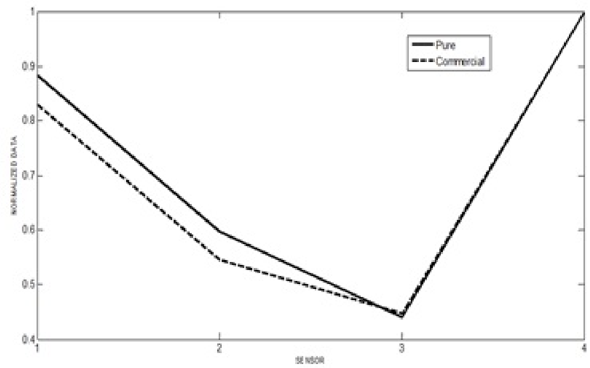

Figure-6. Graph normalized data vs sensor for 5th time.

By referring to Figure-7 and Figure-8, it has shown the boxplot for honey sample. In both boxplot present, sensor 4 shows the best data result because it have same value of mean and did not have other element of range variation. It means that the data show constant value. The line of lower part of boxplot represent in sensor 1 until sensor 3 show the minimum value while the horizontal line at the upper part show the maximum value. Figure-8, which is the boxplot for commercial honey show more clear data because of less repeated value of the datasets. The blue rectangular boxes indicate the interquartile range in which lower line is Q1 and upper line is Q3.

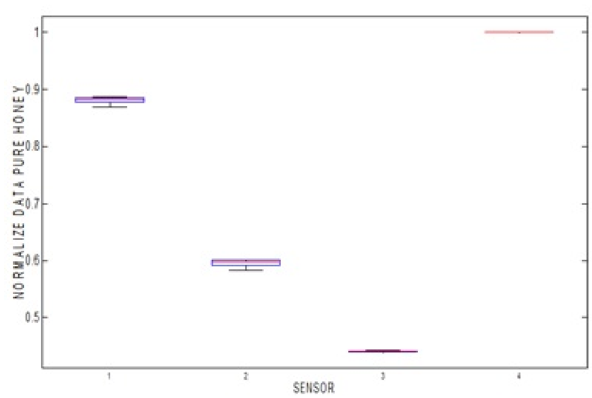

Figure-7. Boxplot for pure honey.

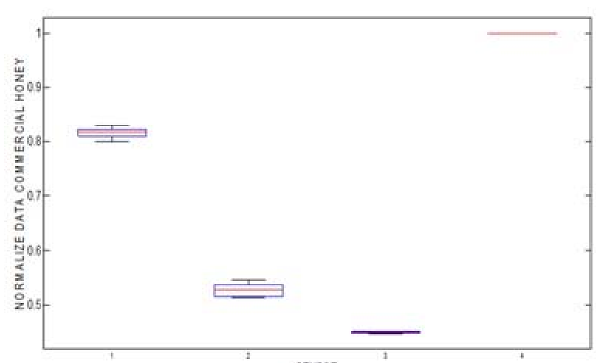

Figure-8. Boxplot for commercial honey.

Table-3 shows the CBR Computation result for percentage similarity when commercial honey from 5th observation was set as the current case. The data in percentage was transferred to the table of CBR performance method. To find the result in CBR computation, the first data from 1st time observation of sample (commercial honey) was store at current case column. After that, another 9 observations throughout the overall samples (4 observations from commercial honey and 5 observations from pure honey) were stored to the store case. Value of percentage similarity is shown and the similarity of current case and stored case is compared.

Table-3. CBR computation result.

Stored Cased	Percentage Similarity (%)
Pure Honey 1st Observation	97.2
Pure Honey 2nd Observation	97.1
Pure Honey 3rdObservation	97.9
Pure Honey 4thObservation	97.2
Pure Honey 5th Observation	97.2
Commercial Honey 1st Observation	98.3
Commercial Honey 2nd Observation	98.7
Commercial Honey 3rdObservation	99.2
Commercial Honey 4thObservation	99.5
Commercial Honey 5th Observation	100.0

In CBR computation, the process of comparing the data from current case and stored case was done and the result is shown in percentage similarity as depicted in Table-3. The result from Table -3 was transferred to CBR performance table to compute the result for odor classification by using voting (k=1, k=2, k=3). The result analysis was finalized by computing the accuracy, sensitivity and specificity of odor classification of both honey samples.

Table-4 indicate the performance evaluation of CBR that has successfully shown 100% accuracy, sensitivity and specificity in classifying odor profile of honey pure honey (PH) and commercial honey (CH) samples for all voting ranking (k=1, k=2 and k=3).

Table-4. Performance evaluation of CBR classification for FF and FPF.

Performance Evaluation	k=1	k=2	k=3
Criteria/ Indices	*Values*	*Values*	*Values*
Total Cases	10	10	10
PH case (P)	5	5	5
CH case (N)	5	5	5
True positive (TP)	5	5	5
False positive (FP)	0	0	0
True Negative (TN)	5	5	5
False negative (FN)	0	0	0
Sensitivity=TP/(TP+FN)	1.0	1.0	1.0
Specificity=TN/(FP+TN)	1.0	1.0	1.0
Accuracy=(TP+TN)/(P+N)	1.0	1.0	1.0

CONCLUSIONS

This paper is successfully presented the odor classification of two type of honey using case-based reasoning approach. It is shown that the odor of pure honey and commercial honey can be classified with 100% accuracy, sensitivity and specificity. The odor of sample honey is trace by sensors array in E-Nose system. Classification of odor profile of pure honey and commercial honey is undergo normalization technique, which is a technique use to simplify the large range of data to small range, as use in this project is from range 0 to range 1. The normalization data is representing in graphs and statistical method using boxplot. Normalized data is transfer to CBR performance and CBR computation table and shown the classification result. This project can be extended by adding extra statistical analysis to analyze the results. As CBR technique is better with the increase of number of samples, this project can be improved by adding more honey samples from different types of honey. For the purpose of comparative study, CBR technique performance measure can be cross-validated using several classifier techniques.

ACKNOWLEDGEMENTS

I would like to acknowledge to my supervisor, Dr Muhammad Sharfi bin Najib, the staff of Faculty of Electrical and Electronic Engineering of University Malaysia Pahang for providing the physical necessities of completing the project as well as generous access to unlimited support, either internal or external resources. It is an honour for me to thank the educators and professors for all the theoretical and life lessons learnt this whole time.

REFERENCES

[1] S. Shafiee, S. Minaei, N. Moghaddam-Charkari, and M. Barzegar. 2014. Honey characterization using computer vision system and artificial neural networks. Food Chem. pp. 143–150.

[2] K. Karabagias, A. V. Badeka, S. Kontakos, S. Karabournioti, and M. G. Kontominas. 2014. Botanical discrimination of Greek unifloral honeys with physico-chemical and chemometric analyses. Food Chem. pp. 181–190,.

[3] C. Chen, E. C. Yang, J. A. Jiang, and T. Te Lin. 2012. An imaging system for monitoring the in-and-out activity of honey bees. Comput. Electron. Agric.pp. 100–109.

[4] S. Seisonen, E. Kivima, and K. Vene. 2015. Characterisation of the aroma profiles of different honeys and corresponding flowers using solid-phase microextraction and gas chromatography-mass spectrometry/olfactometry. Food Chem. pp. 34–40,

ARPN Journal of Engineering and Applied Sciences

[5] M. S. Najib, M. N. Taib, N. A. M. Ali, M. N. M. Arip, and A. M. Jalil. 2011. Classification of agarwood grades using ANN. InECCE 2011 - International Conference on Electrical, Control and Computer Engineering. pp. 367–372.

[6] M. S. Najib, N. A. M. Ali, M. N. M. Arip, A. M. Jalil, and M. N. Taib. 2010. Classification of Agarwood region using ANN. Proceedings - ICSGRC 2010: IEEE Control and System Graduate Research Colloquium. pp. 7–13.

A. Aamodt. 1994. Case-Based Reasoning : Foundational Issues. Methodological Variations , and System Approaches. pp. 39–59.

[7] G.-N. Zhu, J. Hu, J. Qi, J. Ma, and Y.-H. Peng. 2015. An integrated feature selection and cluster analysis techniques for case-based reasoning. Eng. Appl. Artif. Intell. pp. 14–22.

[8] F. A. Halim, M. S. Najib, K. H. Ghazali, and M. F. Zahari. 2015. ISERD - 2nd International Conference on Modern Trends in Engineering and Technology (ICMTET).

[9] M. F. Zahari, J. Thommy Anak, F. A. Halim, M. S. Najib, K. H. Ghazali, and A. M. zoddein. 2015. Intelligent Classification Hazardous Gas Using Sensor Array. J. Adv. Inf. Technoogy. pp. 233–237.